

Crone

Betriebswirtschaftliche Forschung
zur UnternehmensführungHrsg.: Karl-Werner Hansmann, Manfred Layer,
Dietmar Preßmar und Kai-Ingo Voigt

Neuronale Netze zur Prognose und Disposition im Handel

Künstliche Neuronale Netze bieten ein erhebliches Potential zur Verbesserung der Prognose sowie der Disposition eines Handelsbetriebs. Sven F. Crone bietet eine fundierte Analyse der Grundlagen der Absatzprognose, Waren-disposition und der Verfahrensklasse der Neuronalen Netze. Diese erweitert er mithilfe asymmetrischer Zielfunktionen der Entscheidungskosten zu einem Simultanansatz der Vorhersage kostenminimaler Warenbestände, die für den robusten Praxiseinsatz geeignet ist.

Die Forschungsarbeit wurde mit den Wissenschaftspreisen der Universität Hamburg 2008 und der Gesellschaft für Operations Research (GOR) 2009 ausgezeichnet.

Der Inhalt

- Prognose und Disposition im Handel
- Neuronale Netze als Instrument der Prognose
- Empirische Studie von Neuronalen Netzen zur Prognose und Disposition

Die Zielgruppen

Dozenten und Studenten der Betriebswirtschaftslehre mit den Schwerpunkten Operational Research, Wirtschaftsinformatik und Logistik; Praktiker aus Waren- und Materialwirtschaft sowie Supply Chain Management in Industrie und Handel.

Der Autor

Dr. Sven F. Crone promovierte bei Prof. Dr. Dr. h. c. Preßmar am Institut für Wirtschaftsinformatik der Universität Hamburg und ist aktuell Assistant Professor in Management Science an der Lancaster University und Deputy Director des Lancaster Centre for Forecasting in England.

Betriebswirtschaftliche Forschung
zur Unternehmensführung

www.gabler.de

ISBN 978-3-8349-1174-2



RESEARCH

Neuronale Netze zur Prognose
und Disposition im Handel

Sven F. Crone

Neuronale Netze zur Prognose und Disposition im Handel

Sven F. Crone

Neuronale Netze zur Prognose und Disposition im Handel

GABLER RESEARCH

Betriebswirtschaftliche Forschung zur Unternehmensführung

Herausgegeben von

Prof. Dr. Dr. h.c. Herbert Jacob (†), Universität Hamburg

Prof. Dr. Karl-Werner Hansmann, Universität Hamburg

Prof. Dr. Manfred Layer, Universität Hamburg

Prof. Dr. Dr. h.c. Dieter B. Preßmar, Universität Hamburg

Prof. Dr. Kai-Ingo Voigt, Universität Erlangen-Nürnberg

Sven F. Crone

Neuronale Netze zur Prognose und Disposition im Handel

Mit einem Geleitwort von
Prof. Dr. Dr. h. c. Dieter B. Preßmar



RESEARCH

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über
<<http://dnb.d-nb.de>> abrufbar.

Dissertation Universität Hamburg, 2008

1. Auflage 2010

Alle Rechte vorbehalten

© Gabler Verlag | Springer Fachmedien Wiesbaden GmbH 2010

Lektorat: Ute Wrasmann | Sabine Schöller

Gabler Verlag ist eine Marke von Springer Fachmedien.

Springer Fachmedien ist Teil der Fachverlagsgruppe Springer Science+Business Media.

www.gabler.de



Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Umschlaggestaltung: KünkelLopka Medienentwicklung, Heidelberg

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier

Printed in Germany

ISBN 978-3-8349-1174-2

Geleitwort

Das zentrale Gestaltungsproblem eines Handelsbetriebs besteht darin, seine Lieferfähigkeit zu jedem künftigen Zeitpunkt zu garantieren. Dazu ist sowohl eine verlässliche Prognose der Absatzmengen für das Tagesgeschäft notwendig als auch die Bestimmung einer kostenoptimalen Lagerhaltungs- und Bestellpolitik. Für jedes dieser genannten Teilprobleme existieren mathematische Modelle, die im Falle einer praktischen Anwendung zu einer abgestimmten Lösung zusammengeführt werden müssen. Wegen der bekannten Interdependenzprobleme lässt sich diese Planungsaufgabe insbesondere für die praktische Unternehmensführung nur unter großen Schwierigkeiten zufriedenstellend lösen.

Hier zeigt die vorliegende Arbeit neue Wege zu einer bemerkenswerten Problemlösung auf. Maßgebend für die vorgeschlagene Methodenwahl ist der Planungserfolg in der empirischen Konfrontation mit der betrieblichen Praxis. Der Verfasser erweist sich in seiner Analyse der Planungsprobleme als Insider des Handelsgeschäfts; deshalb ist die in seiner Arbeit präsentierte Praxisnähe vorbildlich und aner kennenswert. Vor diesem Hintergrund kann er mit großer Überzeugungskraft zeigen, dass sein Methoden- und Verfahrensvorschlag für den robusten Praxiseinsatz geeignet ist und zugleich den Anforderungen an die Optimalqualität der einschlägigen betriebswirtschaftlichen Theorie genügt.

Die in der vorliegenden Forschungsarbeit entwickelte Alternative zu den bekannten varianzanalytischen Prognoseverfahren und den entsprechenden Lagerhaltungsmodellen ergibt sich aus der Methodenklasse der künstlichen neuronalen Netze. Kennzeichen dieser Methoden ist das umfassende Lern- und Trainingsverfahren, das sich auch an heterogene und komplexe Planungssituationen anpassen kann. In der gegebenen Fragestellung beschreitet der Verfasser neuartige Wege, indem er das gemeinsame Problem von Prognose und Disposition in der Warenwirtschaft auf der Grundlage eines Simultanansatzes mithilfe der neuronalen Netze löst. Dazu werden die bekannten Algorithmen der neuronalen Netze durch asymmetrische lineare und quadratische Bewertungsfunktionen zur Steuerung des Lernprozesses erweitert. Mit diesem Ansatz gelingt es, das Prognoseproblem und die Dispositionsaufgabe miteinander zu verknüpfen und somit simultan zu lösen. Die dabei erzielte Lösungsqualität wird anhand von zahlreichen Datenkonstellationen aus der Praxis überprüft und mit überzeugendem Erfolg nachgewiesen.

Die Arbeit von Sven Crone repräsentiert einen herausragenden und innovativen Beitrag zur Theorie der künstlichen neuronalen Netzen; zugleich wird ein umfassendes Lösungsverfahren für ein zentrales Planungsproblem des Handels aufgezeigt. In Würdigung dieser besonderen Leistung wurde die Forschungsarbeit bereits durch zwei Wissenschaftspreise ausgezeichnet. Ich wünsche dem Buch einen großen und fachlich engagierten Leserkreis.

Prof. Dr. Dr. h. c. Dieter B. Preßmar

Vorwort

Die vorliegende Untersuchung entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Wirtschaftsinformatik der Universität Hamburg unter der Leitung von Prof. Dr. Dr. h.c. Dieter Preßmar und Prof. Dr. Stefan Voß. Die Arbeit wurde am Department of Management Science der Universität Lancaster fertig gestellt und vom Fachbereich Wirtschaftswissenschaften als Inauguraldissertation zur Erlangung des akademischen Grades eines Doktors der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol.) angenommen.

Mein besonderer Dank gilt meinem akademischem Lehrer Herrn Prof. Dr. Dr. h.c. Dieter B. Preßmar für die engagierte Betreuung und Förderung der Arbeit, seine stete Unterstützung und die freundschaftliche Zusammenarbeit. Die Entwicklung des Themas, im Spannungsfeld von wissenschaftlicher Forschung und den Problemstellungen der Praxis, wäre nicht ohne seine Faszination an künstlichen Neuronalen Netzen und dem nachhaltigen Bemühen um praktische, anwendungsorientierte Forschung in Projekten möglich gewesen.

Bei Herrn Prof. Karl-Werner Hansmann möchte ich mich für die Übernahme des Koreferats und bei Herrn Prof. Stefan Voß für die Übernahme des Vorsitzes des Prüfungsausschusses bedanken. Mein weiterer Dank gilt meinen Kollegen, den Mitarbeitern und ehemaligen Mitarbeitern des Lehrstuhls. Ferner danke ich den Juroren zur Verleihung des Prof. Herbert Jacob Preises der Universität Hamburg für das Jahr 2008 sowie den Juroren des nationalen Dissertationspreises der Gesellschaft für Operations Research (GOR e.V.) für das Jahr 2009.

Mein besonderer Dank gilt meiner Familie, ohne deren Unterstützung diese Arbeit niemals fertig gestellt worden wäre und deren Rückhalt wesentlich zum Gelingen dieser Arbeit beigetragen hat. Ebenso bedanke ich mich herzlich bei meiner Freundin für ihr Verständnis, Durchhaltevermögen und moralische Unterstützung während der Fertigstellung der Arbeit.

Sven F. Crone

Inhaltsübersicht

Inhaltsverzeichnis.....	XI
Darstellungsverzeichnisse.....	XVII
Sonstige Verzeichnisse.....	XXIX
1 Einleitung.....	1
1.1 Problemstellung.....	1
1.2 Gang der Untersuchung.....	6
2 Prognose und Disposition im Handel.....	9
2.1 Grundlagen der Warendisposition.....	9
2.2 Modelle der Bedarfsrechnung.....	22
2.3 Modelle der Bestellmengenrechnung.....	32
2.4 Beurteilung von Dispositionsverfahren.....	62
3 Methodische Grundlagen der Prognose.....	67
3.1 Determinanten der Prognosemodelle.....	67
3.2 Berechnung von Prognoseverfahren.....	94
3.3 Beurteilung von Prognoseverfahren.....	128
3.4 Anwendung von Prognoseverfahren.....	142
4 Neuronale Netze als Instrument der Prognose.....	159
4.1 Methodische Grundlagen von Neuronalen Netzen.....	159
4.2 Aufbau und Funktionsweise von Neuronalen Netzen.....	164
4.3 Parametrisierung von Neuronalen Netzen.....	189
4.4 Anwendung von Neuronalen Netzen zur Prognose.....	207
5 Neuronale Netze zur Disposition.....	239
5.1 Lernziele und Zielfunktionen von Neuronalen Netzen.....	239
5.2 Zielfunktionen der statistischen Prognosefehler.....	245
5.3 Zielfunktionen der betriebswirtschaftlichen Entscheidungskosten.....	261
5.4 Studie zur Einsatzfähigkeit von asymmetrischen Kostenfunktionen.....	282
6 Empirischer Vergleich von Neuronalen Netzen zur Prognose und Disposition.....	313
6.1 Gestaltung der Vergleichsstudie.....	313
6.2 Durchführung der Vergleichsstudie.....	318
6.3 Ergebnisse der Vergleichsstudie.....	358
7 Schlussbetrachtung.....	405
7.1 Zusammenfassung.....	405
7.2 Ausblick.....	408
Literaturverzeichnis.....	409
Anhang.....	455

Inhaltsverzeichnis

Abbildungsverzeichnis	XVII
Tabellenverzeichnis	XXV
Abkürzungsverzeichnis.....	XXIX
Symbolverzeichnis	XXXI
1 Einleitung	1
1.1 Problemstellung	1
1.2 Gang der Untersuchung	6
2 Prognose und Disposition im Handel.....	9
2.1 Grundlagen der Warendisposition	9
2.1.1 Definition und Einordnung der Warendisposition.....	9
2.1.2 Warendisposition als Kernprozess des Handels	11
2.1.3 Dispositions- und Lagerstrukturen im Handel.....	14
2.1.4 Funktionsmodell der Warendisposition	19
2.2 Modelle der Bedarfsrechnung	22
2.2.1 Determinanten	22
2.2.1.1 Kennzeichnung des Bedarfs	22
2.2.1.2 Modellauswahl.....	24
2.2.2 Deterministische Modelle.....	26
2.2.3 Stochastische Modelle	28
2.2.3.1 Modelle für stationären Bedarf.....	28
2.2.3.2 Modelle für instationären Bedarf.....	30
2.3 Modelle der Bestellmengenrechnung	32
2.3.1 Determinanten	32
2.3.1.1 Entscheidungsvariablen	32
2.3.1.2 Kostenparameter	34
2.3.1.3 Modellauswahl.....	39
2.3.2 Deterministische Modelle.....	42
2.3.2.1 Modelle für stationär-deterministischen Bedarf	42
2.3.2.2 Modelle für instationär-deterministischen Bedarf.....	46
2.3.3 Stochastische Modelle	47
2.3.3.1 Modelle für stationär-stochastischen Bedarf	47
2.3.3.1.1 Einperiodige stationär-stochastische Modelle.....	47
2.3.3.1.2 Mehrperiodige stationär-stochastische Modelle.....	54
2.3.3.2 Modelle für instationär-stochastischen Bedarf	59
2.4 Beurteilung von Dispositionsverfahren	62
2.4.1 Deskriptive Maße der Dispositionsgüte	62
2.4.2 Zusammenhang von Prognosefehler und Dispositionsgüte.....	65
3 Methodische Grundlagen der Prognose	67
3.1 Determinanten der Prognosemodelle.....	67
3.1.1 Merkmale der Prognose.....	67
3.1.2 Determinanten von Prognosemodellen.....	69
3.1.2.1 Gegenstand der Prognose	69
3.1.2.2 Formalisierungsgrad der Modellbildung	72
3.1.2.3 Gestaltung der abhängigen Modellvariablen	74

3.1.2.4	Gestaltung der unabhängigen Modellvariablen.....	77
3.1.3	Auswahl von Prognosemodell und Prognoseverfahren.....	82
3.1.3.1	Umfang des Erklärungsmodells.....	82
3.1.3.2	Zeitlicher Prognosehorizont.....	85
3.1.3.3	Verfahren der Bedarfsprognose.....	88
3.1.3.4	Prognoseverfahren im Handel.....	90
3.2	Berechnung von Prognoseverfahren.....	94
3.2.1	Subjektive Prognoseverfahren.....	94
3.2.2	Objektive Prognoseverfahren der Verteilungsschätzung.....	96
3.2.2.1	Bestimmung der empirischen Verteilungsfunktion.....	96
3.2.2.2	Approximation durch theoretische Verteilungsfunktionen.....	97
3.2.2.3	Schätzung der Verteilungsparameter.....	100
3.2.3	Objektive Prognoseverfahren der Zeitreihenanalyse.....	103
3.2.3.1	Zeitreihen als Grundlage der Prognose.....	103
3.2.3.2	Verfahren bei konstantem Bedarfsniveau.....	111
3.2.3.2.1	Naive Verfahren und Mittelwerte.....	111
3.2.3.2.2	Mittelwerte.....	112
3.2.3.2.3	Einfache Exponentielle Glättung.....	113
3.2.3.2.4	ARMA-Modelle.....	115
3.2.3.3	Verfahren bei trendbehaftetem Bedarfsverlauf.....	118
3.2.3.3.1	Naives Trendverfahren.....	118
3.2.3.3.2	Doppelte Exponentielle Glättung.....	118
3.2.3.3.3	Lineare Exponentielle Glättung.....	119
3.2.3.3.4	Exponentielle Glättung mit gedämpftem Trend.....	120
3.2.3.3.5	ARIMA-Verfahren.....	121
3.2.3.4	Verfahren bei saisonalem und trend-saisonaltem Bedarfsverlauf.....	122
3.2.3.4.1	Naives Saisonales Verfahren.....	122
3.2.3.4.2	Saisonale Exponentielle Glättung.....	122
3.2.3.4.3	Saisonale ARIMA-Verfahren.....	124
3.2.4	Objektive Prognoseverfahren der Kausalanalyse.....	125
3.2.4.1	Dynamische Regression.....	125
3.2.4.2	Verfahren der Interventionsmodellierung.....	126
3.3	Beurteilung von Prognoseverfahren.....	128
3.3.1	Wirtschaftlichkeit und Prognosegenauigkeit.....	128
3.3.2	Beurteilung der Prognosegenauigkeit.....	129
3.3.2.1	Aspekte der Beurteilung.....	129
3.3.2.2	Statistische Prognosefehler.....	131
3.3.2.3	Statistische Fehlermaße.....	133
3.3.2.4	Komparative statistische Fehlermaße.....	137
3.3.3	Durchführung von empirischen Vergleichsstudien.....	139
3.3.3.1	Erkenntnisse aus Vergleichsstudien.....	139
3.3.3.2	Durchführung von Vergleichsstudien.....	140
3.4	Anwendung von Prognoseverfahren.....	142
3.4.1	Methodologien der Prognoseprozesse.....	142
3.4.2	Phasen des Prognoseprozesses.....	144
3.4.2.1	Problemformulierung.....	144
3.4.2.2	Informationsbereitstellung.....	145
3.4.2.2.1	Identifikation und Sammlung der Daten.....	145
3.4.2.2.2	Datenvorverarbeitung.....	147
3.4.2.3	Auswahl der Verfahren.....	149
3.4.2.3.1	Explorative Datenanalyse.....	149

3.4.2.3.2	Modell- und Verfahrensselektion.....	151
3.4.2.3.3	Verfahrensspezifische Datenvorverarbeitung.....	153
3.4.2.4	Durchführung der Verfahren.....	155
3.4.2.5	Beurteilung der Verfahren.....	156
3.4.2.6	Anwendung der Verfahren.....	156
3.4.3	Fehlerquellen im Prognoseprozess.....	157
4	Neuronale Netze als Instrument der Prognose.....	159
4.1	Methodische Grundlagen von Neuronalen Netzen.....	159
4.1.1	Definition und Abgrenzung.....	159
4.1.2	Motivation zur Analyse von Neuronalen Netzen.....	161
4.1.3	Historische Entwicklung.....	162
4.2	Aufbau und Funktionsweise von Neuronalen Netzen.....	164
4.2.1	Biologisches Vorbild.....	164
4.2.2	Technische Realisierung.....	167
4.2.3	Signalverarbeitung in künstlichen Neuronen.....	169
4.2.3.1	Eingabefunktionen.....	169
4.2.3.2	Aktivierungsfunktionen.....	170
4.2.3.3	Ausgabefunktionen.....	175
4.2.3.4	Kombination der Verarbeitungsfunktionen.....	175
4.2.4	Informationsverarbeitung in Netzwerken von Neuronen.....	179
4.2.4.1	Netzwerktopologie.....	179
4.2.4.2	Struktur der Verbindungsgewichte.....	181
4.2.4.3	Aktivierungsstrategien der Informationsverarbeitung.....	183
4.2.5	Ausgewählte Architekturen.....	184
4.3	Parametrisierung von Neuronalen Netzen.....	189
4.3.1	Grundlagen des Lernens.....	189
4.3.2	Algorithmen des überwachten Lernens.....	194
4.3.2.1	Klassische Lernalgorithmen.....	194
4.3.2.2	Der Backpropagation-Lernalgorithmus.....	197
4.3.2.3	Erweiterungen des Backpropagation-Algorithmus.....	199
4.3.3	Anwendung der Lernalgorithmen.....	203
4.4	Anwendung von Neuronalen Netzen zur Prognose.....	207
4.4.1	Einsatzfähigkeit Neuronaler Netze zur Prognose.....	207
4.4.1.1	Eigenschaften von Neuronalen Netzen.....	207
4.4.1.2	Anwendungsbereiche von Neuronalen Netzen.....	211
4.4.1.3	Anwendungen von Neuronalen Netzen zur Prognose.....	216
4.4.1.4	Empirische Güte der Anwendung.....	223
4.4.2	Netzwerkarchitekturen zur Prognose.....	226
4.4.2.1	Zeitreihenanalytische Modellierung.....	226
4.4.2.2	Kausalanalytische Modellierung.....	228
4.4.2.3	Äquivalenz zu statistischen Prognoseverfahren.....	231
5	Neuronale Netze zur Disposition.....	239
5.1	Lernziele und Zielfunktionen von Neuronalen Netzen.....	239
5.2	Zielfunktionen der statistischen Prognosefehler.....	245
5.2.1	Quadratische Fehlerfunktionen.....	245
5.2.2	Nicht-quadratische Fehlerfunktionen.....	251
5.3	Zielfunktionen der betriebswirtschaftlichen Entscheidungskosten.....	261
5.3.1	Asymmetrische Kostenfunktionen in der Statistik.....	261

5.3.1.1	Erweiterung auf asymmetrische Kostenfunktionen.....	261
5.3.1.2	Lineare asymmetrische Kostenfunktionen	266
5.3.1.3	Nichtlineare asymmetrische Kostenfunktionen.....	269
5.3.1.4	Anwendungsbereiche asymmetrischer Kostenfunktionen.....	272
5.3.2	Neuronale Netze und asymmetrische Kostenfunktionen	275
5.3.2.1	Parametrisierung bei asymmetrischen Kostenfunktionen	275
5.3.2.2	Erweiterung des Backpropagation-Algorithmus	277
5.3.2.3	Auswirkungen von asymmetrischen Kostenfunktionen	280
5.4	Studie zur Einsatzfähigkeit von asymmetrischen Kostenfunktionen	282
5.4.1	Gestaltung des Versuchsaufbaus	282
5.4.1.1	Datenbasis der Studie	282
5.4.1.2	Ziel- und Kostenfunktionen.....	285
5.4.1.3	Modellierung der Verfahren	286
5.4.2	Ergebnisse.....	289
5.4.2.1	Grafische Analyse und Interpretation.....	289
5.4.2.2	Beurteilung der Einsatzfähigkeit	298
5.4.2.3	Beurteilung der Entscheidungsgüte	303
5.4.3	Zusammenfassung und Interpretation	308
6	Empirischer Vergleich von Neuronalen Netzen zur Prognose und Disposition	313
6.1	Gestaltung der Vergleichsstudie	313
6.1.1	Zielsetzung des Verfahrensvergleichs	313
6.1.2	Aufbau der Vergleichsstudie	315
6.2	Durchführung der Vergleichsstudie	318
6.2.1	Problemstellung.....	318
6.2.1.1	Der Warenautomat als Betriebsform des Handels.....	318
6.2.1.2	Prognose und Disposition an Warenautomaten.....	322
6.2.1.3	Modelltheoretische Einordnung der Problemstellung	324
6.2.2	Datenanalyse.....	327
6.2.2.1	Stichprobe der Zeitreihen	327
6.2.2.2	Explorative Datenanalyse	330
6.2.2.2.1	Datenanalyse je Abverkaufsstelle	330
6.2.2.2.2	Datenanalyse je Zeitreihe.....	336
6.2.2.2.3	Kategorisierung der Zeitreihen.....	339
6.2.2.3	Aufteilung der Datenbasis	344
6.2.3	Verfahren zur Bedarfsprognose.....	345
6.2.3.1	Auswahl der Prognoseverfahren.....	345
6.2.3.2	Berechnung der statistischen Prognoseverfahren	348
6.2.3.3	Berechnung der Neuronalen Netze zur Prognose	350
6.2.4	Verfahren zur Bestellmengenrechnung	352
6.2.4.1	Bestimmung des kostenoptimalen Servicelevels.....	352
6.2.4.2	Bestellmengenrechnung für statistische Prognoseverfahren	354
6.2.4.3	Simultane Bestellmengenrechnung mit Neuronalen Netzen	355
6.2.5	Beurteilungsmaße	356
6.2.5.1	Fehlermaße der Prognosegenauigkeit.....	356
6.2.5.2	Kostenmaße der Dispositionsgüte	357
6.3	Ergebnisse der Vergleichsstudie	358
6.3.1	Ergebnisse der Bedarfsprognose	358
6.3.1.1	Gesamtergebnisse	358
6.3.1.2	Ergebnisse nach Absatzstelle.....	369

6.3.1.3	Ergebnisse nach Zeitreihe.....	374
6.3.1.4	Ergebnisse nach Zeitreihencharakteristika	376
6.3.2	Ergebnisse der Warendisposition	381
6.3.2.1	Gesamtergebnisse	381
6.3.2.2	Ergebnisse nach Absatzstelle.....	389
6.3.2.3	Ergebnisse nach Zeitreihen.....	394
6.3.2.4	Ergebnisse nach Zeitreihencharakteristika	395
6.3.3	Zusammenhang zwischen Prognosegenauigkeit und Dispositionskosten.....	400
7	Schlussbetrachtung.....	405
7.1	Zusammenfassung	405
7.2	Ausblick.....	408
	Literaturverzeichnis.....	409
	Anhang	455

Abbildungsverzeichnis

Abbildung 2.1:	Die Warendisposition als Objekt der Logistik, Materialwirtschaft, Beschaffung, Produktion und Distribution	11
Abbildung 2.2:	Kernprozesse und Hauptfunktionen von Handelsbetrieben.....	12
Abbildung 2.3:	Konstituenten eines Lagerhaltungsproblems	14
Abbildung 2.4:	Grundformen von Lagerstrukturen einer Handelsunternehmung	17
Abbildung 2.5:	Teilprozesse der Warendisposition im Handel	20
Abbildung 2.6:	Modelle der Bedarfs- und Bestellmengenrechnung für die Determinanten des Bedarfs	24
Abbildung 2.7:	Stationäre Zeitreihen mit niedriger (a.), mittlerer (b.) und hoher Streuung (c.).....	29
Abbildung 2.8:	Beispiele instationärer Zeitreihen	30
Abbildung 2.9:	Bestellkosten, Lagerungskosten und Gesamtkosten in Abhängigkeit von der Bestellmenge.....	43
Abbildung 2.10:	Lagerbestandsentwicklung im deterministischen Modell nach HARRIS	45
Abbildung 2.11:	Lagerbestandsentwicklung bei stochastischem Bedarf (a.) und bei stochastischer Wiederbeschaffungszeit (b.).....	48
Abbildung 2.12:	Bestimmung der optimalen Bestellmenge q^* im Newsboy-Modell.....	51
Abbildung 2.13:	Bestimmung des Sicherheitsbestandes im Newsboy-Modell	52
Abbildung 2.14:	Lagerbestandsentwicklung bei einer stochastischen (s, q) -Politik	56
Abbildung 2.15:	Lagerbestandsentwicklung bei einer stochastischen (r, S) -Politik	57
Abbildung 2.16:	Lagerbestandsentwicklung bei einer stochastischen (r, s, S) -Politik.....	57
Abbildung 2.17:	Fehlerverlauf deskriptiver Gütemaße des OOU (a.), OVU (b.), OOS (c.) und OVS (d.) im Vergleich zum quadratischen Fehler $SE(e)$	63
Abbildung 3.1:	Strukturmodell der Einzelunternehmung.....	70
Abbildung 3.2:	Prognoseobjekte der Absatzprognose und ihre Interaktion	72
Abbildung 3.3:	Modellbildung im Erklärungs- und Prognosezeitraum.....	78
Abbildung 3.4:	Prognosemodelle der Zeitreihenprognose, der Wirkungsprognose und der kausalen Prognose.....	80
Abbildung 3.5:	Abgrenzung von Zeitreihenprognose und Wirkungsprognose	80
Abbildung 3.6:	Absatzprognose im Produktlebenszyklus	87
Abbildung 3.7:	Systematik objektiver Prognosemodelle und Prognoseverfahren.....	90
Abbildung 3.8:	Anwendung von Prognoseverfahren in der betrieblichen Praxis.....	91
Abbildung 3.9:	Anwendung von Prognoseverfahren im Handel	92
Abbildung 3.10:	Anwendung von Prognoseverfahren in Industrie und Handel.....	92

Abbildung 3.11: Dichtefunktionen und Verteilungsfunktionen der Normalverteilung, der Erlangverteilung und der Poissonverteilung.....	98
Abbildung 3.12: Prozesskette der Bedarfsrechnung mit Verteilungsannahme.....	103
Abbildung 3.13: Regelmäßige und unregelmäßige Komponenten von Zeitreihen.....	104
Abbildung 3.14: Zeitreihenmuster nach der PEGELS-GARDNER-Klassifikation	106
Abbildung 3.15: Instationäre Zeitreihen mit saisonalem Verlauf (a.) und trendbehaftetem Verlauf (b.).....	107
Abbildung 3.16: Instationäre Zeitreihen mit Niveauverschiebung (a.), Ausreißern (b.) und einem Strukturbruch im Trendverlauf (c.)	108
Abbildung 3.17: Dekomposition einer Zeitreihe in ihre Komponenten	110
Abbildung 3.18: Gewichtung der Beobachtungswerte in der Exponentiellen Glättung.....	115
Abbildung 3.19: Prognosemodell der dynamischen Regression zur kausalen Prognose.....	126
Abbildung 3.20: ARX(p)-Prognosemodell zur kausalen Prognose mit Interventionen	126
Abbildung 3.21: Fehlerverlauf des einfachen (a.), absoluten (b.), quadratischen (c.) und der Wurzel des quadratischen Prognosefehlers (d.).....	132
Abbildung 3.22: Kalibrations- und Teststichproben zur Beurteilung von Prognoseverfahren	141
Abbildung 3.23: Phasen des Prognoseprozesses.....	144
Abbildung 3.24: Informationsbereitstellung im Prognoseprozess	146
Abbildung 3.25: Verfahrens unabhängige Datenvorverarbeitung im Prognoseprozess	147
Abbildung 3.26: Datenanalyse im Prognoseprozess.....	150
Abbildung 3.27: Modellselektion und Verfahrensselektion im Prognoseprozess	151
Abbildung 3.28: Baumdiagramm zur Modellselektion.....	152
Abbildung 3.29: Verfahrensspezifische Datenvorverarbeitung im Prognoseprozess.....	154
Abbildung 3.30: Beurteilung der Verfahrensgüte im Prognoseprozess.....	156
Abbildung 4.1: Biologisches Neuron im Neuronenverbund.....	165
Abbildung 4.2: Zentrales Nervensystem und die Ebenen struktureller Organisation.....	166
Abbildung 4.3: Künstliches Neuron im Neuronenverbund.....	167
Abbildung 4.4: Aktivierungsfunktionen und Aktivierungszustände	172
Abbildung 4.5: Aktivierungszustände von Aktivierungsfunktionen	174
Abbildung 4.6: Signalverarbeitung im künstlichen Neuron	176
Abbildung 4.7: Grafische Notation der Informationsverarbeitung im Neuron	178
Abbildung 4.8: Mehrschichtiges Neuronales Netz	179
Abbildung 4.9: Modellierung der Schwellwerte der Aktivierung durch ein Bias-Neuron	180

Abbildung 4.10: Topologien und Verbindungsmatrizen für vorwärtsgerichtete NN ohne (a.) und mit schichtübergreifenden Verbindungen (b.).....	182
Abbildung 4.11: Topologien und Verbindungsmatrizen für rückgekoppelte NN mit direkten (a.), indirekten (b.) und lateralen Rückkopplungen (c.).....	183
Abbildung 4.12: Funktionsverarbeitung in vorwärtsgerichteten Neuronalen Netzen	184
Abbildung 4.13: Funktionsverarbeitung in rekurrenten Neuronalen Netzen.....	184
Abbildung 4.14: Netzwerkarchitektur des Perceptron (a.) und ADALINE (b.).....	185
Abbildung 4.15: Prototypische Netzwerkarchitekturen des MLP (b.), (c.) und (d.), GFFN (d.), RBF NN (e.), HBF NN (f.), PNN (g.) und JNN (h.).....	186
Abbildung 4.16: Zusammenhang von Lernzielen und Lernalgorithmen [erweiterte Darstellung nach Rojas 1993].....	189
Abbildung 4.17: Lernen ohne Feedback (a.) und Lernen mit Feedback zur Fehlerkorrektur (b.).....	192
Abbildung 4.18: Problemklassen der Modellierung und Architekturen Neuronaler Netze	193
Abbildung 4.19: Netzwerkfehler als bivariate Fehlerfunktion mit einem globalen Minimum (a.) und als Gradienten der Fehlerfunktion (b.)	195
Abbildung 4.20: Bivariate Fehlerfunktionen mit je einem globalen und lokalen Minimum (a.) sowie mit multiplen lokalen Minima der Fehlerfläche (b.)	200
Abbildung 4.21: Bivariate Fehlerfunktion mit mehreren lokalen Minima (a.), Höhenlinien (b.) und Gradienten der Fehlerfunktion (c.).....	200
Abbildung 4.22: Underfitting (a.), Fitting (b.) und Overfitting (c.) bei der Approximation von Funktionen durch Neuronale Netze.....	204
Abbildung 4.23: Verlauf der Trainings- und Validierungsfehler im Training.....	205
Abbildung 4.24: Anwendungsbereiche von Neuronalen Netzen.....	210
Abbildung 4.25: Anwendungen von Neuronalen Netzen zur Entscheidungsunterstützung	213
Abbildung 4.26: Wissenschaftliche Veröffentlichungen zur Prognose mit Neuronalen Netzen.....	216
Abbildung 4.27: Betriebswirtschaftliche Veröffentlichungen zur Prognose mit Neuronalen Netzen.....	219
Abbildung 4.28: Architektur eines MLP zur zeitreihenanalytischen Einschritt-Prognose.....	226
Abbildung 4.29: Architektur eines MLP zur zeitreihenanalytischen Mehrschritt-Prognose.....	228
Abbildung 4.30: Architektur eines MLP zur univariaten, nichtlinearen Regression	229
Abbildung 4.31: Architektur eines MLP zur dynamischen Regression.....	230
Abbildung 4.32: Architektur eines MLP zur kausalen Prognose mit Interventionsvariablen.....	231

Abbildung 4.33: Architektur eines MLP zur Modellierung eines linearen AR(1)-Prozesses (a.), eines linearen AR(p)-Prozesses (b.) und eines nichtlinearen AR(p)-Prozesses (c.)	232
Abbildung 4.34: Architektur eines MLP zur Modellierung einer univariaten linearen Regression (a.), einer multivariaten linearen Regression (b.) und einer nichtlinearen logistischen Regression (c.)	233
Abbildung 4.35: Architektur eines MLP zur Modellierung der Naiven Verfahren NF1 (a.), NF2 (b.) und NF3 (c.).....	234
Abbildung 4.36: Architektur eines MLP zur Modellierung eines gleitenden Mittelwertes (a.) und eines exponentiell gewichteten gleitenden Mittelwertes (b.).....	234
Abbildung 4.37: Architektur eines MLP zur Modellierung eines linearen AR(p)-Prozesses	236
Abbildung 5.1: Quadratische Fehlerfunktionen des $SE(e)$, des skalierten quadratischen Fehlers $0,5 SE(e)$ und der Differenz $SE(e) - 0,5 SE(e)$	246
Abbildung 5.2: Minkowski- R -Fehlerfunktionen für $R = 1$, $R = 2$ und $R = 4$ im eindimensionalen Fall (a.) sowie als Distanzen im zweidimensionalen Fall (b.).....	252
Abbildung 5.3: Fehlerfunktionen des absoluten Fehlers $AE(e)$, des quadratischen Fehlers $SE(e)$ und der Differenz beider Fehlerfunktionen $AE(e) - SE(e)$	253
Abbildung 5.4: Fehlerverlauf der $\ln(\cosh)$ -Fehlerfunktion, der absoluten Fehlerfunktion $AE(e)$ und der Differenz von $\ln(\cosh(e)) - AE(e)$ für $c = 10$ (a.) und für unterschiedliche c (b.).....	255
Abbildung 5.5: Epsilon-insensitive Fehlerfunktion der Support Vector Regression.....	258
Abbildung 5.6: Fehlerfunktion und erste Ableitungen des quadratischen Fehler $SE(e)$ im Vergleich zu robusten Funktionen des $AE(e)$, HUBERS winsorierter Fehlerfunktion und TUKEYS Biweight-Fehlerfunktion	259
Abbildung 5.7: Erste Ableitungen von HAMPELS Fehlerfunktion (a.) und einer gekappten Fehlerfunktion (b.).....	260
Abbildung 5.8: Asymmetrische $LINLIN(e)$ -Kostenfunktion im Vergleich zum quadratischen Fehler $SE(e)$ und die Differenz beider Funktionen $LINLIN(e) - SE(e)$	267
Abbildung 5.9: Optimaler Schätzer einer symmetrischen Verteilung bei symmetrischer Kostenfunktion (a.) und asymmetrischer Kostenfunktion (b.)	268
Abbildung 5.10: Optimaler Schätzer einer asymmetrischen Verteilung bei einer asymmetrischen Kostenfunktion.....	269
Abbildung 5.11: $QQC(e)$ -asymmetrische Kostenfunktion im Vergleich zur quadratischen Kostenfunktion $SE(e)$ und ihrer Differenz $QQC(e) - SE(e)$	270

Abbildung 5.12: $LINEX(e)$ -Kostenfunktion im Vergleich zur Kostenfunktion $SE(e)$ und zur Differenz $LINEX(e) - SE(e)$ (a.) sowie unterschiedliche Parametrisierungen von c, d für $LINEX(e)$ (b.).....	271
Abbildung 5.13: Bivariate Fehlerflächen der Fehlerkosten für den quadratischen Fehler $SSE(e)$ (a.) und den modifizierten quadratischen Fehler $\frac{1}{2} SSE(e)$ (b.).....	280
Abbildung 5.14: Bivariate Fehlerfläche der Fehlerkosten für eine asymmetrische Kostenfunktion $LINLIN(e)$ (a.) und der Differenz der Fehlergebirge von $SE(e)$ und $LINLIN(e)$ (b.).....	281
Abbildung 5.15: Histogramme der Wahrscheinlichkeitsverteilungen stationärer Zufallsvariablen mit unterschiedlicher Streuung und überlagerter Darstellung der Dichtefunktion der Normalverteilung.....	283
Abbildung 5.16: Ausschnitte der ersten 120 Werte stationärer Zeitreihen mit niedriger, mittlerer und hoher Nachfrageschwankung der Poissonverteilung, negativen Binomialverteilung und der Normalverteilung.....	284
Abbildung 5.17: Ausschnitte der ersten 120 Werte saisonaler Zeitreihen mit niedrigen, mittleren und hohen Nachfrageschwankungen der Poissonverteilung, negativen Binomialverteilung und der Normalverteilung.....	285
Abbildung 5.18: Originalzeitreihe und Prognosen des MLP. SE auf den stationären Zeitreihen NV_{niedrig} (a.), NV_{mittel} (b.) und NV_{hoch} (c.).....	290
Abbildung 5.19: Streudiagramme der Prognosewerte auf Trainings-, Validierungs- und Testmenge durch MLP. SE auf den Zeitreihen NV_{niedrig} (a.), NV_{mittel} (b.) und NV_{hoch} (c.).....	291
Abbildung 5.20: Originalzeitreihe und Prognosen des MLP. SE auf den saisonalen Zeitreihen $S.NV_{\text{niedrig}}$ (a.), $S.NV_{\text{mittel}}$ (b.) und $S.NV_{\text{hoch}}$ (c.).....	291
Abbildung 5.21: Streudiagramme der Prognosewerte auf Trainings-, Validierungs- und Testmenge durch MLP. SE auf den Zeitreihen $S.NV_{\text{niedrig}}$ (a.), $S.NV_{\text{mittel}}$ (b.) und $S.NV_{\text{hoch}}$ (c.).....	292
Abbildung 5.22: Originalzeitreihe und Prognosen des MLP. SE (a.), des MLP. ACF_1 (b.), des MLP. ACF_2 (c.) und des MLP. ACF_3 (d.) auf der stationären Zeitreihe NV_{mittel}	292
Abbildung 5.23: Originalzeitreihe und Prognosen des MLP. AE (a.), des MLP. ACF_1 (b.), des MLP. ACF_2 (c.) und des MLP. ACF_3 (d.) auf der saisonalen Zeitreihe $S.NV_{\text{mittel}}$	294
Abbildung 5.24: Originalzeitreihe und Prognose des MLP. ACF_3 auf den stationären Zeitreihen NV_{niedrig} (a.), NV_{mittel} (b.) und NV_{hoch} (c.).....	295
Abbildung 5.25: Originalzeitreihe und Prognosen des MLP. ACF_3 auf den saisonalen Zeitreihen $S.NV_{\text{Niedrig}}$ (a.), $S.NV_{\text{Mittel}}$ (b.) und $S.NV_{\text{Hoch}}$ (c.).....	295
Abbildung 5.26: Streudiagramme der Prognosewerte auf Trainings-, Validierungs- und Testmenge für alle Zielfunktionen und alle Zeitreihen mit normalverteiltem Bedarfsverlauf.....	297

Abbildung 5.27: Verteilung der α -Servicegrade über alle MLP der Trainingsmenge (a.), der Validierungsmenge (b.) und der Testmenge (c.) für die Zielfunktionen SE , ACF_1 , ACF_2 und ACF_3	299
Abbildung 5.28: Prozessketten der Disposition mit statistischen Prognoseverfahren (a.), MLP. SE als Prognoseverfahren (b.) und MLP. ACF mit asymmetrischen Zielfunktionen zur Disposition (c.).....	310
Abbildung 6.1: Systematik zur Einteilung von Automatentypen	319
Abbildung 6.2: Lagersystem der Warenautomaten.....	322
Abbildung 6.3: Prozess der Warendisposition und Befüllung an Warenautomaten.....	323
Abbildung 6.4: Konstituenten des Lagerhaltungsproblems an Warenautomaten.....	324
Abbildung 6.5: Zeitreihen des aggregierten Absatzes aller Zeitreihen der Absatzstellen AU-1 (a.), AU-2 (b.), IN-1 (c.), IN-2 (d.) und IN-3 (e.)	331
Abbildung 6.6: ACF (a.) und PACF (b.) der Originalzeitreihe sowie ACF (c.) und PACF (d.) der ersten Differenzen des Absatzes an Abverkaufsstelle AU-1	333
Abbildung 6.7: ACF (a.) und PACF (b.) der Originalzeitreihe sowie ACF (c.) und PACF (d.) der ersten Differenzen des Absatzes an Abverkaufsstelle AU-2	333
Abbildung 6.8: ACF (a.) und PACF (b.) der Originalzeitreihe sowie ACF (c.) und PACF (d.) der ersten Differenzen des Absatzes an Abverkaufsstelle IN-1	334
Abbildung 6.9: ACF (a.) und PACF (b.) der Originalzeitreihe sowie ACF (c.) und PACF (d.) der ersten Differenzen des Absatzes an Abverkaufsstelle IN-2	335
Abbildung 6.10: ACF (a.) und PACF (b.) der Originalzeitreihe sowie ACF (c.) und PACF (d.) der ersten Differenzen des Absatzes an Abverkaufsstelle IN-2	335
Abbildung 6.11: Zeitreihen von Artikeln mit langsamem Warenumsatz (AU-1-08 (a.) und AU-1-10 (b.)) und mit schnellem Warenumsatz (AU-1-01 (c.) und AU-1-09 (d.)).....	337
Abbildung 6.12: ABC-Analyse des mittleren Absatzes über alle Abverkaufsstellen.....	340
Abbildung 6.13: Übersicht der Prognoseverfahren in der Vergleichsstudie.....	347
Abbildung 6.14: Prognosefehler des $mwMAE$ aller Verfahren auf den Testdaten über alle Zeitreihen	359
Abbildung 6.15: Box-Whisker-Grafiken des $mwMAE$ auf den Testdaten über alle Zeitreihen	364
Abbildung 6.16: Korrelationsdiagramm der Trainings-, Validierungs- und Testfehler des $mwMAE$ (a.) und des $Rang\ mwMAE$ (b.) über alle Zeitreihen.	366
Abbildung 6.17: Korrelationsdiagramm der Validierungs- und Testfehler des $mwMAE$ (a.) und des $Rang\ mwMAE$ (b.) über alle Zeitreihen.....	367

Abbildung 6.18: Kumulierte $mwMAE$ je Absatzstelle AU-1, AU-2, IN-1 und IN-2 auf der Testmenge.....	371
Abbildung 6.19: Prognosegüte der multivariaten Verfahren mv.MLP und mv.EXP.ES nach MAE je Zeitreihe auf der Testmenge der Außenautomaten (a.) und der Innenautomaten (b.).....	375
Abbildung 6.20: Prognosegüte der univariaten Verfahren MLP und ausgewählter statistischer Verfahren nach MAE je Zeitreihe an Außenautomaten (a.) und Innenautomaten (b.).....	375
Abbildung 6.21: Dispositionskosten der Prognoseverfahren nach $mwMLLC$ auf den Testdaten über alle Zeitreihen.....	382
Abbildung 6.22: Box-Whisker-Graphen der Dispositionskosten nach $mwMLLC$ der Testdaten über alle Zeitreihen.....	384
Abbildung 6.23: Dispositionsgüte der multivariaten Verfahren mv.MLP.ACF , mv.MLP und mv.EXP.ES nach $MLLC$ je Zeitreihe auf den Testmengen der Außenautomaten (a.) und der Innenautomaten (b.).....	394
Abbildung 6.24: Dispositionsgüte der univariaten Verfahren MLP, NF1, S.ES, und EXP.ES&ARIMA nach $MLLC$ je Zeitreihe an Außenautomaten (a.) und Innenautomaten (b.).....	395
Abbildung 6.25: Streudiagramm der Prognosefehler nach $mwMAE$ und der Dispositionskosten nach $mwMLLC$ je Prognoseverfahren	400
Abbildung 6.26: Streudiagramme der Prognosefehler nach $Rang\ mwMAE$ und der Dispositionskosten nach $Rang\ mwMLLC$ je Prognoseverfahren.....	402

Tabellenverzeichnis

Tabelle 2.1:	Systematik der Bedarfsarten, Modelle und Verfahren der Bedarfsrechnung	26
Tabelle 2.2:	Systematik der Strukturdimensionen von Bestellmengenmodellen	40
Tabelle 2.3:	Entscheidungsvariablen der Bestellmengenmodelle.....	55
Tabelle 2.4:	Deskriptive Fehler zur Beurteilung der Dispositionsgüte.....	62
Tabelle 2.5:	Deskriptive Fehlermaße zur Beurteilung der Dispositionsgüte	64
Tabelle 3.1:	Kategorisierung statistischer Modelle nach dem Skalenniveau der Variablen.....	74
Tabelle 3.2:	Systematik der Determinanten von Erklärungs- und Prognosemodellen.....	85
Tabelle 3.3:	Fristen und Kalenderzeiträume von Prognosenverfahren.....	86
Tabelle 3.4:	Prognosefehler zur Beurteilung der Prognosegenauigkeit.....	131
Tabelle 3.5:	Statistische Fehlermaße zur Beurteilung der Prognosegenauigkeit.....	134
Tabelle 4.1:	Eingabefunktionen in Neuronen	170
Tabelle 4.2:	Kategorisierung Neuronaler Netze nach den Skalenniveaus der Variablen.....	215
Tabelle 4.3:	Publikationen zur Absatzprognose mit Neuronalen Netzen	223
Tabelle 5.1:	Ausgabewerte und Konfusionsmatrix des $SE(e)$ für nominal skalierte Ausgaben.....	248
Tabelle 5.2:	Ausgabewerte und Konfusionsmatrix des $SE(e)$ für binäre Ausgaben.....	248
Tabelle 5.3:	Ausgabewerte und Konfusionsmatrix des $SE(e)$ für metrische Ausgaben.....	248
Tabelle 5.4:	Asymmetrische Kostenfunktionen zur Beurteilung der Entscheidungskosten.....	271
Tabelle 5.5:	Parameter der asymmetrischen Zielfunktionen im Experiment.....	286
Tabelle 5.6:	α -Servicegrade aller MLP je Datenmenge und Zeitreihe	298
Tabelle 5.7:	α - Servicegrade aller MLP je Datenmenge und Zeitreihenmuster	300
Tabelle 5.8:	Fehlermaße und Kostenmaße aller MLP je Datenmenge über alle Zeitreihen	302
Tabelle 5.9:	Ausgewählte Prognosemodelle je Zeitreihe durch das Expertensystem.....	304
Tabelle 5.10:	Servicegrade und Dispositionskosten der Verfahren auf stationären Zeitreihen	305
Tabelle 5.11:	Servicegrade und Dispositionskosten der Verfahren auf saisonalen Zeitreihen	306

Tabelle 5.12:	Mittlere Dispositionskosten der Verfahren nach Verlauf und Verteilung der Zeitreihe	307
Tabelle 5.13:	Mittlere Ränge der Dispositionskosten der Verfahren nach Verlauf und Verteilung der Zeitreihe	308
Tabelle 6.1:	Statistik der Verpflegungsautomaten in Europa (ohne Zigarettenautomaten)	321
Tabelle 6.2:	Determinanten der Disposition an Warenautomaten	327
Tabelle 6.3:	Stichprobe der Absatzstellen und Zeitreihen	329
Tabelle 6.4:	Parameter der aggregierten Zeitreihen je Absatzstelle nach Standort	330
Tabelle 6.5:	Kolmogorov-Smirnow-Tests der Absatzzeitreihen auf unterschiedliche Verteilungsformen	338
Tabelle 6.6:	Korrelationskoeffizienten nach SPEARMANs rho der Zeitreihen der Absatzstelle AU-1	339
Tabelle 6.7:	Lageparameter und Klassifikation der Zeitreihen nach Absatzstellen	341
Tabelle 6.8:	Mittlere Lageparameter und Zeitreihenverläufe nach Absatzstelle und Standort	342
Tabelle 6.9:	Mittlere Lageparameter der Kategorien von Zeitreihen	343
Tabelle 6.10:	Systematik der Zeitreihen Kategorien	344
Tabelle 6.11:	Dimensionen des Verfahrensvergleichs nach Prognosemodell und -verfahren	345
Tabelle 6.12:	Übersicht der Verfahren zur Absatzprognose in der Vergleichsstudie	352
Tabelle 6.13:	Wilcoxon-Vorzeichen-Rang-Tests auf Unterschiede in den Prognosefehlern des <i>mwMAE</i>	360
Tabelle 6.14:	Fehlermaße und Rang der Fehlermaße des MAE über alle Absatzstellen	362
Tabelle 6.15:	Korrelationen von Trainings-, Validierungs- und Testfehler der <i>mwMAE</i>	366
Tabelle 6.16:	Fehlermaße und Rang der Fehlermaße des <i>MAPE</i> über alle Absatzstellen	368
Tabelle 6.17:	Summe, Mittelwert und Rang des <i>MAE</i> je Absatzstelle	370
Tabelle 6.18:	<i>mdMAE</i> und Rang <i>mdMAE</i> nach Absatzstellen	372
Tabelle 6.19:	Summen des Rang <i>mwMAE</i> und entsprechende Rangordnung je Absatzstelle	373
Tabelle 6.20:	Korrelationen der Trainings-, Validierungs- und Testfehler der <i>mwMAE</i> nach Absatzstelle	374
Tabelle 6.21:	<i>mwMAE</i> und <i>mdMAE</i> über alle Zeitreihen nach Absatzverlauf	377
Tabelle 6.22:	<i>mwMAE</i> , <i>mdMAE</i> und Rangordnung über alle Absatzstellen nach Umschlaggeschwindigkeit	378

Tabelle 6.23:	<i>mwMAE</i> und <i>mdMAE</i> über alle Zeitreihen nach Artikelklassifikation.....	380
Tabelle 6.24:	Wilcoxon-Vorzeichen-Rang-Test auf Unterschiede in den Dispositionskosten nach <i>mwMLLC</i>	383
Tabelle 6.25:	<i>MLLC</i> Kosten der Warenbestände und Rang nach <i>MLLC</i> über alle Absatzstellen.....	386
Tabelle 6.26:	Mittlere Servicegrade, Abweichungen vom 96,5%-Soll-Servicegrad und Rang über alle Zeitreihen.....	388
Tabelle 6.27:	Summen und Mittelwerte der <i>MLLC</i> der Warenbestände je Absatzstelle.....	390
Tabelle 6.28:	Mediane der <i>MLLC</i> der Warenbestände je Absatzstelle.....	391
Tabelle 6.29:	Summe der Ränge nach <i>sumMLLC</i> der Warenbestände je Absatzstelle.....	392
Tabelle 6.30:	Mittlere Servicegrade, Abweichungen vom 96,5%-Soll-Servicegrad und entsprechender Rang.....	393
Tabelle 6.31:	<i>mwMLLC</i> und <i>mdMLLC</i> über alle Zeitreihen nach Absatzverlauf.....	396
Tabelle 6.32:	<i>mwMLLC</i> und <i>mdMLLC</i> über alle Absatzstellen nach Umschlagsgeschwindigkeit.....	398
Tabelle 6.33:	<i>mwMLLC</i> und <i>mdMLLC</i> über alle Zeitreihen nach Artikelklassifikation.....	399

Abkürzungsverzeichnis

Abb.	Abbildung
ABC	Artikelklassifikation als A-, B- oder C-Artikel
ACF	Asymmetrische Kostenfunktion
ADALINE	adaptive linear element
ADF	Augmented Dickey Fuller
ADL	Autoregressive Distributed Lag
AKF	Autokorrelationsfunktion
ANOVA	Analysis of Variance
APICS	The Association for Operations Management
AR	Autoregressive
ARCH	Autoregressive Conditional Heteroscedasticity
ARIMA	Autoregressive Integrierte Moving Average
ARIMAX	ARIMA mit kausalen Interventionen X
ARIS	Architektur Integrierter Informationssysteme
ARMA	Autoregressive Moving Average
ARX	Regressionsmodell mit Interventionsmodellierung
ATM	Automat
AU	Außenautomat
Aufl.	Auflage
Bd.	Band
BDV	Bundesverband der Deutschen Vending-Unternehmen
bzw.	beziehungsweise
cdf	cumulative distribution function
CM	Category Management
CPFR	Collaborative Planning Forecasting and Replenishment
D.ES	Doppelte Exponentielle Glättung
d.h.	das heißt
DGOR	Deutsche Gesellschaft für Operations Research
DHI	Deutsches Handelsinstitut
Diss.	Dissertation
DM	Deutsche Mark
DPR	Direkte Produktrentabilität
DT.ES	Dampened Trend Exponential Smoothing
ECR	Efficient Consumer Response
Ed.	Edition
EHI	Europäisches Handelsinstitut
engl.	englisch
ENNS	European Neural Network Society
EOQ	Economic Order Quantity
ES	Exponentielle Glättung
et al.	et alii
etc.	et cetera
EV	Erlangverteilung
EXP	Expertensystem
f.	folgende Seite
FDI	Association of Food Distributors International
ff.	fortlaufend folgende Seiten
FL	Filiallager
FMI	Food Marketing Institute

GB	Gigabyte
GE	Geldeinheiten
GFFN	Generalised-Feed-Forward Neural Network
ggf.	gegebenenfalls
GHz	Gigahertz
GMA	Association of the Grocery Manufacturers of America
GOR	Deutsche Gesellschaft für Operations Research
GRNN	General Regression Neural Network
HBF NN	Hyper-Basisfunktions-Neuronales Netz
Hrsg.	Herausgeber
hrsg.	herausgegeben
i. A.	im Allgemeinen
i. d. R.	in der Regel
i. e. S.	im engeren Sinn
i. w. S.	im weiteren Sinn
IBF	Institute of Business Forecasting
ICANN	International Conference on Artificial Neural Networks
Id	Identitätsfunktion
IEEE	Institute of Electrical Engineers (auch Ieee)
IIF	International Institute of Forecasting
IJCNN	International Joint Conference on Neural Networks
IJF	International Journal of Forecasting
IN	Innenautomat
INNS	International Neural Network Society
IT	Informations- und Kommunikationstechnologie
Jg.	Jahrgang
JNN	Jordan Neural Network
KS	Kolmogorov-Smirnow
Kurt	Kurtosis von Beobachtungen
KW	Kalenderwoche
L.ES	Lineare Exponentielle Glättung
LMS	Least Mean Squares
Log	Logistische Funktion
LTD	Lead Time Demand
LVQ NN	Learning Vector Quantisation Neural Network
MA	Moving-Average
MDE	Mobile Datenerfassung
ME	Mengeneinheiten
MLP	Multilayer Peceptron
mv	multivariate
MW	Prognoseverfahren eines Mittelwertes
NAMA	National Automatic Merchandising Association
NB	Negative Binomialverteilung
NF1	Naives Prognoseverfahren ohne Trend und Saison
NF2	Naives Prognoseverfahren für Trend
NF3	Naives Prognoseverfahren für Saison
NN	künstliche Neuronale Netze
Nr.	Nummer
NuP %	Anteil der Nullperioden von Beobachtungen

NuP	Anzahl der Nullperioden der Zeitreihe
NV	Normalverteilung
o. S.	ohne Seitenangabe
o. V.	ohne Verfasserangabe
OOS	out of stock
PAKF	Partielle Autokorrelationsfunktion
PC	Personal Computer
pdf	probability density function
pmf	probability mass function
PNN	Probabilistisches Neuronales Netz
POS	point of sale
PV	Poissonverteilung
RBF NN	Radiale Basisfunktions Neuronales Netz
RL	Regionallager
RV	Zufallsvariable
S&OP	Sales and Operations Planning
S.	Seite
S.ES	Einfache Exponentielle Glättung
SARIMAX	Saisonale ARIMA-Modelle mit kausalen Interventionen
SCM	Supply Chain Management
SE	Squared Error
Skew	Skewness von Beobachtungen
SKU	stock keeping unit
sog.	sogenannt
STAR	Smooth Transition Autoregression
StdAb	Standardabweichung von Beobachtungen
Sum	Summe
SVR	Support Vector Regression
Tab.	Tabelle
TanH	Hyperbolischer Tangens
TDNN	Time Delay Neural Network
TOUR	Tourenfahrzeug
u. a.	unter anderem
u. U.	unter Umständen
UPoS	unattended point of sale
US	United States of America, Vereinigte Staaten von Amerika
usw.	und so weiter
VaKo	Variationskoeffizient von Beobachtungen
VAR	Vector Autoregressive
vgl.	vergleiche
VMI	Vendor Managed Inventory
Vol.	Volume
VR	Verkaufsraum
XOR	logisches „Entweder-oder“-Problem
z. B.	zum Beispiel
z. T.	zum Teil
ZE	Zeiteinheiten
ZL	Zentrallager
ZNS	Zentrales Nervensystem

Symbolverzeichnis

Warendisposition

a	Anfangsbestand des Lagers am Periodenanfang [ME]
α	Maßzahl des α -Servicegrades von Fehlmengen
β	Maßzahl des β -Servicegrades von Fehlmengen
C_B	Bestellkosten in [GE]
C_F	Fehlmengenkosten in [GE]
$C_{F,L}$	Fehlmengenkosten im Verlustfall in [GE]
$C_{F,V}$	Fehlmengenkosten im Vormerkfall in [GE]
C_L	Lagerungskosten [GE]
c_B	variable Stückkosten der Bestellung in [GE / ME]
$c_{F,L}$	Fehlmengenkosten des entgangenen Gewinns im Verlustfall [GE / ME]
$c_{F,V}$	Fehlmengenkosten der Lieferverzögerungen im Vormerkfall [GE / ME]
$c_{G,V}$	Fehlmengenkosten des Goodwill-Verlustes im Verlustfall [GE / ME]
$c_{G,V}$	Fehlmengenkostensatz des Goodwill-V. im Vormerkfall [GE / ME * ZE]
c_L	variable Lagerhaltungskosten [GE / ME * ZE]
c_o	Kosten der Bedarfsüberdeckung
c_u	Kosten der Bedarfsunterdeckung
g	Schrottwert [GE/ME]
l	Lagerkostensatz in Prozent des gebundenen Kapitals [% / ZE]
p	Einstandspreis der Ware pro Stück [GE / ME], auch Verkaufspreis
q	konstante Bestellmenge (Losgröße) über alle Planungsperioden [ME]
q^*	optimale Bestellmenge [ME]
q_F	Höhe der Fehlmengen in [ME]
q_i	Bestellmenge in der Planungsperiode i [ME]
r	Bestellrhythmus in [ZE], bei konstanten Bestellzyklen $\tau_i = \tau_{i+1} = \dots = r$
S	Bestellniveau des Lagerbestandes [ME]
S^*	optimales Bestellniveau des Lagerbestandes [ME]
s	Bestellpunkt [ME]
s^*	optimaler Bestellpunkt [ME]
τ	Bestellzyklus in [ZE], als Zeit zwischen zwei Bestellterminen t_i und t_{i+1}
τ^*	optimaler Bestellzyklus in [ZE]
t	Zeitpunkt [-], von $t=1, 2, \dots, T$, auch Planungsperiode
t_i	Zeitpunkt in der Planungsperiode i [-], z. B. ein Bestelltermin
t'_i	Wareneingangstermin der Bestellung in t_i in [-]
Δt_A	Auftragsvorbereitungszeit in [ZE]
Δt_E	Einlagerungszeit in [ZE]
Δt_F	zeitliche Dauer einer Fehlmengensituation im Vormerkfall in [ZE]
Δt_L	Lieferzeit in [ZE]
Δt_U	Überprüfungsintervall in [ZE]
Δt_W	Wiederbeschaffungszeit in [ZE]
W	Restbestand des Lagers am Periodenende [ME]
Y	Zufallsvariable des Bedarfs
y_t	Bedarf in der Periode t in [ME/ZE] (ebenso Absatz bzw. Nachfrage)

Prognose

AE_{t+h}	Statistisches Fehlermaß des <i>absolute error</i>
APE_{t+h}	Statistisches Fehlermaß des <i>absolute percent error</i>
$AR(p)$	Autoregressiver Prozess der Ordnung p
a_i	Parameter eines Mittelwertes, für $i=1, 2, \dots, I$
ACF	Kostenmaß der asymmetrischen Kosten
α	Alpha-Glättungsparameter für das Niveau der Zeitreihe zur ES
b_i	Parameter eines Regressionsmodells, für $i=1, 2, \dots, I$
b_0	Konstanter Regressionsparameter
B	Backshift-Operator der ARIMA-Notation
β	Beta-Glättungsparameter für den Trend der Zeitreihe zur ES
d	Grad der Integration der Zeitreihen $I(d)$ Prozesses
D	Grad der saisonalen Integration der Zeitreihen $SI(D)_s$ Prozesses
e_t	Prognosefehler im Zeitpunkt t
E_{t+h}	Statistisches Fehlermaß des <i>simple error</i>
$E(Y)$	Erwartungswert von Y
E_t	irreguläre Komponente einer Zeitreihe
ε	Zufallsvariable, unerklärter Rest der Modellierung
$F_Y(y)$	kumulierte Verteilungsfunktion (cdf) des Auftretens von y
ϕ_i	Parameter eines $AR(p)$ -Prozesses
Φ_i	Parameter eines saisonalen $SAR(p)$ -Prozesses
$GMRAE_{m,h}$	Komparatives Fehlermaß des <i>geometric mean relative absolute error</i>
γ	Gamma-Glättungsparameter für das Niveau der Zeitreihe zur ES
$\gamma(x, y)$	Kovarianz von x und y , auch: $cov(x, y)$
H_i	Nullhypothese i
h	Prognosehorizont [ZE], mit $h=1, 2, \dots, H$
$PV(\mu)$	Poissonverteilung mit dem Parameter μ
$f_Y(y)$	Wahrscheinlichkeitsfunktion (PDF) des Auftretens von y
$I(d)$	Integration der Zeitreihe der Ordnung d
i	Index, u. a. für Planungsperiode [-]
I_t	Informationsgrad im Zeitpunkt t
k	Sicherheitsfaktor für standardisierte Verteilungstabellen
L_t	Zeitreihenkomponente des Niveaus
l	Lag der zeitlichen Verzögerung [ZE], mit $l=1, 2, \dots, L$
l_i	Grad i der Verlustfunktion, auch R in Minkowski- R -Norm
l_1	Lineare Verlustfunktion des $AE(e)$
l_2	Quadratische Verlustfunktion des $SE(e)$
Max	Maximum von Beobachtungen
$MA(q)$	Moving-Average-Prozess der Ordnung q
$MACF$	Kostenmaß der mittleren asymmetrischen Kosten
MAD_{t+h}	Statistisches Fehlermaß des <i>mean absolute deviation</i> , auch MAE_{t+h}
MAE_{t+h}	Statistisches Fehlermaß des <i>mean absolute error</i> , auch MAD_{t+h}
$MAPE_{t+h}$	Statistisches Fehlermaß des <i>mean absolute percent error</i>
$MdAPE_{t+h}$	Statistisches Fehlermaß des <i>median absolute percent error</i>
md	Median, auch Med
ME_{t+h}	Statistisches Fehlermaß des <i>mean error</i>
mw	Arithmetisches Mittel, auch Mw
Med	Median von Beobachtungen, auch md

Prognose (fortgesetzt)

Min	Minimum von Beobachtungen
MPE_{t+h}	Statistisches Fehlermaß des <i>mean percent error</i>
MSE_{t+h}	Statistisches Fehlermaß des <i>mean squared error</i>
Mw	Mittelwert von Beobachtungen
μ	Mittelwert
$N(\mu, \sigma)$	Normalverteilung mit den Parametern μ, σ
n	Gegenwartszeitpunkt, für $t = n$
n	Stichprobengröße der Anzahl von Beobachtungswerten
$OOS(e)$	Deskriptives Fehlermaß des Auftretens einer Fehlmengensituation
$OOU(e)$	Deskriptives Fehlermaß der Anzahl unterbundener Verkäufe
$OVS(e)$	Deskriptives Fehlermaß des Auftretens einer Überbevorrattungssituation
$OVU(e)$	Deskriptives Fehlermaß der Anzahl überbevorrateter Mengen
P	Ordnung des saisonalen autoregressiven $SAR(P)_s$ Prozesses
p	Eintrittswahrscheinlichkeit
p	Ordnung des autoregressiven $AR(p)$ Prozesses
$PB_{m,h}$	Komparatives, statistisches Fehlermaß des <i>percent better</i>
PE_{t+h}	Statistisches Fehlermaß des percent error
$\rho_Y(y)$	Wahrscheinlichkeitsfunktion (pmf) des Auftretens von y
Q	Ordnung des saisonalen moving average $SMA(Q)_s$ Prozesses
q	Ordnung des moving average $MA(q)$ Prozesses
R	Grad der Minkowski- R -Metrik, auch i in l_i
RAE_{t+h}	Komparatives, statistisches Fehlermaß des relative absolute error
$RMSE_{t+h}$	Statistisches Fehlermaß des <i>root mean squared error</i>
RSE_{t+h}	Statistisches Fehlermaß des <i>root of squared error</i>
s	Länge des Saisonzyklus
SAE_{t+h}	Statistisches Fehlermaß des <i>sum of absolute errors</i>
SE_{t+h}	Statistisches Fehlermaß des <i>squared error</i>
SSE_{t+h}	Statistisches Fehlermaß des <i>sum of squared errors</i>
S_t	Saisonfaktor
S_t	Zeitreihenkomponente der Saison
s_Y	empirische Standardabweichung der Stichprobe von Y
s_Y^2	empirische Varianz der Stichprobe von Y
sum	Summe
$\hat{\sigma}_y$	Prognose der Standardabweichung von Y
σ_y	Standardabweichung von Y
σ_y^2	Varianz von Y , auch: $\text{var}(Y)$
T_t	Zeitreihenkomponente des Trends
θ_i	Parameter eines $MA(q)$ -Prozesses
Θ_i	Parameter eines saisonalen $SMA(q)$ -Prozesses
U_t	Komparatives, statistisches Fehlermaß der U-Statistik nach THEIL
WSE_{t+h}	Komparatives, statistisches Fehlermaß des winsorised error
$x_{i,t}$	unabhängige Variable x_i in der Periode t zur Erklärung des Bedarfs y
Y	Zufallsvariable des stochastischen Bedarfs
y_{t-l}	abhängige Variable des Bedarfs y in der Periode $t-l$
\bar{y}	empirischer Mittelwert des Bedarfs y
\hat{y}_Y	Prognose des Bedarfs y
\hat{y}_{t+h}	Prognose des Bedarfs y für die Periode $t+h$

Neuronale Netze

A	Parameter der Netzwerkarchitektur, mit $A=[V, T, W, K]$
B	Methode zur Steuerung des Lernprozesses
Bin	Binäre Aktivierungsfunktion
DF	Direkte Feedback-Struktur der w_{ij} , Spezifikation von W
ε	Epsilon-Parameter der Support Vector Regression
η	Lernrate
FF	Feed-Forward-Struktur der w_{ij} , Spezifikation von W
F_H	Aktivierungsfunktion
F_I	Eingabefunktion
F_O	Ausgabefunktion
f_{act}	Aktivierungsfunktion
f_{out}	Ausgabefunktion
G	Lernalgorithmus des NN
Id	Identitätsfunktion
I^N	Anzahl der Initialisierungen
I^P	Prozedur zur Bestimmung der Startgewichte
IF	Indirekte Feedback-Struktur der w_{ij} , Spezifikation von W
K	Parameter der Aktivierungsstrategie
L	Parameter des Lernprozesses
LF	Laterale Feedback-Struktur von w_{ij} , Spezifikation von W
Log	Logistische Funktion
N_{Hn}	Anzahl Neuronen in der n -ten versteckten Schicht
N_I	Anzahl Neuronen in der Eingabeschicht
net_j	(Netz-)Eingabefunktion in das Neuron j
O	Zielfunktion im Training eines NN
o_j	Ausgabewert des Neurons j
$P_{T,L}$	Lernparameter für jede Phase T je Neuronenschicht L
S	Vorgehensweise zur Auswahl einer der Netzwerkinitialisierungen
SC	Schichtübergreifende w_{ij} , Spezifikation von W
Sum	Summe
δ_j	Fehler der Netzwerkausgabe am Neuron j
T	Parameter der Netzwerktopologie
T	Temperaturparameter zur Kontrolle der Steigung von Log
TanH	Hyperbolischer Tangens
t_j	Lerneingabe, Soll-Ausgabewert des Neurons j
θ_j	Schwellwert der Aktivierung des Neurons j
u_i	Neuron i
V	Parameter der Signalverarbeitung im Neuron, mit $V=[F_I, F_H, F_O]$
W	Gewichtsmatrix aller Verbindungsgewichte w_{ij}
w_{ij}	Verbindungsgewicht zwischen Neuron i und j

1 Einleitung

Die wirtschaftliche Beschaffung und Bereitstellung von Waren zur Bedarfsbefriedigung der Kunden ist ein konstituierendes Merkmal aller Handelsbetriebe. Aufgrund der abweichenden zeitlichen, räumlichen und mengenmäßigen Struktur von Bedarf und Warenangebot ist eine effiziente Planung der Beschaffungsmengen und -zeitpunkte durch die Entscheidungen der Warendisposition erforderlich. Zu hohe Warenbestände führen zu einer Bereitstellung überflüssiger Güter, die als Überbestände Kapital binden, wertvollen Raum in Verkaufsflächen und Lagern belegen und aufgrund des Risikos von Beschädigungen, Diebstahl und Verderb zusätzliche Kosten verursachen. Sind die Bestände dagegen zu niedrig, entstehen Fehlmengen durch ausverkaufte Produkte, die zu Opportunitätskosten in Form von entgangenem Umsatz und Ertrag führen. Bei wiederholt auftretenden Fehlmengen führt die unbefriedigte Kundennachfrage ferner zu einer Beeinträchtigung der Kundenzufriedenheit, bis hin zu einer Verringerung der Einkaufsstättenloyalität und der strategischen Wettbewerbsfähigkeit. Das Treffen korrekter Dispositionsentscheidungen ist somit von besonderer Bedeutung für den wirtschaftlichen Erfolg im Handel.

Das Entscheidungsproblem der Warendisposition im Handel erfordert zunächst die Prognose des stochastischen Bedarfs im anonymen Markt. Hierbei kommen unterschiedliche Modelle und Verfahren der Bedarfsprognose zum Einsatz, die aufgrund ihrer unterschiedlichen Genauigkeit bei der Bedarfsvorhersage die Entscheidungsgüte entscheidend bestimmen. Nachfolgend wird die Verfahrensklasse der künstlichen Neuronen Netze hinsichtlich ihrer Prognosegenauigkeit und Dispositionskosten untersucht und zur simultanen Bedarfsprognose und Bestellmengenrechnung im Sinne einer integrierten Warendisposition weiterentwickelt. Die Genauigkeit der Neuronen Netze wird anschließend in einer objektiven empirischen Vergleichsstudie mit etablierten Verfahren der Bedarfsprognose verglichen.

1.1 Problemstellung

Handelsbetriebe wurden in den vergangenen Jahren mit einer zunehmenden Wettbewerbsintensität, niedrigen Ertragsspannen und hohem Kostendruck konfrontiert. Gleichzeitig sahen sie sich wachsenden, teils auch ganz neuartigen Anforderungen an die Warendisposition gegenüber, die insbesondere durch die Komplexität der Warenversorgung und die steigende Volatilität der Kundennachfrage verursacht wurden.

Durch das wachsende Preis-Leistungs-Bewusstsein und die sinkende Produkt- und Markenloyalität der Konsumenten verändern sich deren Konsummuster in Richtung auf ein hybrides und fragmentiertes Einkaufsverhalten: Die so genannten „smart shopper“ zeigen eine verringerte Einkaufsstättenloyalität [ECR Europe und Roland Berger 2003]. Die wachsende Konkurrenz um loyale Kundengruppen führt zu einem Zuwachs an Werbeaktivitäten in Industrie und Handel, die je nach Produktkategorie eine Vervielfachung der Nachfrage um bis zu 600 % im Aktionszeitraum bewirken und bei einzelnen Nahrungsmitteln bereits bis zu 75 % des Gesamtumsatzes generieren [Aberdeen Group 2005]. Hersteller und Handel führen neue Produkte in einer nie zuvor erreichten Geschwindigkeit und Menge ein [Aberdeen Group 2004]. Gleichzeitig verkürzen sich die Produktlebenszyklen, so dass Sortimente im Handel kontinuierlich erweitert und in immer kürzeren Zyklen erneuert werden. Die wachsende Sortimentstiefe und -breite führt aufgrund des beschränkten Verkaufsraums zu häufigen Ein- und Auslistungen von Produkten und Produktvarianten.¹ Als Resultat steigt die Volatili-

¹ Bei konstanten Konsumausgaben führt dies zu steigenden Logistikkosten und sinkenden Umsätzen je Artikel und somit zu einer weiteren Erhöhung des Kostendrucks.

tät und Komplexität der Nachfrage in Form von Absatzspitzen, vielfältigen Verschiebungen des Absatzniveaus und lokalen Trends. Dies erschwert zunehmend die Bedarfsprognose und verringert so die Planungsgenauigkeit bei der Warendisposition [Aberdeen Group 2004, S. 3; KPMG und EHI 2006, S. 30; KPMG 2005].

Als Folge erhöhter Bedarfsschwankungen steigen die Kapitalkosten. Die Aufrechterhaltung der Lieferfähigkeit und der Liefertermintreue bei wachsenden Nachfrageschwankungen erfordert eine Erhöhung der Warenbestände im Handel. Im Jahr 2005 investierten die einhundert größten Konsumgüterhersteller über 100 Mrd. US-Dollar in die Bestände in den handelsseitigen Fertigwarenlagern [Page 2005; GMA und IBM 2005]. Gleichzeitig stieg der Anteil unverkäuflicher Ware durch Beschädigung und Verderb auf 1,2 % des Umsatzes, was vermeidbaren Kosten von über 5 Mrd. US-Dollar pro Jahr entspricht [FDI et al. 2002]. Dennoch sinkt die Warenverfügbarkeit in den Absatzstellen des Handels. Unabhängige Studien belegen, dass 8 % aller neu eingeführten Produkte und 17 % der Produkte mit Sonderaktionen täglich ausverkauft sind, mit noch größeren und längeren Präsenzlücken für einzelne Produktkategorien, Wochentage und Absatzregionen [GMA 2002; GMA et al. 2002; ECR Europe und Roland Berger 2003]. Dies ist weit entfernt von der handelsüblichen Zielsetzung eines Servicegrades von 99 % und hat erhebliche Konsequenzen. Den Absatzstellen des europäischen Handels entgehen so 3 bis 4 Prozent (etwa 4 Mrd. Euro) ihres Umsatzes pro Jahr, da Konsumenten fehlende Produkte nur zu 53 % durch alternative Packungsgrößen oder andere Marken substituieren [ECR Europe und Roland Berger 2003; GMA et al. 2002]. Fehlmengen führen zudem zu einem Verlust an Kundenzufriedenheit, bei wiederholtem Auftreten auch zu einer Verringerung der Marken- und Einkaufsstättenloyalität und somit zum Verlust von Wettbewerbsvorteilen [ECR Europe und Roland Berger 2003].

Darüber hinaus führen Fehlbestände zu weiteren Komplikationen in der Bedarfsprognose, da die Verlagerung der Nachfrage auf alternative Packungsgrößen und konkurrierende Produkte, verzögerten Kaufhandlungen oder der Kauf in einer anderen Einkaufsstätte eine zusätzliche Verstärkung der Nachfragevolatilität verursachen [GMA et al. 2002]. Die weiter verringerte Prognosegenauigkeit erfordert wiederum noch höhere Sicherheitsbestände [APICS 2005, S. 2]. Entsprechend gilt die Vermeidung von Fehlbeständen als Hauptziel der Warendisposition, noch vor der Reduzierung der Lagerhaltungs- und Logistikkosten [EHI 2006]. Die wachsenden Bedarfsschwankungen verstärken sich zudem entlang der Supply Chain vom Konsumenten über den Handel bis hin zu Zulieferern in Form des so genannten Bullwhip-Effekts, der insbesondere in Industrieunternehmen zur Untersuchung neuartiger Modellen der Bedarfsprognose und Bestellmengenrechnung führte [Hansmann 2001].

Aus diesen Gründen wurden in Wissenschaft und Praxis vielfältige Modelle und Verfahren entwickelt, um die Güte der Warendisposition im Handel zu verbessern. Zu den bestehenden Ansätzen der Bedarfsprognose und der Bestellmengenrechnung auf einzelnen Lagerstufen wurden überwiegend Modelle zur Bestellmengenrechnung in mehrstufigen Lieferketten des Handels entwickelt. Diese zielen durch die Koordination von Beständen entlang den Lagerstufen des Handels bis zur Industrie und die Flexibilisierung der Distributionslogistik auf eine kostenoptimale Warenverfügbarkeit (vgl. Abschnitt 2.3). Empirische Analysen zu den Ursachen für Fehlmengen zeigen jedoch, dass 67 bis 85 % aller Fehlmengen durch Aktivitäten auf der letzten Lagerstufe der Handelsfiliale verursacht werden [ECR Europe und Roland Berger 2003]. Dies bestätigt sich auch in einer sinkenden Warenverfügbarkeit entlang der Supply Chain, die von 99 % im Fertigwarenlager der Industrie auf 98 bis 99 % im Zentrallager des Handels und auf 90 bis 93 % in der Handelsfiliale sinkt [ECR Europe und Roland Berger 2003, S. 15]. Die Probleme einer suboptimalen Warendisposition in der Handelsfiliale lassen sich jedoch nicht durch Ansätze der Flexibilisierung oder Beschleunigung der zuliefernden Logistikketten, Optimierung der Warenbestände in übergeordneten Lagerstufen oder

Informationsweiterleitung zur Vermeidung des Bullwhip-Effekts beheben [Aberdeen Group 2004, S. 4]. Somit zeigen Ansätze zur Bedarfsprognose und zur Bestellmengenrechnung in der Filiale das größte Potenzial für Verbesserungen der Dispositionsentscheidungen, Warenbestände und somit Servicegrade im Handel.

In den Handelsfilialen führt jedoch eine einfache Erhöhung der Sicherheitsbestände nicht zu einer erhöhten Warenverfügbarkeit im Regal oder einer Vermeidung von Fehlmengen. Im Gegenteil belegen Studien der Handelspraxis eine positive Korrelation zwischen der Höhe der Sicherheitsbestände und der Anzahl von Fehlmengen, die auf die steigende Komplexität der Zwischenlagerung und der Warenverräumung im Regal zurückgeführt wird [GMA et al. 2002, S. 46]. Verbesserte theoretische Modelle der Bestellmengenrechnung unter Verwendung spezifischer Verteilungsformen des Bedarfs sowie Mehrproduktmodelle zur Berechnung effektiver Sicherheitsbestände sind zudem weitgehend etabliert und bieten nur geringes Verbesserungspotenzial.

Aus diesen Gründen gewinnen Ansätze zur Verbesserung der Bedarfsprognose immer mehr an Bedeutung. ECR Europe identifiziert die inkorrekte Prognose von Standard- und Aktionswaren auf der Ebene der Handelsfiliale als einen Hauptverursacher von Ineffizienzen [ECR Europe und Roland Berger 2003, S. 33]. Dabei können 35 bis 47 % aller Fehlmengen allein auf fehlerhafte Bedarfsprognosen in den Filialbestellungen zurückgeführt werden [GMA et al. 2002; ECR Europe und Roland Berger 2003; Stölzle und Helm 2005].¹ Diese Erkenntnis wird auch durch eine Studie der Bundesvereinigung Logistik gestützt, die als Hauptursache für eine mangelnde Warenverfügbarkeit „unzureichende Prognosefähigkeit“ nennt [Straube et al. 2005]. Nach Einschätzung des US-Bundesverbands beruhen von den 100 Mrd. US-Dollar an Beständen in Fertigwaren 50 % allein auf Fehlern in der Absatzprognose [GMA und IBM 2005]. Dies ist darauf zurückzuführen, dass Prognosen der zukünftigen Bedarfsmengen und Bedarfszeitpunkte die Güte der Bestellmengenrechnung und somit die Güte der Warendisposition in zweierlei Hinsicht beeinflussen: (1.) Die Genauigkeit der Punktprognose bestimmt das Niveau der Bestellmenge und somit die Höhe der Abweichung vom tatsächlichen Bedarf. (2.) Die Prognosegenauigkeit vorangegangener Perioden bestimmt durch die Streuung historischer Prognosefehler die Höhe der Sicherheitsbestände, welche Schwankungen in der Nachfrage und Ungenauigkeiten der Prognose ausgleichen sollen.

Aufgrund der Bedeutung der Prognosegenauigkeit auf die Entscheidungsgüte, auch in anderen Disziplinen, wurde eine Vielzahl quantitativer Prognoseverfahren zur Abbildung linearer sowie nichtlinearer Prognosemodelle entwickelt, die sich durch eine unterschiedliche Genauigkeit, Fehlerrobustheit, Komplexität und Angemessenheit für unterschiedliche Formen von Zeitreihen auszeichnen.² Aufgrund der besonderen Struktur der Absatzzeitreihen im Handel können viele dieser Verfahren jedoch nicht valide, reliabel und automatisiert angewendet werden. Die Zeitreihen im Handel sind aufgrund der hohen Artikelfluktuation und vielfacher Ein- und Auslistungen kurz und von hoher Volatilität. Dies geht einher mit hoher Instationarität durch Niveauverschiebungen, lokale Trends und Ausreißer, Heteroskedastizität der Nachfrage und mehrfacher, sich überlagernder Saisonalitäten. Viele Absatzmuster zeigen

¹ Suboptimale Dispositionsentscheidungen sind jedoch nicht die einzigen Ursachen für Präsenzlücken und Überbestände im Handel. STÖLZLE und HELM {, 2005, sfc-39d520a21} identifizieren ferner Leistungsdifferenzen, Unzulänglichkeiten bei der Regalbefüllung sowie Liefer- und Platzierungsprobleme als Ursachen. Ferner existieren Interdependenzen zu Prozessen der Filiallogistik, zu den Lagerstrukturen, zur Infrastruktur der Informationssysteme sowie zu der Aufbau- und Ablauforganisation [Loderhose 2006]. Da eine Verbesserung der Bedarfsprognose und der Dispositionsgenauigkeit auf die vorangehend genannten Aspekte jedoch keinen Einfluss hat, wird nachfolgend nicht weiter auf sie eingegangen.

² Neben den etablierten Verfahren der Exponentiellen Glättung und der Autoregressive Integrated Moving Average (ARIMA) Modelle finden auch nichtlineare Verfahren der Smooth Transition Autoregression (STAR) oder der Autoregressive Conditional Heteroskedasticity (ARCH) für heteroskedastische Zeitreihen weite Beachtung bei Prognosen auf Aktien- und Finanzmärkten (vgl. Abschnitt 3.2)

zudem signifikante Einflüsse durch externe Einflussfaktoren in Form der Kommunikations-, Preis- und Sortimentspolitik in Handel und Industrie, Kalendereffekten durch Feiertage und Urlaubszeiten mit individuellem zeitlichem Vorlauf und Verzögerungen der Effekte sowie durch Wettereinflüsse. Die sich überlagernden Einflussgrößen sind vielfach nichtlinearer Natur und erlauben daher keine Modellierung der Systemzusammenhänge mittels linearer Erklärungsmodelle. Die Absatzzeitreihen von Artikeln zeigen selbst innerhalb einer Warengruppe heterogene Bedarfsmuster und selektive Abhängigkeiten von externen Effekten, die zudem für einen Artikel je nach Absatzstelle unterschiedlich ausfallen können. Aufgrund der Komplexität der Interaktionen und ihrer Heterogenität existieren vielfach keine A-priori-Annahmen über ihre Struktur selbst innerhalb einer Warengruppe, was die Bildung expliziter Erklärungsmodelle zusätzlich erschwert.

Zudem erfordert die tägliche oder wöchentliche, absatzstellen- und produktspezifische Warendisposition einer Vielzahl von Verkaufseinheiten je Absatzstelle die weitgehende Automatisierung der Prognose. Viele Prognoseverfahren können aber die Komplexität der Zeitreihen und Erklärungsmodelle nicht automatisiert abbilden. Der Handel setzt daher vielfach weiterhin auf einfache zeitreihenanalytische Verfahren der Mittelwertbildung und der Exponentiellen Glättung, vielfach sogar ohne Berücksichtigung von regelmäßig auftretenden Wochentags- oder Monatssaisonalitäten oder kausalanalytischen Einflüssen externer Faktoren.¹ Einer Befragung von 560 Absatzplanern zufolge nutzen lediglich 38 % Prognosemodelle die Saisonalität abbilden können, und nur 6 % kausalanalytische Verfahren zur Verwendung exogener Erklärungsvariablen [Jain 2003, S. 60-64].

Als Alternative zu etablierten statistischen Verfahren können auch neuartige Ansätze der Künstlichen Intelligenz zur Bedarfsprognose verwendet werden. Künstliche Neuronale Netze (NN)² bilden eine Klasse mathematisch-statistischer Verfahren, deren Funktionsweise dem Vorbild der neurobiologischen Informationsverarbeitung in Nervenzellen und Nervenverbänden des Menschen nachempfunden wurde. Als semiparametrische, datengetriebene und nichtlineare Verfahren weisen NN insofern attraktive Eigenschaften für die Prognose auf, als sie die Fähigkeit besitzen, jede beliebige Funktion bis zu einem beliebigen Grad der Genauigkeit zu approximieren. Damit eignen sie sich sowohl zur Modellierung linearer Zusammenhänge als auch subtiler nichtlinearer funktionaler Zusammenhänge der zeitreihenanalytischen und kausalanalytischen Prognose, die von gängigen statistischen Verfahren nicht erfasst werden können, jedoch in der betrieblichen Praxis überwiegen. Als datengetriebene Verfahren erlernen NN die der Zeitreihe zugrunde liegende komplexe Modellstruktur direkt aus historischen Beobachtungswerten ohne A-priori-Annahmen über die Relevanz und die Interaktion der unabhängigen Variablen. Somit formulieren NN auch dann valide Prognosemodelle, wenn den Experten keine formalen Erkenntnisse über die tatsächliche Modellform der Zeitreihen vorliegen – etwa hinsichtlich der Interaktion von Absatz, Werbung, Preis und Kalendereffekten.³ Darüber hinaus besitzen NN die Fähigkeit zur Generalisierung der erlernten Erklärungsmodelle, in dem sie den Systemzusammenhang auf neue, bisher unbekannte Beobachtungen in der Zukunft im Sinne der Prognose übertragen. Sie umgehen auf diese Weise die Probleme der Modellspezifikation, der Modellselektion und der Parameterbe-

¹ Kalendereffekte, Preis- und Werbemaßnahmen, vergangene Fehlmengen oder Wettereinflüsse sind somit manuell durch den Disponenten zu korrigieren und für zukünftige Perioden zu berücksichtigen.

² Die Bezeichnung des Verfahrens mit dem Kürzel „NN“ erfolgt in Übereinstimmung mit der englischsprachigen Literatur zur Abgrenzung gegen die etablierte Bezeichnung „KNN“ der „K-nearest-neighbour“-Algorithmen der Clusteranalyse. In der Einleitung wird aus Gründen der Übersichtlichkeit nicht zwischen NN im Allgemeinen und der Verfahrensklasse der „Multilayer Perceptrons“ (MLP), die in dieser Arbeit vorwiegend untersucht werden, unterschieden.

³ Darüber hinaus werden ihnen weitere Eigenschaften der Fehlerrobustheit, Unanfälligkeit gegenüber Störungen nachgesagt, die jedoch nicht universell und insbesondere nicht im Kontext der Absatzprognose bestätigt werden können.

stimmung statistischer Verfahren sowie die Probleme der formalen Modellbildung bei den Expertensystemen.

Die Einsatzfähigkeit von NN für Prognosen wurde seit Beginn der 1990er-Jahre in unterschiedlichen Anwendungsbereichen experimentell demonstriert. Trotz ihrer theoretischen Fähigkeiten sind NN jedoch bislang nicht als Prognoseverfahren in der betriebswirtschaftlichen Praxis etabliert, im Gegensatz zu ihrer breiten Akzeptanz als Verfahren der Klassifikation (vgl. Abschnitt 4.4). Die wenigen wissenschaftlichen Studien zur Absatz- und Bedarfsprognose zeigen nur uneinheitliche Ergebnisse hinsichtlich ihrer Prognosegüte und Generalisierungsfähigkeit. Neue Veröffentlichungen haben zudem erhebliche Zweifel an den Fähigkeiten von NN bei der Prognose selbst einfacher saisonaler oder trendbehafteter Zeitreihen ohne adäquate Vorverarbeitung aufkommen lassen [Hill et al. 1996; Zhang und Qi 2005]. Meta-Studien zeigen jedoch, dass der Großteil der publizierten Studien aufgrund von mangelhaften, unzureichend dokumentierten Experimentaufbauten sowie dem mangelndem Vergleich der Ergebnisse mit der Genauigkeit etablierter statistischer Verfahren nur eingeschränkt valide und reliable Erkenntnisse zulassen [Adya und Collopy 1998; Crone und Preßmar 2006; Liao und Fildes 2005]. Ferner ist anzuführen, dass Erkenntnisse über die Prognosegüte eines Verfahrens nicht uneingeschränkt auf die Zeitreihen einer anderen Anwendungsdomäne übertragbar sind, sodass die besondere Struktur der verrauschten, kurzen, saisonalen und trendbehafteten Zeitreihen im Handel eine eigenständige Untersuchung der Fähigkeiten und der Güte von NN erfordert.

Über die besonderen Fähigkeiten zur nichtlinearen, datengetriebenen Prognose bieten NN theoretisch weiteres Potenzial für Effizienzsteigerungen in der Warendisposition. Ebenso wie viele lineare statistische Verfahren der Regression finden NN adäquate Prognosemodelle und -parameter durch die Minimierung einer quadratischen Zielfunktion, obwohl sie nicht durch eine zugrunde liegende statistische Theorie an quadratische Prognosefehler gebunden sind. Die Zielfunktion ist insofern eine bedeutende Determinante von NN, als sie mittels Bewertung der Fehlerkosten die Zielsetzung und somit das Verhalten des NN als Prognoseverfahren wesentlich bestimmt. Sie dominiert insbesondere die Entscheidungen über die Architektur, die Topologie und die Lernparameter in der Modellierung von NN, die lediglich das erfolgreiche Minimieren der Zielfunktion auf den Daten ermöglichen sollen. Die nahezu ausschließliche Verwendung der quadratischen Zielfunktion der l_2 -Norm zum Training von NN ist vielmehr auf die mathematische Einfachheit, die historische Entwicklung der Lernverfahren der NN sowie die etablierte Nutzung der Kleinste-Quadrate-Schätzer in der Statistik zurückzuführen [Reed und Marks 1999].

Vereinzelt wurde die Nutzung abweichender statistischer Fehlermaße zur Prognose mit NN vorgeschlagen, so zum Beispiel die Funktion der absoluten Fehler der l_1 -Norm [Bishop 1995] sowie die ε -insensitive Verlustfunktion zur robusten Prognose [Vapnik 1999; Reed und Marks 1999], die beide symmetrisch um Null verlaufen. Da in der betriebswirtschaftlichen Praxis die Kosten fehlerhafter Entscheidungen vielfach quantifizierbar sind, können diese unmittelbar als Ziel- bzw. Kostenfunktion modelliert werden. In der zu betrachtenden Anwendung der Warendisposition steigen die Kosten einer Überprognose oder Unterprognose in der Regel nicht symmetrisch und nicht quadratisch mit dem Betrag des Prognosefehlers. Aufgrund der empirischen Kostenstruktur im Handel führen Fehlmengen durch zu niedrige Prognosen zu höheren Kosten als Überbestände durch zu hohe Prognosen. Können diese Kosten auf der Filialebene für jedes Produkt geschätzt werden, erlaubt dies die Bildung von asymmetrischen, nichtquadratischen Kostenfunktionen, die als asymmetrische Zielfunktionen die tatsächlichen betrieblichen Entscheidungskosten von positiven und negativen Prognosefehlern aus der Bestandsentscheidung der Warendisposition repräsentieren. Dies ermöglicht theoretisch das Training der NN mit betriebswirtschaftlichen Zielfunktionen, und somit eine abweichende

Form der Entscheidungsfindung. Durch die neuartige Zielfunktion wird ein NN auf die Prognose des Warenbestands mit dem minimalen Erwartungswert der Entscheidungskosten für einen zukünftigen Dispositionszeitpunkt trainiert. Das NN schätzt auf diese Weise ein durch die asymmetrischen Kosten spezifiziertes Quantil auf der kumulativen Verteilungsfunktion der Prognosefehler und nicht wie andere statistische Verfahren den bedingten Erwartungswert des arithmetischen Mittels. Somit erlernt das NN die Vorhersage eines kostenminimalen Warenbestandes unmittelbar aus den vergangenen Beobachtungswerten der Absatzzeitreihen. Unter Verwendung einer asymmetrischen Kostenfunktion aus den tatsächlichen Entscheidungskosten kann die Funktionsweise des NN somit zu einem integrierten Entscheidungsmodell der simultanen Bedarfsprognose und Bestandsberechnung in der Warendisposition erweitert werden. Somit wird die herkömmliche Verwendung von NN als Prognoseverfahren auf die unmittelbare Lösung geeigneter betrieblicher Entscheidungsprobleme durch die Formulierung asymmetrischer Zielfunktionen generalisiert, und somit eine neuartige Form der Modellbildung und Entscheidungsunterstützung durch die Verfahrensklasse der NN erschlossen.

Das Ziel der vorliegenden Arbeit ist die Ermittlung der Eignung und der Güte von NN bei der betrieblichen Bedarfsprognose und Warendisposition. Dabei steht die Entwicklung von NN für die Disposition mithilfe von asymmetrischen, betriebswirtschaftlichen Zielfunktionen der Entscheidungskosten, und ihre Beurteilung im Vergleich zu konventionellen Fehlerfunktionen und statistischen Verfahren, im Mittelpunkt.

1.2 Gang der Untersuchung

Die vorliegende Untersuchung ist in sieben Kapitel gegliedert, deren Inhalt und Bedeutung im Folgenden kurz aufgezeigt wird.

Im Anschluss an diese Einleitung werden in Kapitel 2 zunächst die wesentlichen Grundlagen der Bedarfsprognose und der Bestellmengenrechnung für die Warendisposition im Handel vorgestellt. Dabei wird die Warendisposition in die operativen Kernprozesse der Beschaffung, Lagerung und Distribution eingeordnet und der interdependente Charakter von Bedarfs- und Bestellmengenrechnung als sequenzielle Teilprozesse hervorgehoben. Nach einer Einführung in die grundlegenden Begrifflichkeiten der Warendisposition folgt eine Erörterung deterministischer und stochastischer Modelle der Bedarfs- und Bestellmengenrechnung. Aufgrund der besonderen Bedeutung der Bedarfsprognose bei der Disposition in Handelsfilialen beschränkt sich die weitere Diskussion auf ein grundlegendes, einstufiges, einperiodiges Lagerhaltungsmodell für ein Produkt zur Bestellmengenrechnung bei stochastisch-instationärem Bedarf, das sogenannte Zeitungsjungen- oder Newsboy-Modell. Im Anschluss wird sodann die Bedeutung der Prognosegüte in Bezug auf die Güte der Warendisposition theoretisch sowie anhand empirischer Beispiele herausgearbeitet.

In Kapitel 3 werden die Grundlagen, Modelle und Verfahren der Bedarfsprognose im Detail analysiert. Da in der Literatur die Begriffe und Determinanten der Prognose uneinheitlich verwendet werden, erfolgt ihre ausführliche Besprechung unter Verwendung der Wirtschaftswissenschaftlichen Literatur für eine spätere Einordnung und Diskussion der NN als Prognoseverfahren. Es folgt eine Diskussion objektiver Prognoseverfahren, geordnet nach ihrer Modellform für stationäre, trendbehaftete, saisonale und trendsaisonale Zeitreihenverläufe, um die Herausforderungen einer zeitreihenspezifischen Modellbildung und Verfahrensauswahl in der Anwendung konventioneller statistischer Prognoseverfahren der Zeitreihenanalyse zu verdeutlichen. Diese Struktur wird zudem in der sich anschließenden

Herleitung eines prototypischen Prognoseprozesses für statistische Verfahren vertieft, die eine Abgrenzung zu dem semiparametrischen Ansatz der Prognose mit NN zulässt. Zum Vergleich der konkurrierenden Erklärungsmodelle und Prognoseverfahren werden im Anschluss statistische und deskriptive Fehlermaße systematisiert und ausführlich diskutiert. Die Systematik der Gütekriterien bildet den Rahmen für die spätere Herleitung betriebswirtschaftlicher Zielfunktionen der Entscheidungskosten und erlaubt einen direkten Bezug zu den Kostenstrukturen statistischer Fehlermaße.

Gegenstand des 4. Kapitels ist die Modellierung von NN zur Bedarfsprognose. Die Grundlagen der NN werden hinsichtlich ihres Aufbaus, ihrer Architektur und ihrer Topologie, der Funktionsweise ihrer Informationsverarbeitung und ihrer Lernverfahren einschließlich der historischen Entwicklung aus dem biologischen Vorbild ausführlich vorgestellt. Dies erfolgt um in Hinblick auf die nachfolgende Weiterentwicklung der NN zu verdeutlichen, dass keine architektonische oder modelltheoretische Begründung für eine Beschränkung auf quadratische Zielfunktionen in der Anwendung von NN zur Prognose besteht. Im Anschluss an die Erörterung der Grundlagen von NN, die durch eine entsprechende Konfiguration sowohl zur Vorhersage zukünftiger Werte als Klassifikation, Regression oder Clusteranalyse angewendet werden könnten, erfolgt die Spezifizierung der NN zur Punktprognose im Sinne der Regression. Neben der theoretischen Eignung von NN zur Prognose werden empirische Anwendungsbereiche und die Prognosegüte von NN auf der Basis einer repräsentativen Literaturanalyse diskutiert. Nachfolgend wird eine Übersicht über die Modellierung von NN als Verfahren der zeitreihenanalytischen sowie kausalanalytischen Bedarfsprognose gegeben, und es werden die Freiheitsgrade ihrer Architektur durch die Möglichkeiten zur Modellierung von etablierten statistischen Prognoseverfahren durch NN aufgezeigt.

In Kapitel 5 wird sodann ein neuartiger Ansatz des Trainings von NN zur Entscheidungsfindung mithilfe betriebswirtschaftlicher Zielfunktionen vorgestellt. Ausgehend von den konventionellen statistischen Fehlerfunktionen der quadratischen Fehler sowie neuerer Erweiterungen robuster Verlustfunktionen werden asymmetrische Kostenfunktionen als Zielfunktionen der Parametrisierung von NN aus den empirischen Fehlerkosten der betrieblichen Dispositionsentscheidung hergeleitet. Nach einer methodischen Erweiterung des Backpropagation-Lernalgorithmus auf betriebswirtschaftliche Kostenfunktionen folgt eine experimentelle Überprüfung der Eigenschaften und Fähigkeiten von NN mit asymmetrischen Kostenfunktionen bei der Disposition unter Verwendung von synthetischen Zeitreihendaten. Das NN minimiert, wie gezeigt werden wird, die asymmetrische Kostenfunktion valide und reliabel und kann daher unmittelbar betriebliche Entscheidungen wie die der Warendisposition in Form einer simultanen Bedarfsprognose und Bestellmengenrechnung treffen.

Zur Beurteilung der praktischen Einsatzfähigkeit und der Güte von NN bei der Disposition mit asymmetrischen Kostenfunktionen folgt in Kapitel 6 eine ausführliche Untersuchung anhand eines empirischen Vergleichs. Um mögliche Zugewinne der Dispositionsgüte durch die neuartigen asymmetrischen Kostenfunktionen zu identifizieren werden zudem NN zur konventionellen Prognose modelliert. Die Ergebnisse der NN aus der Bedarfsprognose und der Warendisposition auf 51 Zeitreihen über 21 Wochen werden einer Auswahl von etablierten Prognoseverfahren und Expertensystemen aus der Praxis gegenübergestellt und hinsichtlich ihrer Prognosegüte und ihrer Dispositionsgüte beurteilt.¹ Da NN bei der Prognose und bei der Disposition sowohl eine zeitreihenanalytische als auch eine kausalanalytische

¹ Obwohl Prognosen nicht zum Selbstzweck durchgeführt werden, sondern zur Unterstützung betrieblicher Entscheidungen, erfolgt die wissenschaftliche oder empirische Analyse sowie die Weiterentwicklung neuartiger Prognoseverfahren häufig losgelöst vom originären Entscheidungskontext anhand einfacher statistischer Fehlermaße. Eine abschließende Beurteilung der Güte eines Verfahrens – selbst auf einem einzelnen spezifischen Datensatz – kann nur unter Berücksichtigung der Entscheidungskosten erfolgen.

Modellierung zulassen, werden beide Ansätze berechnet und mit statistischen Verfahren der Kausalanalyse verglichen. Somit lassen sich sowohl Erkenntnisse zur Prognosegüte statistischer Verfahren im Vergleich zu NN treffen, als auch Erkenntnisse aus dem Vergleich von zeitreihenanalytischen zu kausalanalytischen Prognosemodellen, Erkenntnisse der Anwendung eines Verfahrens auf alle Zeitreihen im Vergleich zu Ansätzen der zeitreihenspezifischen Modellselektion durch Expertensysteme. Die Beurteilung der Verfahrensgüte erfolgt durch die Messung der erzielten Prognosegenauigkeit anhand mehrerer statistischer Fehlermaße losgelöst vom Entscheidungskontext der Disposition, als auch der Güte der Dispositionsentscheidung gemessen anhand der resultierenden Kosten aus den Bestellmengen. Schließlich erlaubt der Verfahrensvergleich eine Untersuchung der Auswirkungen der Prognosegüte auf die Dispositionsgüte bei stochastisch-instationärem Bedarf, und lässt Aussagen über Interdependenz von Prognosefehlern und Dispositionskosten sowohl für statistische Verfahren als auch für NN mit unterschiedlichen Zielfunktionen zu.

In Kapitel 7 werden die wesentlichen Ergebnisse der Untersuchung zusammenfasst und ein Ausblick auf weiterführende Untersuchungen und Einsatzmöglichkeiten von NN bei der Prognose und bei der Disposition gegeben.

2 Prognose und Disposition im Handel

Im Folgenden wird zunächst der Begriff der Warendisposition in den Betrachtungsgegenstand des Handels eingeordnet und gegenüber anderen Verwendungen abgegrenzt. Bei der Betrachtung der Kernprozesse, der Lagerstrukturen und des Funktionsmodells der Warendisposition wird die besondere Bedeutung der Bedarfs- und Bestellmengenrechnung im Handel verdeutlicht.

Anschließend werden die Determinanten und Modelle der Bedarfsrechnung für deterministischen und stochastischen Bedarf analysiert, gefolgt von einer Diskussion grundlegender Modelle zur Bestellmengenrechnung. Die Determinanten des Bedarfs werden im Detail analysiert, da ihre Ausprägungen sowohl die Modellauswahl der Bedarfsrechnung als auch der Bestellmengenrechnung bestimmen. Dabei erhalten die Modelle des stationären sowie des instationären stochastischen Bedarfs besondere Aufmerksamkeit, um die Relevanz der Bedarfsprognose für die Warendisposition herauszustellen.

Abschließend werden deskriptive Gütekriterien zur Beurteilung von Prognoseverfahren in der Warendisposition untersucht und die Auswirkungen von Fehlern in der Bedarfsprognose auf die Entscheidungsqualität der Bestellmengenrechnung analytisch entwickelt, bevor diese in einem späteren Kapitel empirisch überprüft wird.

2.1 Grundlagen der Warendisposition

2.1.1 Definition und Einordnung der Warendisposition

Wenngleich der Begriff Disposition im Handel unmittelbar mit den Vorgängen der Beschaffung und Lagerhaltung assoziiert wird, ist er in der Betriebswirtschaftslehre weiter gefasst. Daher erfolgt zunächst eine definitorische Einordnung der Disposition in die Bereiche Materialwirtschaft, Logistik und Beschaffung.

In der Betriebswirtschaftslehre bezeichnet die dispositive Arbeit grundsätzlich einen betrieblichen Produktionsfaktor, der mit den Elementarfaktoren der ausführenden Arbeit, der Betriebsmittel und der Werkstoffe im Betrieb kombiniert wird [Gutenberg 1990, S. 43ff.].¹ Die dispositive Arbeit gestaltet den Einsatz der genannten Faktoren durch Planung, Organisation und Kontrolle, in Abgrenzung zur ausführenden bzw. objektbezogenen Arbeit am Erzeugnis.² Im Sinne der Planung bezeichnet die betriebliche Disposition eine Entscheidung zur situationsabhängigen Regelung eines Einzelfalls im Gegensatz zu einer Improvisation oder Ausnahmeregelung [Arentzen 1993, S. 816]. Vor diesem Hintergrund betrachten VOß und GUTENSCHWAGER die Dispositionsentscheidung als eigenständigen Prozess der Planung und Entscheidung „mit dem Ziel der Auswahl und Realisation einer Handlungsalternative unter Unsicherheit“ [Voß und Gutenschwager 2001, S. 122]. Die Disposition bildet folglich in der entscheidungsorientierten Betriebswirtschaftslehre eine bedeutende Unternehmensteilfunktion [Wöhe und Döring 1990, S. 91-93], deren Betrachtungsgebiet weit über die handelsbetrieblichen Prozesse der Warenbeschaffung und -bereitstellung hinausgeht und die somit einer schrittweisen Eingrenzung bedarf.

¹ Etymologisch bedeutet Disposition sinnvolle Einteilung und Planung [Kluge 2002, S. 185; Klappenbach und Steinitz 1967-1977, S. 120] oder auch Verfügung [Bichler und Lorsch 1985, S. 648].

² Vereinzelt wird der dispositive Faktor zudem in den originären Faktor der Unternehmensführung und den derivativen Faktor der Planung, Organisation und Kontrolle unterteilt [Fricke und Rube 2001], ein Aspekt, der hier jedoch vernachlässigt wird.

Im Rahmen dispositiver Entscheidungen des Elementarfaktors Werkstoffe¹ wird der Begriff Disposition synonym mit der operativen Bedarfsermittlung und Bedarfsdeckung von Materialien im Sinne der Materialdisposition verwendet [Stickel 1997, S. 199; Tempelmeier 2003, S. 2]. „Unter Materialdisposition versteht man die kurzfristige Planung der Versorgung ... mit Erzeugnis- und Betriebsstoffen in quantitativer und terminlicher Hinsicht unter Beachtung von Kostenkriterien“ [Arnolds et al. 2001, S. 41]. Sie hat als Teilfunktion der Materialwirtschaft die Hauptaufgabe, die Bedarfe an eigengefertigten und fremdbezogenen Teilen zur Deckung des Primärbedarfes einer bestimmten Zeitperiode zu ermitteln [Stickel 1997, S. 199].² Die Materialdisposition umfasst dementsprechend alle Tätigkeiten, die zur termingerechten und wirtschaftlichen Beschaffung und Bereitstellung der Materialien durch eine zielgerichtete Bestimmung von Bedarfsmenge und Bedarfszeitpunkt sowie von Bestellmenge und Bestellzeitpunkt erforderlich sind [Bichler und Lörsch 1985, S. 13]. Ist eine unmittelbare Bereitstellung nicht wirtschaftlich, umfasst und steuert die Disposition neben der Beschaffung auch alle Entscheidungen zur Lagerung der Materialien unter Berücksichtigung relevanter Randbedingungen einschließlich der Lagerbestände, Bestellbestände, Warenreservierungen, Auftragsrückstände, Lieferzeiten und Überprüfungszeiten. Somit nimmt die Materialdisposition in solchen Betrieben von Industrie und Handel eine zentrale Rolle ein, die Verbrauchsfaktoren im betrieblichen Leistungsprozess einsetzen und lagern. Da eine effiziente und effektive Materialdisposition wesentlich zum Unternehmenserfolg beiträgt [Bichler und Lörsch 1985], wird der Disposition vielfach die zentrale materialwirtschaftliche Planungsaufgabe zugeschrieben [Stickel 1997, S. 199].

Nach TEMPELMEIER [1992, S. 3] umfassen die Objekte der Materialwirtschaft die eingesetzten Roh-, Hilfs- und Betriebsstoffe, bezogene Teile in Form von vorbehandelten Materialien mit einem höheren Bearbeitungsgrad als Rohstoffe sowie Halb- und Fertigerzeugnisse, die als Bestandteile in die Enderzeugnisse eingehen oder zum Betrieb von Betriebsmitteln erforderlich sind. Davon sind die Sachgüter der Handelswaren zu unterscheiden, die von anderen Wirtschaftssubjekten produziert werden und nicht Bestandteil des Produktionsprogramms sind. Die Disposition dieser Handelswaren in Unternehmen des Handels sowie der Industrie, unter Fortfall von Interdependenzen zur Produktion, wird als Warendisposition bezeichnet. Nach ARNOLDS et al. [2001] bezeichnet die Warendisposition die kurzfristige Planung der wirtschaftlichen Versorgung mit Handelswaren in quantitativer und terminlicher Hinsicht. Diese Eingrenzung folgt einer funktionellen Definition des Handels als eigene wirtschaftliche Tätigkeit neben dem Produzieren und Konsumieren [Müller-Hagedorn 1998, S. 15]. „Handel im funktionellen Sinne liegt vor, wenn Marktteilnehmer Güter, die sie in der Regel nicht selbst be- oder verarbeiten (Handelswaren), von anderen Marktteilnehmern beschaffen und an Dritte absetzen“ [Ausschuss für Begriffsdefinitionen aus der Handels- und Absatzwirtschaft 1995, S. 28]. Die Warendisposition ist somit nicht auf die Institutionen der Handelsbetriebe beschränkt, und zudem eng mit den Problemstellungen der Materialdisposition und der Beschaffung in Produktionsbetrieben verwandt. Dies erfordert eine genauere Abgrenzung der Warendisposition hinsichtlich der sie betrachtenden wissenschaftlichen Disziplinen und der relevanten Literatur.

Da die Waren- und Materialdisposition im Schnittpunkt von Beschaffung, Logistik und Materialwirtschaft in Industrie und Handel einzuordnen ist, erfolgt eine uneinheitliche Verwendung der Terminologie ohne trennscharfe Abgrenzung von Begriffen [Bichler und

¹ Nach einer produktionswirtschaftlichen Betrachtung bezeichnen TEMPELMEIER [2003] sowie GROCHLA [1992] diese auch als Verbrauchsfaktoren, in Anlehnung an die Unterscheidung der Elementarfaktoren in Repetierfaktoren mit Verbrauchscharakter und Potenzialfaktoren mit mehrfachem Leistungspotenzial.

² Der Primärbedarf wird als Bedarf an Enderzeugnissen und verkaufsfähigen Zwischenprodukten berechnet und davon ausgehend die Bedarfsmengen und -termine der benötigten Baugruppen, Teile, Rohstoffe sowie Hilfs- und Betriebsstoffe als Sekundär- und Tertiärbedarfe ermittelt (vgl. Abschnitt 2.2.1.1).

Lörsch 1985; Müller-Hagedorn 1998, S. 497-498]. Abbildung 2.1 zeigt eine funktionsorientierte Einordnung in die vorgenannten Disziplinen nach TEMPELMEIER [1992].

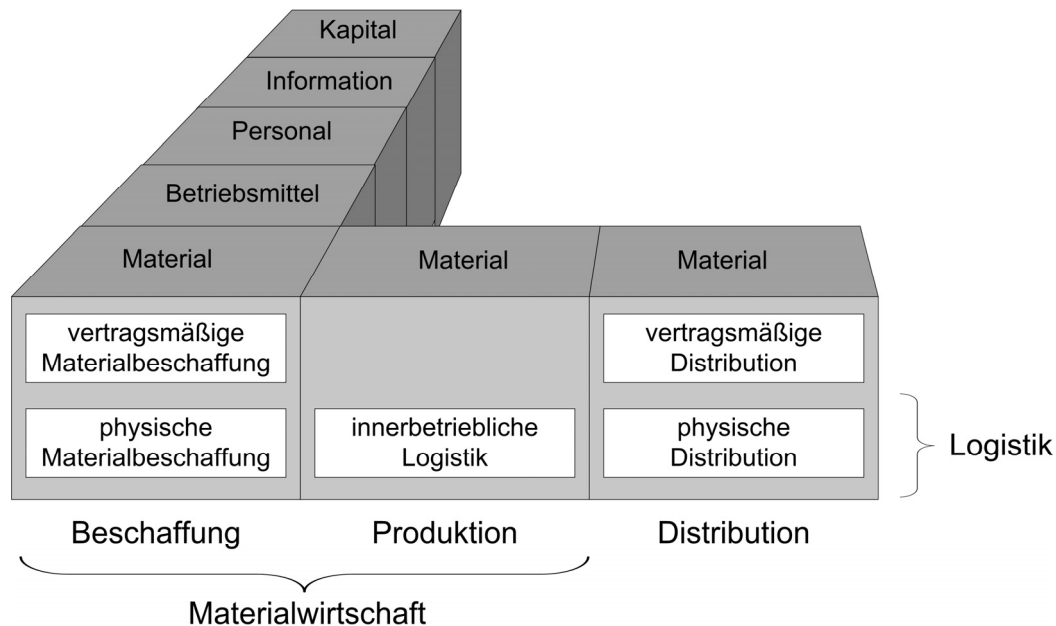


Abbildung 2.1: Die Warendisposition als Objekt der Logistik, Materialwirtschaft, Beschaffung, Produktion und Distribution¹

In Abbildung 2.1 sind die vorgenannten Elementarfaktoren dargestellt, wobei das Material als übergeordnete Klasse der Handelswaren in den Kernfunktionen Beschaffung, Produktion und Distribution dargestellt ist. In der Abbildung wird deutlich, dass die Warendisposition zur physischen Materialbeschaffung je nach Betrachtungsgegenstand als Objekt der Beschaffung, Logistik oder Materialwirtschaft, der Beschaffungs-Logistik oder auch der Material-Logistik betrachtet werden kann [Tempelmeier 1992, S. 2-6]. Die Disposition von Handelswaren kann daher nicht ausschließlich im Rahmen der Teildisziplin Handelslogistik betrachtet werden, sondern muss auch die Fragestellungen und Veröffentlichungen übergeordneter Disziplinen der Materialwirtschaft, Beschaffung und Logistik von Produktionsbetrieben berücksichtigen. Dabei sind jedoch die Besonderheiten der Warendisposition in den Abverkaufsstellen und Lagern des stationären Handels einzubeziehen, die nachfolgend näher untersucht werden.

2.1.2 Warendisposition als Kernprozess des Handels

In Abgrenzung zu der vorangehenden, funktionellen Definition der Handelstätigkeit folgen die anschließenden Ausführungen zur Disposition in Handelsbetrieben einer institutionellen Definition. Sie umfassen jene Institutionen, „deren wirtschaftliche Tätigkeit ausschließlich oder überwiegend dem Handel im funktionellen Sinn zuzurechnen ist“ [Ausschuss für Begriffsdefinitionen aus der Handels- und Absatzwirtschaft 1995, S. 28].² Unter Fortfall von Produktionsaufgaben gilt der Absatz von Handelswaren als die zentrale Aufgabe eines Handelsbetriebes [Müller-Hagedorn 1998, S. 493], wobei die Waren üblicherweise das bedeutendste Investitionsobjekt eines Handelsunternehmens repräsentieren [Götz 1999, S. 4]. Entsprechend rücken alle unmittelbar auf die Ware gerichteten Tätigkeiten in den Mittelpunkt

¹ Die Abbildung zeigt eine erweiterte Darstellung in Anlehnung an TEMPELMEIER [1992, S. 4], der die dargestellte Taxonomie der Elementarfaktoren nach GUTENBERG um einen fünften Elementarfaktor, nämlich den der Information, erweitert [vgl. u.a. Tempelmeier 1992; Gutenberg 1990]. Auf eine Diskussion dieser bedeutenden Erweiterung wird an dieser Stelle jedoch verzichtet.

² Auf eine detaillierte Auseinandersetzung mit den vielfachen Definitionen des Handelsbegriffes wird hier verzichtet. Für eine Diskussion siehe beispielsweise MÜLLER-HAGEDORN [1998, S. 16-25].

der handelsbetrieblichen Betrachtung; diese werden in der Summe als Warenwirtschaft bezeichnet.

MÜLLER-HAGEDORN [1998] untersucht die Warendisposition in Handelsbetrieben als Teilfunktion des Bestandsmanagements und der Handelslogistik, die den Fluss der Waren und der für ihren Transport benötigten Verpackungsmittel vom Lieferanten bis zum gewünschten Verwendungspunkt gestaltet. Obwohl Transport und Lagerung inklusive Rück- und Abtransport von Waren ebenfalls als Dispositionen über andere Faktoren wie Betriebsmittel, Personal, finanzielle Mittel etc. angesehen werden können, zählen diese Fragestellungen nicht zum Betrachtungsfeld der Warendisposition im eigentlichen Sinne [Müller-Hagedorn 1998, S. 497]. Ebenso werden weiterführende warenlogistische Fragestellungen des Lagermanagements, des Transports und der Auftragsverwaltung vernachlässigt, da sie nicht die Warendisposition im Sinne der physischen Materialdisposition betreffen. Somit ist die Warendisposition im Handel auf Fragestellungen des wirtschaftlichen Bestandsmanagements durch die Bestimmung der Entscheidungsparameter der Bestellmenge sowie des Bestell- und des Lieferzeitpunktes beschränkt [Müller-Hagedorn 1998, S. 498].

Eine nähere Betrachtung der Warendisposition im Schnittpunkt von Beschaffung, Logistik und Lagerung erfordert ihre Einordnung in die Kernprozesse des Handels. BECKER und SCHÜTTE [1996] charakterisieren Handelsbetriebe durch drei konstituierende Kernprozesse: Beschaffung, Lagerung und Distribution.¹ Diese sind in Abbildung 2.2 prozessorientiert unter Zuordnung der jeweiligen Hauptfunktionen und der ausführenden Informationssysteme als Handels-H-Modell dargestellt.

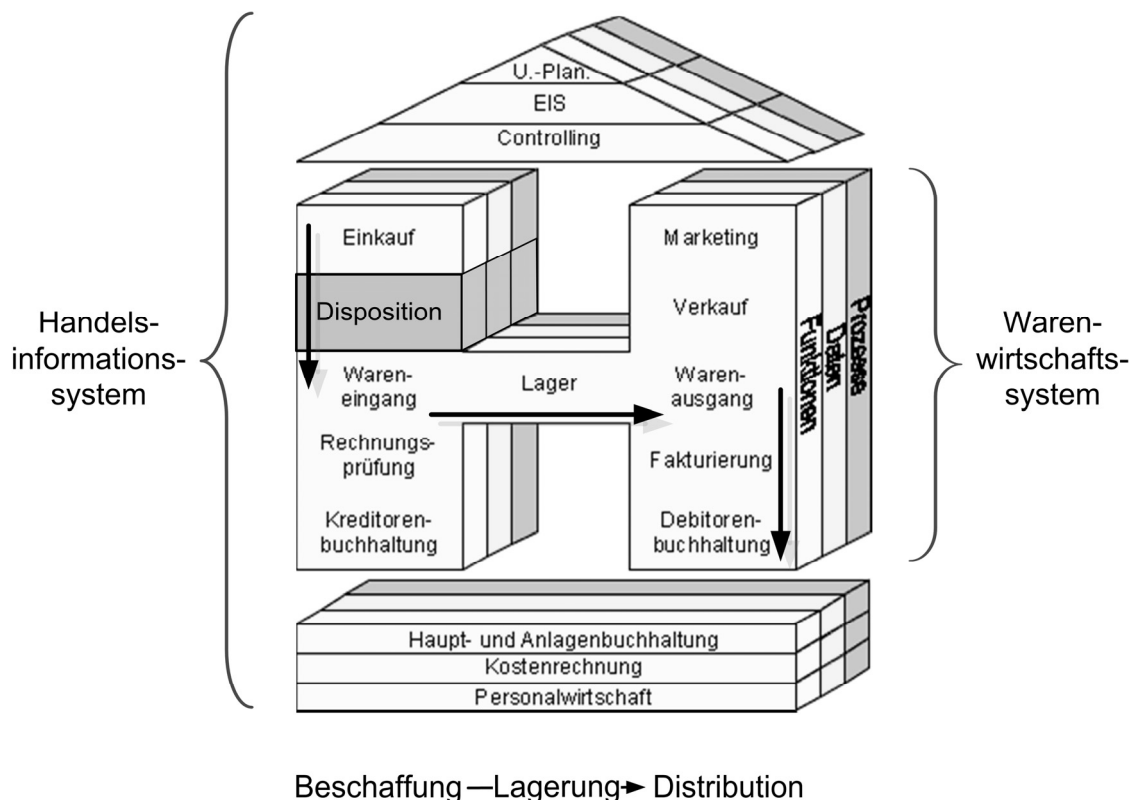


Abbildung 2.2: Kernprozesse und Hauptfunktionen von Handelsbetrieben²

¹ Als Prozess definieren sie eine inhaltlich abgeschlossene, zeitliche und sachlogische Abfolge von Funktionen, die zur Bearbeitung eines betriebswirtschaftlich relevanten Objektes notwendig sind [Becker und Schütte 1996].

² Erweiterte Darstellung in Anlehnung an BECKER und SCHÜTTE [1996, S. 123]. Die Pfeile symbolisieren die operativen warenlogistischen Prozesse.

Die vertikal dargestellten Kernprozesse der Beschaffung und Distribution werden durch den horizontalen warenlogistischen Kernprozess der Lagerung miteinander verbunden. Der Kernprozess der Beschaffung verfolgt die Aufgabe der „bedarfsgerechten und wirtschaftlichen Versorgung des Handelsunternehmens mit Waren“ [Becker und Schütte 1996, S. 148]. Diese operativen Kernprozesse werden durch administrative Supportprozesse und Funktionen der Buchhaltung, der Kostenrechnung und der Personalwirtschaft unterstützt und durch übergeordnete Systeme und Prozesse der Unternehmensplanung und des Controllings koordiniert.

BECKER und SCHÜTTE ordnen die Warendisposition in den Kernprozess der Beschaffung ein und bezeichnen sie als „Prozess von der Mengenplanung bis zur Auslösung einer konkreten Bestellung“ [Becker und Schütte 1996, S. 460]. Die administrativen Aufgaben der Einkaufsfunktion bilden die Grundlage für den operativen Prozess der Disposition innerhalb eines vorbestimmten Sortiments und gelisteter Lieferanten [Müller-Hagedorn 1998, S. 513]¹ sowie für die nachfolgenden beschaffungslogistischen Prozesse Wareneingang, Rechnungsprüfung und Kreditorenbuchhaltung.² Die Bedeutung der Warendisposition (in Abbildung 2.2 farblich hervorgehoben) geht aus ihrer Position in den Kernprozessen des Handels hervor: sie initiiert die lagerlogistischen Prozesse und ermöglicht so die distributionsseitigen Prozesse des Warenausgangs zur Bedarfsbefriedigung der Kunden. Durch die Warendisposition wird der Lagerbestand zeitlich und mengenmäßig durch die Bestimmung kostenoptimaler Bestellmengen und Bestellzeitpunkte anhand des Tripels „Artikel – Menge – Zeit“ in der Bestellung gesteuert. Der Waren- oder Lagerbestand gilt dabei als zentrale Größe aller beschaffungs- und handellogistischen Aktivitäten. Die Warendisposition gestaltet so den physischen Warenfluss in Intensität und Frequenz entlang der gesamten Wertschöpfungskette vom Lieferanten bis zum Endkunden [Becker und Schütte 1996, S. 148]. Disposition stellt somit die zentrale Entscheidungsfunktion in den Prozessen der Warenwirtschaft, der Beschaffung und der Logistik des Handels dar [Becker und Schütte 1996], wodurch ihr nach GROCHLA [1992, S. 17] eine herausragende Position zukommt.³

Innerhalb der Kernprozesse des Handels unterscheiden BECKER und SCHÜTTE [1996, S. 419] fünf etablierte Geschäftsarten, die durch das Auseinanderfallen der korrespondierenden Güter-, Geld und Informationsflüsse eine unterschiedliche Ausgestaltung der Prozesse der Warendisposition erfordern: Lagergeschäft, Streckengeschäft, Aktionsgeschäft, Dienstleistungsgeschäft und Zentralregulierungsgeschäft. Im Streckengeschäft erfolgt der Warenfluss direkt zwischen Lieferant und Kunde, während der dispositionsbezogene Informationsfluss zwischen Kunde und Handel erfolgt. Im Zentralregulierungsgeschäft werden der dispositionsbezogene Informationsfluss und der logistische Warenfluss unmittelbar zwischen Lieferant und Kunde abgewickelt; allein die Forderungsabwicklung erfolgt durch das Handelsunternehmen. Das Aktionsgeschäft unterscheidet sich durch den einmaligen Aktions-Charakter und die beschaffungs- und distributionssynchrone Planung von Einkauf und Verkauf deutlich von den logistischen Anforderungen der übrigen Geschäftsprozesse. Das Dienstleistungsgeschäft

¹ Vorgelagerte Fragestellungen der Disposition von Handelswaren im Rahmen der mittel- bis langfristigen Sortimentspolitik oder des Einkaufs stellen für Entscheidungen der Warenlogistik exogene Größen dar und werden daher nachfolgend vernachlässigt [Müller-Hagedorn 1998, S. 130].

² Organisatorisch wird die Warendisposition häufig durch Organisationseinheiten des Einkaufs durchgeführt, obwohl durch die wachsende Warenflussorientierung und die Konditionsvereinbarungen in Rahmenverträgen unabhängig von einzelnen Bestellmengen die Disposition als Bevorratungsfunktion eine stärker logistische und weniger einkaufsorientierte Natur aufweist [Becker und Schütte 1996, S. 173]. Eine Übersicht über die aufbauorganisatorischen Aspekte der Disposition geben [Bierhals 1993; Lensing und Sonnemann 1995].

³ Allerdings bezieht GROCHLA [1992] seine Aussage auf den materialwirtschaftlichen Bereich von Sachleistungsbetrieben in Form von Fertigungs-, Veredelungs- und Gewinnungsbetrieben, weniger auf den warenwirtschaftlichen Bereich von Dienstleistungs- und Handelsbetrieben. Unter Berücksichtigung des höheren Anteils der Dispositionsprozesse an den Unternehmensfunktionen und -prozessen im Handel kann die Gültigkeit seiner Aussage jedoch ebenso für Handelsbetriebe angenommen werden.

unterscheidet sich von den anderen Geschäftsfällen durch das abweichende Objekt der Dienstleistung, zum Beispiel in Form von Leistungen der Beratung oder der Verkaufsförderung, das zusätzlich zu den eigentlichen Handelsfunktionen angeboten wird. Von allen Geschäftsarten repräsentiert das Lagergeschäft „die wichtigste Geschäftsart, mit der der Handel die Aufgabe der mengenmäßigen, zeitlichen und räumlichen Überbrückung erfüllt“ [Becker und Schütte 1996, S. 147]. Entsprechend konzentrieren sich die nachfolgenden Untersuchungen ausschließlich auf die Funktionen, Modelle und Verfahren des Lagergeschäftsprozesses, der im Anschluss an eine Darstellung der Grundlagen von Lagersystemen detailliert dargestellt wird.¹

Aus Gründen der Anschaulichkeit und der begrifflichen Klarheit wird in dieser Arbeit nur die Lagerung physischer Güter in Handelsbetrieben untersucht. Die Beschränkung auf Dispositionsentscheidungen in Handelsbetrieben hat gemäß ASSFALG [1976, S. 14] mehrere Vorteile: (1.) In Handelsbetrieben ist die Warendisposition von besonderer Bedeutung, weil der größte Teil des Betriebskapitals im Lager gebunden ist und dort die meisten Kosten verursacht werden. (2.) Warendisposition und Lagerhaltung werden keinem anderen Bereich nachgeordnet, etwa dem Produktionsbereich in Industriebetrieben. Die für Dispositionsentscheidungen typischen Einflussfaktoren können in Handelsbetrieben daher besser isoliert und kontrolliert werden als in Produktionsbetrieben, in denen die besondere Bedeutung des Produktionsprozesses die Entscheidungen der Warendisposition einengt. (3.) Die Nachfrage im Lagergeschäftsprozess des Handels ist stochastisch und instationär. Sie erfordert besondere Berücksichtigung durch Modelle und Verfahren der Absatzprognose. Ohne zu vernachlässigen, dass sich in den Lagerstufen der Industrie grundsätzlich ähnliche Probleme stellen, beziehen sich die folgenden Ausführungen ausschließlich auf das Problem der Prognose und Warendisposition in Handelsbetrieben.

2.1.3 Dispositions- und Lagerstrukturen im Handel

Lager bilden in allen Wirtschaftsformen einen Puffer zwischen eingehenden (engl. *input flows*) und ausgehenden Güterströmen (engl. *output flows*) mit abweichender zeitlicher und quantitativer Struktur und erzielen so eine Entkopplung der Ströme [Schneeweiß 1981, S. 3-4; Hochstädter 1969, S. 1]. In Abbildung 2.3 sind die Konstituenten eines einstufigen Lagerhaltungssystems mit einem Kunden, einem Unternehmen mit einem einzelnen Lager und einem Lieferanten dargestellt. Die Eingangs- und Ausgangsströme zeigen sowohl den Warenfluss vom Lieferanten zum Kunden als auch den entgegengesetzt verlaufenden Informationsfluss vom Kunden zum Lieferanten.

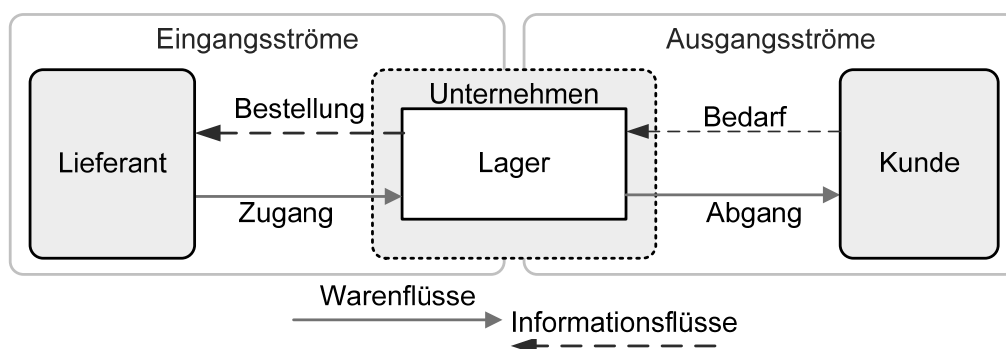


Abbildung 2.3: Konstituenten eines Lagerhaltungsproblems²

¹ Eine Darstellung und Diskussion der Funktions- und Prozessmodelle der anderen Geschäftsfälle findet sich in [Becker und Schütte 1996].

² Erweiterte Darstellung in Anlehnung an SCHNEEWEIß [1981, S. 7]. Zusätzlich dazu verlaufen die Geldflüsse im Lagergeschäftsprozess des Handels wie die Warenflüsse entgegengesetzt zu den Informationsflüssen.

Nach WYBARK und WILLIAMS [1976, S. 598] werden die Unsicherheiten in Lagersystemen nach dem Ursprung der Unsicherheit aus den Ausgangsströmen der Nachfrageseite (engl. *demand side*) und der Unsicherheit aus den Eingangsströmen der Versorgungsseite (engl. *supply side*) unterschieden. Ferner wird nach Unsicherheiten der Zeitpunkte und der Mengen unterschieden. Unsicherheiten im Zeitpunkt von Warenezugängen entstehen durch Abweichungen der Lieferzeiten in Form verzögerter oder verfrühter Anlieferungen durch Lieferanten bzw. vorgelagerte Lager- oder Produktionsstufen, während Unsicherheiten in der Menge der Zugänge durch verringerte Liefermengen, Schwund, Verderb oder Ausschuss entstehen.¹ Unsicherheit der Bedarfsmengen entsteht durch Schwankungen in der Menge der nachgefragten Waren, während Unsicherheit der Bedarfszeitpunkte durch frühere oder spätere Nachfrage der Konsumenten oder vorgelagerte Lager- und Verarbeitungsstufen entstehen [Minner 2000, S. 11].² Unsicherheit in den Lagerabgängen in Menge und Zeit entsteht hingegen, wenn der Kundenbedarf nicht zeitsynchron aus dem Lagerbestand erfüllt werden kann und somit Fehlmengen verursacht oder Nachlieferungen erfordert.

Sowohl Eingangs- als auch Ausgangsströme von Waren und Information können hinsichtlich ihrer Menge und des Zeitpunktes deterministischer oder stochastischer Natur sein. Die Lagerabgänge werden durch die Nachfrage außerhalb des Systems hervorgerufen und gelten somit als unbeeinflussbar [Hochstädter 1969, S. 2-6]. Die Eingangsströme zur Auffüllung des Lagers können jedoch durch die Entscheidungsvariablen des Zeitpunktes und der Menge einer Bestellung gesteuert werden, die durch die Warendisposition bestimmt werden. Die Warendisposition steuert somit die Höhe des Lagerbestands durch die Bestimmung der Bestellmengen und -zeitpunkte unter Koordination der Eingangs- und Ausgangsströme. Während SCHNEEWEIß [1981, S. 6-7] die Steuerung der Lagerbestandshöhe als Lagerdisposition bezeichnet und von der Bestelldisposition zur Steuerung der Eingangsströme abgrenzt, wird unter Berücksichtigung der direkten Interdependenz zwischen Ausgangs- und Eingangsströmen durch den Fortfall der Produktion in Handelsbetrieben nachfolgend die Steuerung des Gesamtsystems unter dem Begriff der Warendisposition subsumiert.

Die Ausprägungen deterministischer bzw. stochastischer Eingangs- und Ausgangsströme erfordern ihre Berücksichtigung in darauf abgestimmten Bestellmengenmodellen. Ein Bestellmengenmodell „is given by the coordinated set of rules and procedures that allow for routine decisions on when and how much to order of each item that is needed ... to fill customer demand“ [Hax und Candea 1984, S. 129].³ Die Determinanten der Ausgangsströme des Bedarfs sowie der Eingangsströme der Bestellungen werden in den Abschnitten 2.2.1 sowie 2.3.1 charakterisiert; anschließend werden geeignete Modelle der Warendisposition abgeleitet. Die Güte der Warendisposition wird dabei anhand unterschiedlicher Beurteilungskriterien gemessen, welche die Höhe der erforderlichen Warenbestände und den Anteil der befriedigten Kundenbedarfe evaluieren; sie werden in Abschnitt 2.3.3.2 dargestellt.

Die Lagerbildung zur mengenmäßigen und zeitlichen Entkoppelung von Eingangs- und Ausgangsströmen erfolgt für unterschiedliche Zielsetzungen. SCHNEEWEIß [1981, S. 3] unterscheidet eingangs- und ausgangsbezogene Ziele, welche er in Anlehnung an ARROW et al. [1958] als Lagerhaltungsmotive bezeichnet.⁴ Motive der Eingangsströme der Versorgungs-

¹ Darüber hinaus treten zeitliche und mengenmäßige Effekte vielfach auch kombiniert auf, zum Beispiel durch verzögerte Teillieferungen von Bestellungen [Minner 2000, S. 11].

² Vielfach treten beide Effekte gleichzeitig auf, so dass die Einzeleffekte in den nachfolgend dargestellten Bestellrhythmusverfahren in der Wahrscheinlichkeitsverteilung des Bedarfs kombiniert werden. Dagegen erlauben Bestellpunkt-Verfahren eine eigenständige Modellierung beider Effekte in separaten Verteilungen.

³ Bestellmengenmodelle werden auch als Lagerhaltungsmodelle (engl. *inventory systems*), als Bestellregeln oder als Bestellpolitiken [Toporowski 1996] bezeichnet (vgl. Abschnitt 2.3.1.3).

⁴ HAX und CANDEA [1984, S. 125-127] sowie SILVER et al. [1985, S. 30-32] kategorisieren die Motive ebenfalls nach ARROW [Arrow et al. 1958] in transaktions-, sicherheits- und spekulationsorientierte Motive, jedoch ohne Zuordnung zu den Input- und Output-Strömen. MINNER [2000, S. 8-10] leitet aus den Motiven

seite umfassen die einkaufsbedingte Lagerung zur Ausnutzung von Mengenrabatten, die transportbedingte Lagerung zur Ausnutzung von Transportkapazitäten in der Bestellung und spekulationsorientierte Motive der Lagerung. Ferner bestehen Motive der produktionsbedingten Lagerung für Waren mit Reifeprozessen sowie das verbreitete Motiv der beschaffungsbedingten Lagerung zur Bildung von Sicherheitsvorräten für Unsicherheiten auf den Beschaffungsmärkten. Motive der Nachfrageseite umfassen dagegen die absatzbedingte Lagerung, die eine Lagerbildung aufgrund der Struktur der Ausgangsströme aus einer Vielzahl von Bedarfen mit abweichenden Mengen und Zeitpunkten erfordert [Schneeweiß 1981, S. 3-4]. Je nach der Position des Lagers im Distributionssystem können sowohl versorgungs- als auch nachfrageseitige Motive überwiegen und unterschiedliche Entscheidungsmodelle der Disposition bestimmen. In den Handelsfilialen mit allgemein stochastischen Bedarfsstrukturen dominieren nachfrageseitige Motive einer unmittelbaren Bedarfsbefriedigung die versorgungsseitigen Motive. Dagegen treten in den Zentrallagern des Handels auch spekulationsorientierte Motive der einkaufs- und transportbedingten Lagerung auf der Versorgungsseite auf, so dass die Distributionssysteme im Handel unterschiedliche Lagerhaltungsmotive an ihren Schnittstellen koordinieren.

Die Dispositionsentscheidungen der Warenbestände innerhalb einer Handelsunternehmung können somit nicht losgelöst von der Filial- und Lagerstruktur im Distributionssystem betrachtet werden [siehe u.a. Toporowski 1996, S. 47; Wahl 1999]. Die Struktur der Warenströme wird insbesondere durch die Anzahl der unterschiedlichen Lagerstufen (vertikale Struktur), die Anzahl und die Distanz der Lagerstandorte auf jeder Stufe (horizontale Struktur) und die Beziehungen zwischen den Lagern¹ sowie durch ihre Lagerhaltungsmotive im logistischen Gesamtsystem bestimmt [Schneeweiß 1981, S. 4; Toporowski 1996, S. 47; Schulte 1991].² Im Handel werden die Logistikeinheiten der Zentrallager (ZL) und der untergeordneten Regionallager (RL) als Auslieferungslager bezeichnet und von den Logistikeinheiten der Filiallager (FL) und der Lagerung im Verkaufsraum (VR) auf der Ebene der Filiallager unterschieden. Im Handel entsprechen die Kunden der letzten Lagerstufe den Konsumenten, weshalb VR und FL auch als Abverkaufsstelle (engl. *point of sale*, POS) bezeichnet werden.³ In Abbildung 2.4 sind die grundlegenden Lagerstrukturen, Warenflüsse und korrespondierenden Lagersysteme einer Unternehmung dargestellt. Die Darstellung der Warenflüsse zwischen Lieferant und Kunde über unterschiedliche Lagerstufen erweitert entsprechend die Darstellung der Konstituenten des einstufigen Lagerhaltungsmodells aus Abbildung 2.3.⁴

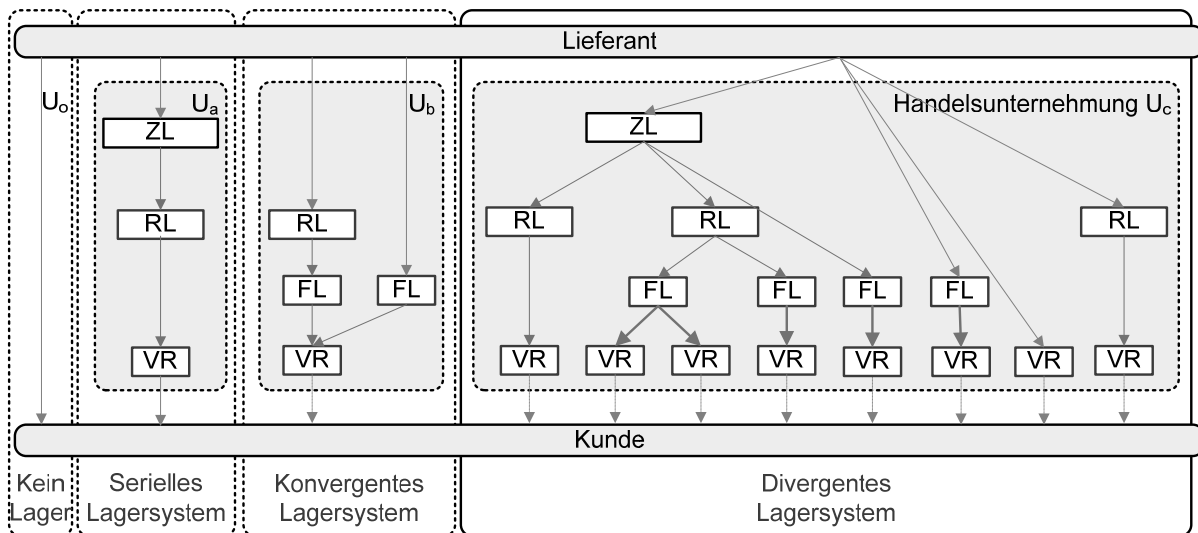
zudem Kategorien von Bestandsarten aus Produktions- und Distributionsentscheidungen ab. Diese Zuordnung erscheint jedoch wenig hilfreich, da Bestände in unterschiedlichen Betriebsformen durch unterschiedliche Motive gebildet werden können. Auf eine weiterführende Diskussion dieses Aspekts wird daher verzichtet.

¹ Eine detaillierte Diskussion alternativer Lagerstrukturen im Handel und ihrer Motive aus unterschiedlichen Schnittstellen findet sich in [Toporowski 1996; Wahl 1999] sowie handelsübergreifend in [Minner 2000].

² Die Festlegung der horizontalen und der vertikalen Lagerstruktur zählt zu den längerfristigen Entscheidungen des taktisch-strategischen Entscheidungsrahmens [Toporowski 1996]. Sie wird daher für den operativen Planungshorizont der Warendisposition der vorliegenden Untersuchung als konstant gegeben angenommen und hier lediglich zur Einordnung der Dispositionsprobleme in mehrstufige Distributionssysteme dargestellt.

³ Aus Sicht der Konsumenten werden sie auch als Einkaufsstätten (engl. *points of purchase*) bezeichnet.

⁴ Abbildung 2.4 wurde im Vergleich zur horizontalen Darstellung der Warenflüsse in Abbildung 2.3 zur besseren Übersichtlichkeit gedreht. Sie ist auf eine Darstellung der Strukturen von Distributionslagern im Handel beschränkt. Eine Einordnung industrieller Einkaufs-, Zwischen-, Ersatzteil- und Fertigproduktlager im Produktionsprozess gibt SCHNEEWEIß [1981, S. 4-5].

Abbildung 2.4: Grundformen von Lagerstrukturen einer Handelsunternehmung¹

Nach PFOHL [2004] lassen sich azyklische Distributionsstrukturen nach direkten, indirekten und kombinierten Warenflüssen unterscheiden, die zu unterschiedlichen Lagerstrukturen führen. Im Fall direkter Warenflüsse vom Hersteller bzw. Lieferanten zum Kunden (U_0 in Abbildung 2.4) liegt ein so genanntes nullstufiges System ohne Lagerstufen vor [Wahl 1999], bei dem der Distributionsvorgang ohne zusätzliche Lagerungs- oder Handhabungsprozesse durch ein Handelsunternehmen erfolgt.² Indirekte Güterflüsse vom Lieferanten zum Kunden beinhalten mindestens einen zusätzlichen Lagerstandort des Handels zur Konzentration oder Auflösung der Warenmengen [Wahl 1999, S. 11]. Lagerstrukturen können anhand der Anzahl von Lagerstufen und der Anzahl der Lagerstandorte je Stufe in ein- und mehrstufige Systeme unterteilt werden. „Bei einstufigen Systemen werden die Kunden von einer Stufe mit Lagerstandort bedient“ [Wahl 1999, S. 61], während in mehrstufigen Systemen mehrere Lagerstufen hintereinander geschaltet agieren (U_a , U_b und U_c in Abbildung 2.4). Kombinierte Systeme vereinigen direkte und indirekte Warenflüsse aus null-, ein- und mehrstufigen Teilsystemen für unterschiedliche Kernprozesse oder Sortimente im Distributionssystem [Pfohl 2004]. Diese Taxonomie kann auf zyklische Warenflüsse erweitert werden, zum Beispiel durch Warenrücknahmen und Recycling (engl. *reverse logistics*) [Tang und Teunter 2006], die spezifische Probleme in der Modellierung aufwerfen.

Mehrstufige Lagersysteme bestehen aus einer hierarchischen Folge von Lagerstandorten mit einem Verbund von Materialbeziehungen [Wahl 1999, S. 39], die nach CLARK und SCARF [1960] als Echelon bezeichnet werden. Sie treten in unterschiedlichen Topologien auf, deren Basistopologien serieller, konvergenter und divergenter Lagersysteme in Abbildung 2.4 dargestellt sind. Serielle Lagersysteme (U_a) besitzen höchstens einen Vorgänger- und einen Nachfolger-Knoten. Sie erlauben einfache Analysen, treten in der Praxis jedoch nur selten auf [Wahl 1999]. In konvergenten Lagerhaltungssystemen oder -subsystemen (U_b), die insbesondere in Montagesystemen, durch Konzentrationspunkte in Distributionsstrukturen oder durch mehrere Lieferanten entstehen, besitzen die einzelnen Lagerknoten mehrere Vorgänger und maximal einen Nachfolger [Wahl 1999]. In divergenten Systemen (U_c) besitzen die Lagerknoten maximal einen Vorgänger und beliebig viele Nachfolger, was zu den typisch

¹ Erweiterte Darstellung in Anlehnung an TOPOROWSKI [1996, S. 51].

² Abbildung 2.4 vernachlässigt die Lagerstufen der Einkaufs-, Zwischen-, Ersatzteil-, Fertigprodukte-, Auslieferungs- und Handelslager in den Produktions- und Distributionssystemen der zuliefernden Produktionsbetriebe zugunsten einer einheitlichen Betrachtung hierarchischer Lagerstrukturen in Distributionssystemen des Handels [Grochla 1992]. Ferner werden hier lediglich Vorgänge immobiler, standortbezogener Lagerhaltung berücksichtigt, unter Vernachlässigung der Aspekte mobiler und transportmittelbezogener Lagerhaltung im Distributionssystem [Becker und Schütte 1996].

arboreszenten Strukturen von Distributionssystem führt, die eine zulässige Annäherung an die empirischen Lagerhaltungsstrukturen im Handel repräsentieren. Empirische Lagerhaltungssysteme bestehen jedoch in der Regel als Mischformen der Basistopologien unter Erweiterung mehrstufiger Lagerhaltungssysteme um laterale, horizontale und zyklische Beziehungen durch gegenläufige Warenflüsse zwischen Lagerknoten einer Stufe [Wahl 1999, S. 59].¹

Eine nachfrageorientierte Koordination aller Warenflüsse eines Distributionssystems erfordert zunächst die genaue Disposition an der Schnittstelle der Einzelhandelsfilialen zur originären Kundennachfrage. Dies hebt die besondere Bedeutung der Filialdisposition in Abgrenzung zur Disposition in Auslieferungslagern hervor. Die nachfolgende Untersuchung der Bedarfsprognose und Warendisposition mit Neuronalen Netzen beschränkt sich daher auf die Problemstellungen und Entscheidungsparameter der einstufigen Filialdisposition zur Bestandsplanung unter Annahme der Konstanz aller Entscheidungsparameter des Lager- und Distributionssystems und unter Vernachlässigung spezifischer Problemstellungen der Disposition auf höheren Lagerstufen (in Übereinstimmung mit Becker und Schütte {, 1996, 193967936} sowie Toporowski {, 1996, 257781889}).²

Zur isolierten Warendisposition an einzelnen Lagerknoten des Distributionssystems sind die spezifischen warenlogistischen Eigenschaften der jeweiligen Handelslagerstufe zu berücksichtigen. ATZBERGER [1998a] hebt besonders die Unterschiede zwischen der Disposition auf Filiallager- und auf Auslieferungslagerebene hervor. Der Warenbestand der Auslieferungslager dient der Bündelung der Warenzugänge von Herstellern und ihrer Verteilung an Regionallager und Handelsfilialen unter Berücksichtigung spezifischer Lagerhaltungsmotive. Die Warenflüsse sind durch Eingangsströme großer Warenmengen von einer Vielzahl von Lieferanten und regelmäßige Ausgangsströme großer Warenmengen an wenige Handelsfilialen gekennzeichnet. Bedarfsschwankungen einzelner Filialen gleichen sich auf der übergeordneten Lagerstufe gegeneinander aus, so dass die Volatilität der Nachfrage und somit die Bedeutung der nachfrageseitigen Motive als auch die Komplexität der Warendisposition verringert wird.

Der Warenbestand in Handelsfilialen dient dagegen der direkten Befriedigung der Nachfrage von Konsumenten in der Zeit zwischen den Warenzugängen [Toporowski 1996, S. 106], deren regelmäßige und entsprechend rhythmischen Eingangsströme durch große Warenmengen aus einer einzelnen vorgelagerten Systemstufe und entsprechend geringfügiger durch versorgungsseitige Motive geprägt sind. Nach ATZBERGER [1998b, S. 7] unterscheiden sich Handelsfilialen vor allem durch den unmittelbaren Kontakt zu den Konsumenten und die resultierende Struktur hochfrequenter, geringer Warenabflüsse und somit die gesteigerte Volatilität der Zeitreihen. Fehlmengen besitzen eine sofortige Umsatzrelevanz durch verlore-

¹ Über die Betrachtung der innerbetrieblichen Lagersysteme hinaus erfordern auch die zwischenbetrieblichen Warenflüsse zwischen einem oder mehreren Lieferanten der Supply Chain eine Berücksichtigung in den Modellen der Bedarfs- und Bestellmengenrechnung. Aufgrund der hohen Komplexität der Koordination von zwischenbetrieblichen Dispositionsprozessen über Distributionsnetzwerke und Lagerstufen hinweg wird auf ihre Modellierung in der Bestellmengenrechnung vielfach verzichtet, so auch hier. Ebenso werden Ansätze der zwischenbetrieblichen Bedarfsrechnung in Form des Vendor Managed Inventory (VMI) oder des Collaborative Planning Forecasting and Replenishment (CPFR) durch den Austausch von Bedarfsinformationen über Dispositions- und Lagerstufen hinweg im Folgenden nicht untersucht. In diesem Zusammenhang wird auf die Publikationen der Initiative Efficient Consumer Response (ECR) verwiesen, der beide Verfahrensansätze entspringen [ECR Europe und Roland Berger 2003; Seifert 2001; Seifert 2004].

² Untersuchungen zur optimalen Lager- und Distributionsstruktur im Handel finden sich unter anderem in [Toporowski 1996; Dittrich 2000]. Weiterführende Aspekte, zum Beispiel der personalwirtschaftlichen oder aufbau- und ablauforganisatorischen Unterschiede der Disposition in Handelsfilialen und Zentrallagern, werden ebenfalls vernachlässigt. Modelle der Disposition und Bestandsoptimierung in zwei- und mehrstufigen Systemen des Handels, häufig allerdings unter Vernachlässigung der Probleme der Bedarfsrechnung, finden sich in [Toporowski 1996; Wahl 1999]. WAHL unterscheidet insbesondere zwischen dezentralen Pull-Systemen und zentralen Push-Systemen in der Koordination der Warendisposition über mehrere Dispositionsstufen [Wahl 1999; ebenso Minner 2000].

ne Kaufakte. Da Bedarfs- und Absatzmuster der Konsumenten im Wettbewerb lokal unterschiedlich sind, ist eine Disposition pro Filiale und Artikel erforderlich. Konventionelle Problemstellungen der Disposition in Zentrallagern zur Lieferantenselektion, Auftragsbündelung und Mengenkonditionen zur Berechnung optimaler Bestellmengen spielen in Handelsfilialen dagegen eine untergeordnete Rolle, da sie von Lagerstufen desselben Betriebs beliefert werden [Atzberger 1998b]. Durch die Belieferung aller Artikel aus einer oder wenigen übergeordneten Lagerstufen erfolgt üblicherweise eine rhythmische Planung, Bestellung und Belieferung. Die Disposition an den Abverkaufsstellen der Handelsfilialen unterscheidet sich auch in Menge und Frequenz und somit dem erforderlichen Ressourcenaufwand von der Zentrallagerlogistik. Je nach Betriebsform des Handels disponiert ein Discounter ca. 780 bis 1.600 Artikel, ein Supermarkt 7.000 bis 12.000, Verbrauchermärkte 21.000 bis 40.000 und Selbstbedienungs-Warenhäuser zwischen 33.000 und bis zu 63.000 Artikel [Atzberger 1998a]. Die Sortimentsbreite und -tiefe in den Filialen ist vielfach nur geringfügig kleiner als im Zentrallager. Aufgrund der geringen Lagerfläche in der Handelsfiliale muss das Sortiment häufiger, oft wöchentlich oder täglich, disponiert werden, während in der Zentrallagerdisposition längere Dispositionsintervalle je Lieferant und für unterschiedliche Warengruppen realisiert werden können. Die gesamte Dispositionsleistung einer Lagerstufe aus mehreren hundert Filialen eines Unternehmens übersteigt somit die Komplexität und Dispositionsleistung eines Zentrallagers bei Weitem. Entsprechend grenzt ATZBERGER [1998b] die Disposition in Handelsfilialen auch begrifflich ab und unterscheidet die Filialdisposition von der konventionellen Zentrallagerdisposition.

2.1.4 Funktionsmodell der Warendisposition

Die Betrachtung der Kernprozesse in Abschnitt 2.1.2 wird in Richtung auf eine prozessorientierte Analyse der Entscheidungen der Warendisposition in ihren Teilprozessen und Funktionen erweitert. Im Rahmen eines Architekturkonzeptes von Handelsinformationssystemen entwickeln BECKER und SCHÜTTE [1996, S. 173] ein erstes Referenzmodell der Warendisposition in Zentrallagern für den Geschäftsprozess des Lagergeschäftes von Handelsbetrieben.¹ Sie beschreiben die Warendisposition als zeitlich aufeinander folgende Hauptfunktionen der Bedarfsrechnung, der Bestellmengenrechnung, der Liefermengenrechnung, der Aufteilung, der Bestellübermittlung und der Bestellüberwachung, die in Abbildung 2.5 als Prozess dargestellt sind.² Der Hauptprozess der Warendisposition wird durch das Informationsobjekt der Bestellung geprägt, das den Lagerbestand zeitlich und mengenmäßig determiniert [Becker und Schütte 1996, S. 184].

¹ BECKER und SCHÜTTE [1996] spezifizieren ihr Funktionsmodell für eine einstufige Warendisposition in Handelszentralen bzw. Zentrallagern des Handels, wie aus den Funktionen zur Aufteilung der Bestellmengen für untergeordnete Lagerstufen der Regionallager, Warenverteilzentren und Filialen ersichtlich wird. Es wird hier unter Vernachlässigung einzelner Prozessschritte auf einen Prozess zur Filialdisposition ausgedehnt.

² Die Limitrechnung zur Beschaffungssteuerung durch die Festlegung von Beschaffungshöchstwerten für Warengruppen und Lieferanten wird von BECKER und SCHÜTTE [1996] sowie von OELDORF und OLFERT [2002] ebenfalls als Funktion der Warendisposition betrachtet. Sie wird hier in Anlehnung an [Götz 1999, S. 51] der vorherigen Kernfunktion des Einkaufs zugeordnet, da sie nur geringe prozessorientierte Verknüpfungen mit den Funktionen der Warendisposition aufweist und zudem nur geringe Relevanz für die nachfolgenden Untersuchungen der Filialdisposition besitzt.

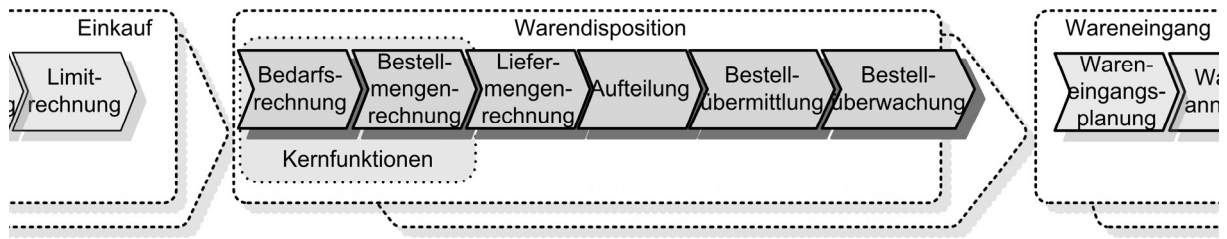


Abbildung 2.5: Teilprozesse der Warendisposition im Handel¹

Voraussetzung für die Deckung zukünftiger Bedarfe durch die Warendisposition ist zunächst eine Bedarfsfeststellung durch die Kernfunktion der Bedarfsrechnung [Lensing und Sonnemann 1995]. Die Bedarfsrechnung verfolgt das Ziel einer möglichst genauen Ermittlung der deterministischen oder stochastischen Bedarfsmengen und Bedarfszeitpunkte und betrachtet ausschließlich die Ausgangsströme der Nachfrageseite. Dabei kommen in Abhängigkeit von der Struktur des Bedarfs sowohl Modelle der programmorientierten Bedarfsrechnung als auch der verbrauchsorientierten Bedarfsprognose zum Einsatz, die in Abschnitt 2.2 näher betrachtet werden. Ausgehend von den ermittelten Bedarfsdaten und -zeitpunkten errechnet die Bestellmengenrechnung anschließend kostenminimale Bestellmengen und Bestellzeitpunkte zur effizienten Bedarfsdeckung unter Koordination aller Determinanten der Eingangs- und Ausgangsströme. Die verwendeten Modelle der Bestellmengenrechnung werden insbesondere durch die Ausgestaltung der Waren- und Informationsströme sowie der Lagerstrukturen im Unternehmen bestimmt, die in Abschnitt 2.3 dargestellt werden.

Die Liefermengenrechnung betrachtet die Auswahl eines Lieferanten anhand quantitativer und qualitativer Kriterien, z. B. Preis und Lieferservice, sowie die Festlegung der Lieferzeiten und Liefermengen jedes Lieferanten aus den vorangehend erhobenen Bestellmengen und -zeitpunkten der Bestellmengenrechnung [Becker und Schütte 1996, S. 183].² Die Aufteilung umfasst alle Funktionen der effizienten Verteilung einer zentralen Gesamtbeschaffungsmenge der zentralen Warendisposition durch eine Handelszentrale auf untergeordnete Lagerstufen oder Filialen im Logistiksystem der Unternehmung [Becker und Schütte 1996, S. 184]. Anschließend erfolgt die Bestellübermittlung zu den Lieferanten unter Aggregation der Filialbedarfe und -bestellungen durch die Zentrale [Becker und Schütte 1996, S. 184]. Nach Übersendung der zeitlich und mengenmäßig definierten Bestellung wird der Dispositionsprozess durch die Bestellüberwachung abgeschlossen, der die Einhaltung der Liefertreue der beauftragten Lieferanten hinsichtlich Menge und Zeit beurteilt. Der sich anschließende Kernprozess des Wareneingangs wird – zeitlich entkoppelt von der Warendisposition – durch die Warenannahme der Bestellmengen angestoßen.

Die vorliegende Untersuchung beschränkt die Analyse der Prognosegenauigkeit und Dispositionsgüte von Neuronalen Netzen auf die Filialdisposition des Handels. Dies erlaubt nachfolgend eine Vernachlässigung der Teilprozesse der Liefermengenrechnung und Aufteilung (siehe Referenzmodell in Abbildung 2.5), da diese ausschließlich in der Disposition mit zentraler Bestellung und Belieferung in Handelszentralen und Zentrallagern entstehen. Ebenso werden die Funktionen Bestellübermittlung und Bestellüberwachung vernachlässigt,

¹ Eigene Darstellung in Anlehnung an BECKER und SCHÜTTE [1996, S. 173 ff.]. Die Darstellung erfolgt als Wertschöpfungskettendiagramm in Anlehnung an die Modellierungsrichtlinien der Architektur Integrierter Informationssysteme (ARIS) nach SCHEER [2001; 2002].

² Im Gegensatz zur taktisch-strategischen Lieferantenvorauswahl des Einkaufs ist die Lieferantenauswahl eine operative Aufgabe der Beschaffungslogistik. BECKER und SCHÜTTE [1996, S. 176] dekomponieren so das Problem der simultanen Planung der Bestellmengen und -zeitpunkte für mehrere Lieferanten durch eine sequenzielle Liefermengenrechnung gemäß der Handelspraxis. Diese Vereinfachung erscheint auch vor dem Hintergrund der Belieferung von Handelsfilialen aus einem unternehmenseigenen Zentrallager zulässig, so dass nachfolgend die Problemstellungen einer simultanen Planung der Bestellmengen und -zeitpunkte für unterschiedliche Lieferanten in der Belieferung im Zuge der Filialdisposition vernachlässigt werden können.

da diese in Handelsinformationssystemen standardisiert und automatisiert sind und keine Interaktion mit unterschiedlichen Modellen der Warendisposition bieten.

Bedarfsrechnung und Bestellmengenrechnung bilden somit die Kernfunktionen der Disposition, unabhängig von der disponierenden Lagerstufe und der Betriebsform. Die prozessorientierte Betrachtung der Warendisposition in ihren Hauptfunktionen verdeutlicht die Dekomposition der übergeordneten Problemstellung in disjunkte Teilfunktionen. Das übergeordnete Ziel aller Teilprozesse der Warendisposition ist die wirtschaftliche, artgerechte, mengengerechte und zeitgerechte Deckung des Warenbedarfs eines Handelsunternehmens zur Bedarfsbefriedigung seiner Kunden [Oeldorf und Olfert 2002, S. 155]. Bedarfs- und Bestellmengenrechnung werden dabei als disjunkte Teilfunktionen aufgefasst, die eigenständig untersucht werden können. Die Aufteilung in die sequenziellen Teilprobleme einer exakten Bedarfsvorhersage durch die Bedarfsrechnung und einer Bestimmung optimaler Bestellmengen und -zeitpunkte durch die Bestellmengenrechnung erlaubt eine Reduktion der Modellierungskomplexität und eine entsprechende Vereinfachung der Betrachtung unter sequenzieller Lösung der Teilkomponenten [Grochla 1992]. Die Lösung der dekomponierten Problemstellungen erfolgt in eigenständigen Modellen mit individuell abgeleiteten Unterzielen unter Berücksichtigung ihrer Interdependenzen. Der grundlegende Zielkonflikt der Warendisposition aus den antinomen Zielen der Kostenminimierung durch Verringerung der Warenbestände und der Lieferservicemaximierung durch deren Erhöhung kann so im Rahmen der Bestellmengenrechnung weitgehend getrennt von der Zielsetzung einer möglichst genauen und unverzerrten Bedarfsvorhersage gelöst werden.¹

Nachfolgend werden die beiden relevanten Kernfunktionen der Bedarfsrechnung und der Bestellmengenrechnung hinsichtlich ihrer Strukturdimensionen, Determinanten, Modelle und Verfahren strukturiert dargestellt, um die empirische Problemstellung der Warendisposition in Handelsfilialen genauer analysieren zu können.

¹ Dabei sind allerdings die statistischen Voraussetzungen einer unverzerrten Schätzung ohne systematische Über- oder Unterprognose zu überprüfen, siehe auch Abschnitt 3.

2.2 Modelle der Bedarfsrechnung

2.2.1 Determinanten

2.2.1.1 Kennzeichnung des Bedarfs

Die Bedarfsrechnung dient einer möglichst validen und reliablen Quantifizierung des zukünftig auftretenden Bedarfs in unterschiedlichen Perioden.¹ Der Bedarf ist die am Markt tatsächlich auftretende Nachfrage nach Gütern [Hadelers 2000, S. 132]; er kann nicht direkt durch die Unternehmung manipuliert werden [Assfalg 1976].²

Da Bedarf und Lagerbestand in Menge und Zeitpunkt nicht übereinstimmen müssen, unterscheidet sich der Bedarf durch Fehlmengen vom Absatz. Der Bedarf entspricht dabei der Menge eines Gutes, die einer Lagerstufe zur Befriedigung der Kundennachfrage entnommen werden müsste [Naddor 1971, S. 134], während der Absatz die befriedigte Nachfrage einer Lagerstufe bezeichnet. Da der Absatz durch die Mengen des physischen Lagerabgangs leichter in konventionellen Geschäftsprozessen zu erheben ist, werden Bedarf und Absatz vielfach synonym betrachtet und Untersuchungen auf den Absatz beschränkt [Assfalg 1976; Klingst 1979]. Im Selbstbedienungshandel führt die räumliche Distanz der Kasse zudem zu einem Verlust des Kundenkontaktes und somit zur Unkenntnis des wahren Bedarfs am Regalplatz. Aufgrund des Messproblems des wahren Bedarfs bildet der Absatz hier die einzige zugängliche Informationsquelle zur Bestimmung des Bedarfs [Götz 1999, S. 50]. Nachfolgend werden die Begriffe Bedarf und Absatz synonym verwendet, obwohl der Bedarf die eigentliche Einflussgröße auf die Warenströme und das Lagerhaltungssystem darstellt [Götz 1999, S. 49].³

Der Bedarf ist zunächst hinsichtlich seiner Determinanten zu charakterisieren, um daraus geeignete Modelle der Bedarfsrechnung und später der Bestellmengenrechnung abzuleiten. Bedarfssituationen werden nachfolgend nach den materialwirtschaftlichen Planungsaspekten, dem Grad der Unsicherheit der zukünftigen Bedarfssituation und der Bedarfsentwicklung im Zeitablauf über mehrere Planungsperioden hinweg unterschieden.

Nach den Planungsaspekten der Materiallogistik werden drei Bedarfsarten unterschieden: Primär-, Sekundär- und Tertiärbedarf. Der Primärbedarf stellt den absatzbestimmenden Bedarf an Handelswaren bzw. Enderzeugnissen, Fertigprodukten oder verkaufsfähigen Ersatzteilen dar [Hadelers 2000, S. 155; Schulte 1991; Tempelmeier 2003, S. 8]. Sekundär- und Tertiärbedarf bezeichnen dagegen den Bedarf an Rohstoffen, Einzelteilen und Baugruppen sowie an Hilfs- und Betriebsstoffen, die zur Erzeugung des Primärbedarfs benötigt werden und die aus der Erzeugnisstruktur des Primärbedarfs mittels Bedarfsauflösung durch Stücklisten bestimmt werden [Hadelers 2000, S. 132]. In Handelsbetrieben entspricht der Sekundärbedarf den einzelnen Komponenten aus kombinierten Aktionswaren, speziellen Werbemittel- und Warenträgern, die direkt über Stücklisten aus den Primärbedarfen der disponierten Aktionswaren ermittelt werden können. Tertiärbedarfe bestehen aus den Verbrauchsmaterialien, u. a. dem Verpackungsmaterial und den Transportträgern, die in den logistischen Prozessen des Handels aufgehen. Aufgrund der geringen Bedeutung von Sekundär- und Tertiärbedarf im Lagerprozess des Handels kann die materialwirtschaftliche

¹ Die Bedarfsrechnung wird auch als Bedarfsermittlung, Bedarfsmengenplanung, Beschaffungsdisposition oder Materialbedarfsermittlung bezeichnet [Oeldorf und Olfert 2002, S. 154-204; Hadelers 2000].

² In der Literatur zu Lagerhaltungssystemen wird der Bedarf auch als *lead time demand* (LTD) bezeichnet.

³ Die systematischen Verzerrungen durch die Verwendung von Absatzmengen anstelle der tatsächlichen Bedarfsmengen werden somit – in Übereinstimmung mit der gängigen Literatur – vernachlässigt.

Unterscheidung für Handelswaren vernachlässigt werden [Becker und Schütte 1996, S. 174]. Nachfolgende Ausführungen beschränken sich daher auf die Primärbedarfe.¹

Ferner werden Bedarfsstrukturen nach dem Grad der Unsicherheit in der zukünftigen Bedarfssituation in deterministischen und stochastischen Bedarf unterteilt.² Verfügt der betriebliche Entscheidungsträger im Planungszeitraum über genaue Kenntnis der Höhe und der Zeitpunkte des zukünftigen Bedarfs, liegt eine deterministische Bedarfsstruktur im Sinne einer Entscheidung unter Sicherheit vor [Götz 1999, S. 50]. Liegen dagegen keine sicheren Informationen über Höhe und Zeitpunkt des Bedarfs vor, besteht eine stochastische Bedarfssituation unter Risiko oder Ungewissheit [Götz 1999, S. 50; Assfalg 1976, S. 16]. Im Fall des stochastischen Bedarfs unter Risiko wird der Bedarf als Zufallsvariable betrachtet, die einer bekannten und vollständig spezifizierten Wahrscheinlichkeitsverteilung folgt (vgl. Abschnitt 2.2.3). Dem Disponenten sind somit die Eintrittswahrscheinlichkeiten der möglichen Umweltzustände des Bedarfs bekannt, so dass der Bedarf zukünftiger Planungsperioden durch stochastische Modelle der Bedarfsrechnung prognostiziert werden kann. Ist die Verteilungsfunktion des Bedarfs hingegen unbekannt oder ihre Lage- und Verteilungsparameter nur teilweise bekannt, handelt es sich um einen stochastischen Bedarf bei Ungewissheit oder objektiver Unsicherheit [Assfalg 1976, S. 16; Starr und Miller 1962]. Prognosen über die zukünftige Höhe des Bedarfes können in diesem Fall nicht oder nur unter zusätzlichen Annahmen erfolgen [Götz 1999, S. 50; Hochstädter 1969].³

Darüber hinaus wird der Bedarf nach der Konstanz der Bedarfsentwicklung im Zeitablauf über mehrere Planungsperioden hinweg unterschieden. DOMSCHKE et al. [1997, S. 64] unterscheiden bei der Formulierung und Lösung von Modellen zur Lagerhaltung zwischen statischen und dynamischen Modellen, nach denen alle oder nur einige Parameterwerte im Planungszeitraum unverändert bleiben oder aber zeitlichen Schwankungen unterliegen. Für den Parameter des Bedarfs wird ein zeitinvarianter Bedarfsverlauf als stationärer Bedarf bezeichnet, ein Bedarfsverlauf mit zeitlichen Veränderungen dagegen als instationärer Bedarf.⁴ Bei stationärem Bedarf wird somit unabhängig von der Planungsperiode im Zeitablauf jeweils eine konstante Menge \bar{y} nachgefragt, während bei einer instationären Bedarfsstruktur die nachgefragten Mengen y_t je Planungsperiode t variieren [Stickel 1997, S. 199]. Die Unterscheidung der Bedarfsentwicklungen entspricht der Abgrenzung von stationären und instationären stochastischen Prozessen der statistischen Literatur zur Prognose [Müller-Hagedorn 1998, S. 542; Schlittgen 1997]. Sie unterscheidet stationäre stochastische Prozesse, in denen Mittelwert und Varianz der Bedarfsverteilung im Zeitablauf konstant bleiben, von instationärem Bedarf mit veränderlichen Parametern der Verteilungsfunktion im Zeitablauf im Sinne einer Strukturveränderung der Bedarfsverteilung. Aufgrund der besonde-

¹ Ferner fasst TRUX [1972, S. 18-19] für Waren, die sowohl Primär- als auch Sekundärbedarfe aufweisen, beide zum Periodenbedarf zusammen. Ferner unterscheidet er den Brutto-Bedarf einer Periode von dem Netto-Bedarf durch die Bedarfsmenge, die unmittelbar aus Lagerbeständen befriedigt werden kann [Trux 1972, S. 19-20]. Diese Unterscheidung von Brutto- und Netto-Bedarfen wird nachfolgend jedoch durch die Modelle der Bestellmengenrechnung und nicht der Bedarfsrechnung betrachtet.

² Der Grad der Unsicherheit wird in der betriebswirtschaftlichen Entscheidungstheorie nach der Kenntnis der Eintrittswahrscheinlichkeiten der Umweltzustände unterschieden. DOMSCHKE et al. [1997, S. 63] erweitern die Unterscheidung in deterministische und stochastische Modelle nach dem Grad der Unsicherheit des Bedarfs auf alle Parameter von Lagerhaltungsmodellen. Ihre Ausführungen betrachten jedoch ausschließlich stochastische Ausprägungen des Bedarfs, nicht der anderen Eingangs- und Ausgangsströme, so dass die Einschränkung der Abgrenzung nach dem Grad der Unsicherheit des Bedarfs zulässig erscheint.

³ Darüber hinaus unterscheidet die Entscheidungstheorie auch stochastische Entscheidungen bei Unwissen, in denen auch die Auswirkungen der untersuchten Handlungsalternativen nicht vollständig bekannt sind. Diese sind jedoch aufgrund der bekannten Interaktion des Bedarfs in der Entscheidungssituation der Bedarfs- und Bestellmengenrechnung für die Warendisposition zu vernachlässigen.

⁴ Diese Bezeichnung erfolgt nicht einheitlich. GÖTZ [1999] unterscheidet statischen von dynamischem Bedarf, in Übereinstimmung mit der Einteilung der Modellparameter durch DOMSCHKE et al. [1997]. Da die nachfolgenden Ausführungen ausschließlich die Dynamik des Bedarfs in den Modellen berücksichtigen, werden hier gemäß der Literatur zur Prognose stationäre und instationäre Modelle unterschieden.

ren Bedeutung für die Modelle der Bedarfsrechnung wird das Konzept der Stationarität in Abschnitt 2.2.3.1 formal dargestellt.

2.2.1.2 Modellauswahl

Die Kennzeichnung des Bedarfs erfolgt nach dem Grad der Unsicherheit und der zeitlichen Bedarfsentwicklung. Diese bestimmen maßgeblich die Gestaltung der zulässigen Modelle in der Bedarfsrechnung und der Bestellmengenrechnung, und werden in Abbildung 2.6 einander zugeordnet.

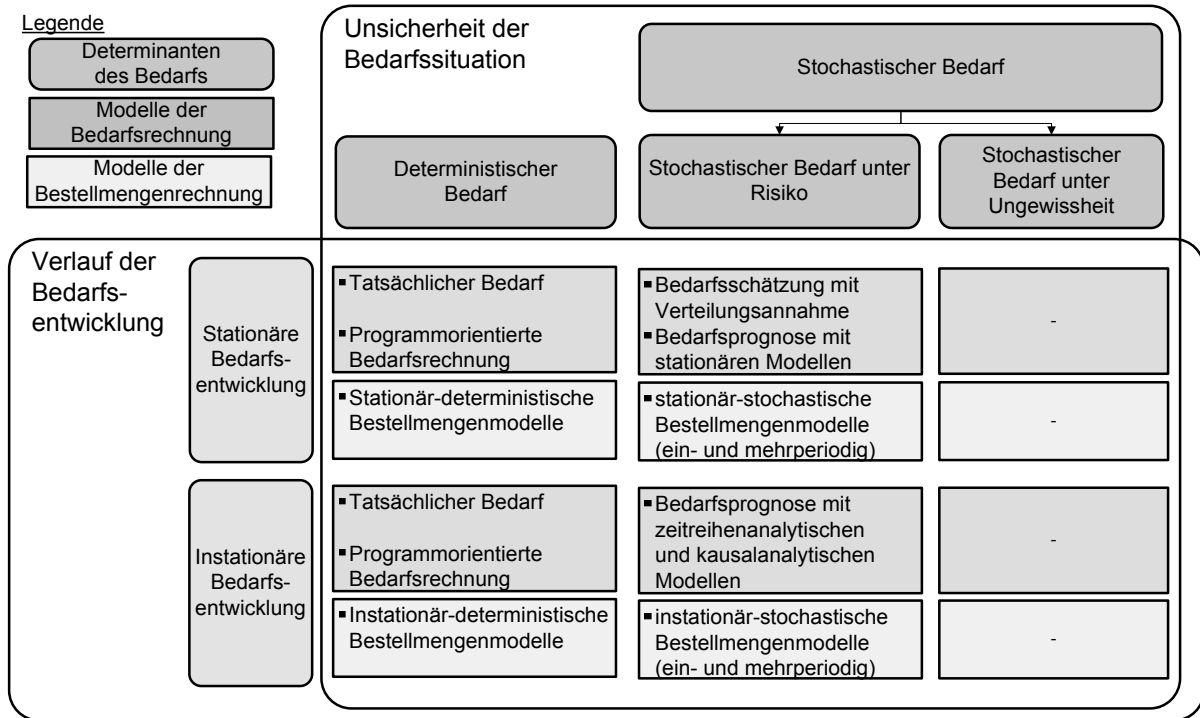


Abbildung 2.6: Modelle der Bedarfs- und Bestellmengenrechnung für die Determinanten des Bedarfs

Die Modelle der Bedarfsberechnung in Abbildung 2.6 werden dabei zunächst nach der Berücksichtigung zukünftiger, stochastischer Einflussgrößen in der Planung unterschieden.¹ Da in deterministischen Bedarfssituationen der zukünftige Bedarf bekannt ist, erfordern sie grundsätzlich keine Vorhersage, sondern erlauben die Verwendung des tatsächlichen Bedarfs

¹ SCHULTE [1991, S. 23] unterteilt darüber hinaus instationäre Bedarfsformen weiter nach der Bedarfskontinuität in kontinuierlichen und unregelmäßigen Bedarf sowie die Sonderformen des erstmaligen und des einmaligen Bedarfs. Als unregelmäßiger Bedarf wird dabei ein unvorhersehbares oder sporadisches, stochastisches Bedarfsauftreten charakterisiert, das spezifische Verfahren der Bedarfsrechnung zur Bedarfsermittlung erfordert. Auch wenn unregelmäßiger Bedarf nach TOPOROWSKI [1996, S. 108] die Anwendung konventioneller Verfahren im Rahmen der Bestellmengenrechnung verbietet, werden sie hier als Sonderformen des stochastischen, instationären Bedarfs betrachtet, die durch die Selektion zutreffender Modelle und Verfahren der Absatzprognose berücksichtigt werden (siehe Abschnitt 3.2). Die Bedarfsformen des erstmaligen oder einmaligen Bedarfs können für deterministische Bedarfe problemlos vorhergesagt werden, da diese bekannt sind. Die Problemstellung der Prognose stochastischer erst- und einmaliger Bedarfe umfasst die Prognose von neuen Produkten ohne vorherige Bedarfsinformationen, was empirisch beispielsweise durch subjektive Prognoseverfahren der Analogieschlüsse erfolgt. Diese Prognose einmaliger Bedarfe oder auch einmaliger Ereignisse im Sinne von Ausreißern bildet ein eigenständiges Forschungsfeld, auf das hier nicht eingegangen werden soll. GÖTZ [1999, S. 51] sowie NADDOR [1971] unterscheiden ferner nach der zeitlichen Verteilung der stochastischen Bedarfsentwicklung innerhalb einer Planungsperiode, was sie als Bedarfsform pro Periode bezeichnen. Die Probleme einer konzentrierten Bedarfsintensität zum Beginn oder Ende einer Planungsperiode im Gegensatz zu einer gleichmäßig verteilten Bedarfsentwicklung kann jedoch durch adäquate Bestimmung der Länge der Planungsperioden und somit der Zyklen der Bedarfsrechnung und Bestellmengenrechnung berücksichtigt werden. Das Problem der Selektion einer äquidistanten Planungsperiode in Abhängigkeit von der Variabilität der Bedarfsentwicklung stellt jedoch ebenfalls ein eigenständiges Planungsproblem dar, auf das hier nicht eingegangen werden kann.

oder die Bedarfsrechnung zur Aufteilung der Primärbedarfe auf Sekundär- und Tertiärbedarfe [Bichler und Lörsch 1985, S. 118; Tempelmeier 1992; Stickel 1997].¹ Die Bestimmung des deterministischen Bedarfs wird auch als programmorientierte Bedarfsrechnung bezeichnet und in Abschnitt 2.2.2 näher analysiert.

Modelle der stochastischen Bedarfsrechnung werden eingesetzt, wenn exakte deterministische Verfahren aufgrund unbekannter zukünftiger Bedarfe nicht anwendbar sind. Sie ermitteln den Bedarf zukünftiger Perioden aufgrund von empirisch beobachteten Verbrauchswerten der Vergangenheit unter Prognose in die Zukunft [Stickel 1997, S. 200; Becker und Schütte 1996, S. 174; Oeldorf und Olfert 2002, S. 186 ff.]. In der Auftragsfertigung mit deterministischen Bedarfsstrukturen werden Modelle der stochastischen Bedarfsprognose zudem zur effizienten Planung geringwertiger Sekundär- und Tertiärbedarfe eingesetzt. Dabei ist abzuwägen, ob die höhere Planungsgenauigkeit und der hohe Rechenaufwand einer deterministischen Bedarfsrechnung durch Stücklistenauflösung für stationäre oder instationäre Sekundär- und Tertiärbedarfe gerechtfertigt ist oder ob auf die Anwendung der effizienten, aber ungenaueren stochastischen Verfahren der Bedarfsprognose zurückgegriffen werden soll [Haderler 2000, S. 154; Tempelmeier 2003; Schulte 1991]. Liegt eine stochastische Bedarfsituation vor, so erfordert diese jedoch zwingend die Vorhersage unbekannter Bedarfsmengen und –zeitpunkte durch stochastische Modelle und Verfahren.

Modelle der stochastischen Bedarfsrechnung werden in der Materialwirtschaft auch als Modelle der verbrauchsorientierten Bedarfsrechnung oder Bedarfsprognose bezeichnet, in Abgrenzung zu den programmorientierten Modellen bei deterministischem Bedarf. Zur eindeutigen sprachlichen Abgrenzung werden nachfolgend die deterministischen Modelle der Bedarfsrechnung von den stochastischen Modellen der Bedarfsprognose unterschieden. Aufgrund ihrer besonderen Bedeutung für die Warendisposition im Handel werden stochastische Modelle nachfolgend nur kurz dargestellt und anschließend in Kapitel 3 formal hergeleitet und analysiert. Im Anschluss daran werden in Kapitel 4 und 5 neuartige Verfahren zur Bedarfsprognose und unmittelbaren Disposition durch Neuronale Netze entwickelt. Die Bedarfsrechnung für Prozesse eines stationären, stochastischen Bedarfs kann sowohl durch Modelle mit konstanter Verteilungsannahme als auch durch Prognosemodelle für stationären Verlauf erfolgen, die in Abschnitt 2.2.3.1 dargestellt werden. Dagegen erfordern instationäre stochastische Bedarfe die Identifikation der zugrunde liegenden Modellform einschließlich Trends, Saisonalitäten, Ausreißern und Niveauverschiebungen und die Auswahl und Anwendung eines passenden Prognosemodells zur Bedarfsrechnung. Diese werden zunächst in Abschnitt 2.2.3.2 skizziert und aufgrund ihrer besonderen Bedeutung in der Filialdisposition in Abschnitt 3 eingehend diskutiert.

In der Handelspraxis ist eine Unterscheidung der Bedarfsstruktur nach dieser Taxonomie vielfach unnötig, da der Bedarf im anonymen Markt grundsätzlich stochastischer Natur ist und fast ausschließlich einer instationären Bedarfsentwicklung folgt. Dies ist insbesondere auf die diskrete Nachfrage vieler Verbraucher, Unstetigkeiten im Einkaufsverhalten der Nachfrager, Saison- und Modeeinflüsse auf den Bedarf, Störungen in den Waren- und Informationsflüssen der Distributionskette sowie den Einsatz absatz- und marketingpolitischer Instrumente der Preis-, Produkt-, Sortiments-, Kommunikations- und Distributionspolitik durch das eigene Unternehmen oder die Konkurrenz zurückzuführen [Müller-Hagedorn 1998, S. 517]. Dennoch werden aus Gründen der Vereinfachung häufig Modelle und Verfahren für stationäre stochastische und deterministische Bedarfsformen in der Praxis angewendet,

¹ Darüber hinaus werden weitere Arten der deterministischen Bedarfsrechnung wie die auftragsgesteuerte Bedarfsermittlung durch Einzelbedarfs- und Sammeldisposition [Lensing und Sonnemann 1995] sowie die projektgesteuerte Bedarfsrechnung [Hochstädter 1969] charakterisiert, die aufgrund ihrer ausschließlich produktionswirtschaftlichen Relevanz hier vernachlässigt werden.

obwohl die stark vereinfachten Annahmen der Bedarfsverläufe so systematische Fehler durch die Fehlspezifikationen des Bedarfs verursachten.¹ Die nachfolgende Diskussion der Bedarfsrechnung und der darauf aufbauenden Bestellmengenrechnung berücksichtigt daher sowohl deterministische als auch stochastische Modelle, um die Erkenntnisse und Unzulänglichkeiten einzelner Modelle aufzuzeigen und sie in die Literatur zur Waren- und Materialdisposition einzuordnen. Die Systematik der Bedarfsarten sowie die Modelle der Bedarfsrechnung werden in dem morphologischen Kasten in Tabelle 2.1 einander gegenübergestellt.

Tabelle 2.1: Systematik der Bedarfsarten, Modelle und Verfahren der Bedarfsrechnung²

Grad der Sicherheit des Bedarfs	deterministischer Bedarf			stochastischer Bedarf				
Regelmäßigkeit des Bedarfs	regelmäßiger Bedarf							unregelmäßiger B.
Stationarität des Bedarfs	stationärer Bedarfsverlauf			instationärer Bedarfsverlauf				
Verfahren der Bedarfsrechnung	Programmorientierte Verfahren der deterministischen Bedarfsrechnung			Verbrauchsorientierte Verfahren der stochastischen Bedarfsprognose				
Modelle der Bedarfsrechnung	Analytische Modelle	Synthetische Modelle	Gozinto-Modelle	Stationäre Modelle	Trend-Modelle	Saisonale Modelle	Trend-Saison-M.	Sporadische Modelle

Darüber hinaus beeinflusst die Bedarfsstruktur maßgeblich die Auswahl geeigneter Modelle der Bestellmengenrechnung. Ein stationär-deterministischer Bedarf erlaubt die Anwendung einfacher deterministischer Modelle und Verfahren der Bestellmengenrechnung für konstante Bedarfs- und Bestellmengen, da keine Sicherheitsbestände zur Absicherung von Planungsungenauigkeiten des Bedarfs erforderlich sind. Dagegen erfordert eine instationär-deterministische Bedarfsentwicklung die Ermittlung variierender optimaler Bestellmengen durch spezifische Verfahren dynamischer Mehrperiodenmodelle unter Vernachlässigung der Probleme der Bedarfsrechnung, siehe Abschnitt 2.3.2. Stochastische Bedarfsituationen erfordern die Berücksichtigung der Planungsungenauigkeit des Bedarfs durch stochastische Modelle der Bestellmengenrechnung. Für die Berücksichtigung des Risikos wird bei stationären Bedarfsentwicklungen angenommen, dass die Zufallsvariable des Bedarfs einer stationären Verteilung folgt. Im Fall instationärer Bedarfsentwicklungen wird dagegen die Stationarität der Residuen stochastischer Prognoseverfahren unterstellt [Müller-Hagedorn 1998, S. 542]; somit werden die Überlegungen zur Berechnung von Bestellmengen durch geeignete Prognoseverfahren von stationäre auf instationäre Bedarfsverläufe übertragen, siehe Abschnitt 2.3.3 und 2.3.3.2.

2.2.2 Deterministische Modelle

Modelle der deterministischen Bedarfsrechnung erlauben eine „exakte Bestimmung des Materialbedarfs nach Menge und Termin“ [Schulte 1991, S. 155] für deterministischen Bedarf. Sie werden in der materialwirtschaftlichen Bedarfsplanung auch als programmorientierte Modelle und Dispositionsverfahren bezeichnet [Hadelar 2000, S. 73; Arnolds et al. 2001].

Der deterministische Bedarf y_t [in Mengeneinheiten (ME) pro Zeiteinheit (ZE)]¹ bezeichnet die nachgefragte Menge eines Gutes in der Planungsperiode t . Bei deterministischem Bedarf

¹ GÖTZ [1999] vernachlässigt - trotz seiner mehrfachen Hinweise auf die Bedeutung der stochastischen Bedarfsunsicherheit und der Prognoseverfahren - die Stochastik des Bedarfs in seiner Arbeit. Vielmehr nimmt er stochastische Bedarfe als deterministisch vorgegeben an und verwendet sie in deterministischen Modellen der Bestellmengenrechnung. Dies vereinfacht die tatsächliche Problemstellung stark und führt zu einer erheblichen Verfälschung der Lösungen [Hartmann 2002].

² Eigene Darstellung in Anlehnung an TOPOROWSKI [1996, S. 113]. Die Darstellung vernachlässigt die materialwirtschaftliche Unterscheidung von Primär-, Sekundär- und Tertiärbedarf, da diese Bedarfskategorien nicht überschneidungsfrei mit den anderen Kategorien und Verfahren abgebildet werden können und Hinsichtlich der Bestellmengenrechnung von untergeordneter Relevanz sind.

werden die Mengen und Termine als ein mit Sicherheit bekanntes und unveränderliches Datum angenommen, unter Vernachlässigung aller stochastischen Einflussgrößen auf die Planungsperioden der Zukunft [Hadelar 2000]. Entsprechend erfordern deterministische Modelle der Bedarfsrechnung keine Schätzung der Primärbedarfe. Bei der Berücksichtigung des deterministischen Bedarfs ist keine Unterscheidung zwischen stationären und instationären Bedarfen erforderlich, da auch zukünftige dynamische Veränderungen des Bedarfs im Zeitablauf deterministisch bekannt sind und keine Vorhersage erfordern [Lee und Nahmias 1993, S. 6]. Somit sind keine Modelle zur Bedarfsvorhersage sowohl stationärer als auch instationärer deterministischer Bedarfsverläufe des Primärbedarfs erforderlich.

Deterministische Modelle der Bedarfsrechnung beziehen sich vielmehr auf die Ermittlung zukünftiger Bedarfsmengen der Sekundärbedarfe aus den deterministisch bekannten, zukunftsbezogenen Plandaten der Primärbedarfe [Schulte 1991, S. 155-156; Oeldorf und Olfert 2002; Arnolds et al. 2001, S. 85]. Die Bedarfsrechnung erfolgt aus Primärbedarfen der bereits erfassten Kundenaufträge, zum Beispiel in der industriellen Kundenauftragsfertigung oder bei einem ausreichenden Vorlauf des Auftragseingangs. Im Handel tritt die programmorientierte Bedarfsrechnung insbesondere im Aktionsgeschäft auf, wenn Aktionsbestellmengen von Filialen oder Kunden mit ausreichendem zeitlichem Vorlauf vorliegen und so die Primär- und Sekundärbedarfe ex ante bekannt sind [Becker und Schütte 1996, S. 174]. In diesem Fall gehen die deterministischen Mengen der Sekundär- und Tertiärbedarfe als Datum in die weiteren Teilprozesse der Disposition ein. TRUX [1972, S. 21-23] weist insbesondere auf die Problemstellung der unterschiedlichen Vorlaufzeiten der Bedarfsarten hin, so dass Sekundär- und Tertiärbedarfe ggf. früher verfügbar sein müssen, um in die Erstellung der Primärbedarfe einzugehen. Die Dekomposition besitzt somit sowohl eine mengenmäßige als auch eine zeitliche Dimension. Modelle der Bedarfsrechnung werden vorwiegend in der exakten Berechnung von Sekundärbedarfen aus deterministischen Primärauftragsmengen mittels Stücklisten- oder Rezepturauflösung angewendet [Lensing und Sonnemann 1995, S. 68]. Als Verfahren der Bedarfsauflösung werden u. a. analytische Fertigungs- oder Dispositionsstufenverfahren, Renetting- oder Gozinto-Verfahren sowie synthetische Verfahren der linearen Optimierung und der Graphentheorie angewendet [Tempelmeier 2003, S. 119-135; Oeldorf und Olfert 2002, S. 173-185].²

Die vereinfachende Annahme deterministischer Bedarfe durch den notwendigen zeitlichen Vorlauf zwischen Bestellung und Bedarfsbefriedigung steht vielfach im Widerspruch zur zeitlichen Überbrückungsfunktion von Handelsbetrieben, insbesondere im hier untersuchten Kernprozess des Lagergeschäftes [Becker und Schütte 1996]. Somit besitzen deterministische Modelle und Verfahren der Bedarfsrechnung kaum Relevanz für die Warendisposition von Handelsbetrieben und werden daher nachfolgend vernachlässigt.³

¹ Die Notationen der Modelle und Verfahren der Bedarfsrechnung, der Bestellmengenrechnung und insbesondere der Verfahren der Neuronalen Netze ist in der Literatur uneinheitlich. Sie werden so weit wie möglich unter Angabe der Abweichung von der Ursprungsnotation harmonisiert. So wird anstelle der in der stochastischen Bedarfsrechnung mittels Verteilungsannahmen üblichen Bezeichnung des Bedarfs x_t , nachfolgend die in der Prognoseliteratur üblichere Bezeichnung y_t verwendet.

² Die Dekomposition der aggregierten Bedarfe auf Zentrallagerebene in die einzelnen Bedarfe der Lagerstufen auf Filialebene durch die Hauptfunktion der Aufteilung (vgl. Abschnitt 2.1.4) wird jedoch nicht als Problem der programmorientierten Bedarfsrechnung betrachtet, sondern im Rahmen der mehrstufigen Disposition durch die Bestellmengenrechnung der Bedarfe auf untergeordneten Ebenen behandelt.

³ Aufgrund der Dominanz stochastischer Verfahren zur Bedarfsrechnung in Handelsbetrieben wird der Begriff Bedarfsrechnung häufig synonym mit der stochastischen Bedarfsrechnung, der Prognoserechnung oder der Bedarfsprognose verwendet. Dies erscheint jedoch irreführend und ist hinsichtlich der Bedeutung der Verfahrensauswahl aus den spezifischen Bedarfsstrukturen zu vermeiden.

2.2.3 Stochastische Modelle

2.2.3.1 Modelle für stationären Bedarf

Modelle der stationären, stochastischen Bedarfsprognose werden zur Vorhersage von stochastischem Primär-, Sekundär und Tertiärbedarf verwendet. Als Unterscheidungsmerkmal der Anwendung unterschiedlicher stochastischer Modelle dient die Stationarität des Bedarfs, der nachfolgend formal dargestellt wird.

Der Bedarf wird als Zufallsvariable Y betrachtet, die beliebige Werte y_t im Zeitablauf t annehmen kann. Eine stationäre Zeitreihe bezeichnet einen mittelwertstationären und kovarianzstationären Prozess, sofern keine systematischen Veränderungen der ersten beiden Momente von Erwartungswert $E(Y)$ und Varianz σ_Y^2 auftreten [Schlittgen und Streitberg 1987, S. 3; 79-85].¹ Ein stochastischer Prozess gilt als mittelwertstationär, wenn sein empirischer Mittelwert \bar{y}_t im Zeitablauf $t = 1, 2, \dots, T$ konstant ist:

$$\bar{y}_t =: \mu \quad \text{für alle } t. \quad (2.1)$$

Ein stochastischer Prozess gilt als kovarianzstationär, wenn die Kovarianzfunktion $\gamma(t, t+h)$ des Prozesses nur von der Entfernung $t - (t+h)$ der Zeitpunkte t und $t+h$ abhängt,

$$\gamma(t, t+h) =: \gamma(t - (t+h)) \quad \text{für alle } t, h. \quad (2.2)$$

Ist der Prozess kovarianzstationär, ist seine empirische Varianz s_t^2 somit auch varianzstationär:

$$s_t^2 =: \sigma^2 \quad \text{für alle } t, \text{ da} \quad (2.3)$$

$$\sigma_t^2 = \gamma(t, t) = \gamma(0) = \gamma(t+h, t+h) = \sigma_{t+h}^2$$

Varianzstationäre Zeitreihen von Beobachtungen heißen homoskedastisch, instationäre Zeitreihen mit heterogener Varianz dagegen heteroskedastisch. Die Verteilung eines stationären Bedarfs ist somit zeitinvariant. Die Stationarität einer Zeitreihe wird durch statistische Tests ermittelt, insbesondere den AUGMENTED DICKEY-FULLER-Test (ADF) zur Prüfung auf Stationarität des Mittelwertes sowie den GOLDFELD-QUANDT-Test zur Prüfung der Homoskedastizität von Prognoseresiduen [Makridakis et al. 1998].

In Abbildung 2.7 sind drei mittelwert- und varianzstationäre, homoskedastische Zeitreihen eines normalverteilten Bedarfsverlaufs y_t mit zunehmender Streuung unter Berücksichtigung der zeitlichen Ordnung der Bedarfs-Zeitpunkte t als Zeitreihendiagramme sowie als empirische Wahrscheinlichkeitsfunktionen $P_Y(y)$ ohne Zeitbezug y_i dargestellt.

¹ Die Instationarität umfasst für stochastischen Bedarf neben einer Veränderung des Bedarfsniveaus auch die Möglichkeit zur Veränderung der gesamten Dichte- und Verteilungsfunktion der Bedarfsverteilung [Domschke et al. 1997, S. 159], obwohl dies auch in theoretischen Studien weitgehend vernachlässigt wird. Eine formale Herleitung und ausführliche Diskussion der Stationarität von Zeitreihen findet sich bei NERLOVE et al. [1995, S. 23-36]

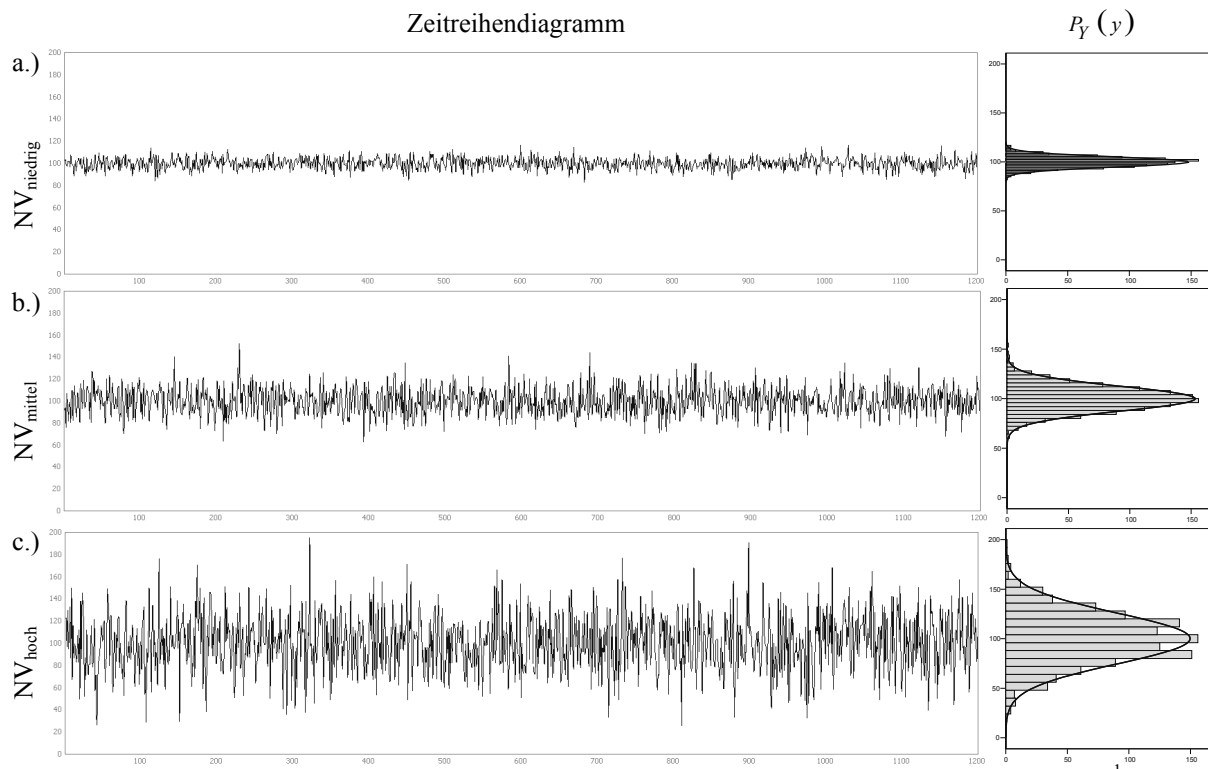


Abbildung 2.7: Stationäre Zeitreihen mit niedriger (a.), mittlerer (b.) und hoher Streuung (c.)¹

Abbildung 2.7 verdeutlicht die Auswirkungen einer ansteigenden Varianz der Zeitreihen von NV_{niedrig} über NV_{mittel} auf NV_{hoch} bei konstantem Mittelwert auf das Zeitreihendiagramm sowie auf die Wahrscheinlichkeitsverteilung des Bedarfs. Im Sonderfall stationärer Zeitreihen ist die Varianz der Zeitreihe um den stationären Mittelwert identisch mit der Varianz der Residuen der Prognose im Sinne der Prognosefehler um einen konstanten Mittelwert, nicht aber für Vorhersagen und Verfahren bei instationärem Bedarf.

Im Fall eines stationären, stochastischen Bedarfs bei Risiko kann der zukünftige Bedarf durch Modelle der Verteilungsannahme geschätzt werden. Der Bedarf wird als diskrete oder stetige Zufallsvariable Y betrachtet, die beliebige Werte y_i , einer bekannten, vollständig spezifizierten Wahrscheinlichkeitsverteilung annehmen kann.² Die Beobachtungen von y_i sind stochastisch unabhängig und ohne serielle Autokorrelation [Domschke et al. 1997, S. 159]; sie repräsentieren somit eine unabhängige Folge von Zufallszahlen ohne Berücksichtigung der zeitlichen Ordnung ihres Auftretens. Die Bedarfsschätzung durch Verteilungsannahme erfordert zunächst die Ermittlung der empirischen diskreten Wahrscheinlichkeitsfunktion bzw. der stetigen Dichtefunktion oder die Approximation durch eine theoretische Wahrscheinlichkeitsfunktion unter Schätzung der empirischen Verteilungsparameter. Der Erwartungswert $E(Y)$ des Bedarfs bezeichnet jenen Wert, der unter Berücksichtigung aller Ausprägungen der Zufallsvariablen y_i und der Wahrscheinlichkeit ihres Auftretens im Durchschnitt erwartet wird, und gilt als unverzerrter Schätzer für den Mittelwert der Grundgesamtheit μ .

¹ Die Zeitreihen besitzen eine normalverteilte Streuung von $Y \sim N(\mu, \sigma)$, mit den Parametern für die Zeitreihe der niedrigen Streuung NV_{niedrig} mit $N(100, 5)$ in (a.), der mittlerer Streuung NV_{mittel} mit $N(100, 12,5)$ in (b.) und der hoher Streuung NV_{hoch} mit $N(100, 25)$ in (c.). Alle Zeitreihen sind den nachfolgenden Experimenten aus Kapitel 5 entnommen und werden dort eingehender beschrieben.

² Abweichend zur Notation der Materiallogistik und der Bestellmengenrechnung erfolgt die Darstellung des Bedarfs durch Y bzw. y [siehe u.a. Toporowski 1999a; 1999b; Götz 1999] anstelle von X bzw. x [siehe u.a. Naddor 1971] oder D und d [Lee und Nahmias 1993; Hillier und Lieberman 1997]. Dies dient der Anpassung der Notation an die Prognosemodelle und Verfahren, in denen x unabhängige Variablen und y die abhängige Variable des Prognosegegenstandes repräsentiert (siehe Kapitel 3).

Eine Bedarfsschätzung durch Verteilungsannahme ist jedoch nur für Bedarfsverteilungen mit stationärem Mittelwert und Varianz zulässig, was in der Handelspraxis nur einen seltenen Sonderfall empirischer Bedarfsverteilungen repräsentiert. In Abbildung 2.8 sind Beispiele instationärer Bedarfsverläufe dargestellt: (a.) Bedarf mit stationärem Mittelwert und zunehmender Varianz, (b.) instationärem Mittelwert und konstanter Varianz und (c.) instationärem Mittelwert und heteroskedastischer Varianz. Aufgrund der Instationarität können diese Zeitreihen nicht valide durch Modelle der Verteilungsannahme prognostiziert werden, sondern erfordern Modelle der Bedarfsprognose für instationären Bedarf.

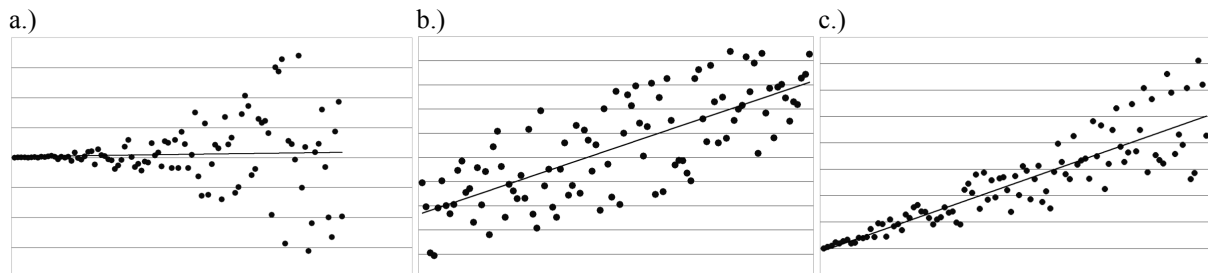


Abbildung 2.8: Beispiele instationärer Zeitreihen

Das Konzept der Stationarität kann auf die Residuen eines instationären Prognosemodells erweitert werden. In Abbildung 2.8 (b.) ist die Verteilung der Residuen um eine lineare Regressionsgerade als homoskedastische Varianz zu interpretieren, im Gegensatz zur heteroskedastischen Verteilung in Abbildung 2.8 (c.). Ist die Prämisse der Stationarität durch Veränderungen der Momente der Verteilung nicht erfüllt, müssen stochastische Modelle der Bedarfsprognose für instationären Bedarf angewendet werden.

2.2.3.2 Modelle für instationären Bedarf

Modelle der instationären, stochastischen Bedarfsprognose werden zur Vorhersage von Primär-, Sekundär und Tertiärbedarf verwendet, wenn eine deterministische Bedarfsrechnung nicht möglich oder nicht wirtschaftlich ist und instationäre Bedarfsverläufe eine Bedarfsschätzung durch Verteilungsannahme verbieten.

Die stochastische Bedarfsprognose muss unterschiedliche regelmäßige und unregelmäßige Komponenten der Bedarfsverläufe im Zeitablauf berücksichtigen. Zu den regelmäßigen Komponenten von Zeitreihen zählen das mittlere Bedarfsniveau sowie unterschiedliche Formen der Saisonalität und des Trends, die jeweils eine mittelwert-instationäre Bedarfsverteilung bedingen. Die Zeitreihenkomponenten sowie ihre Kombination zu den archetypischen Bedarfsmustern¹ wird in Abschnitt 2.3.2.1 eingehend untersucht. Diese regelmäßigen Komponenten werden zusätzlich durch unregelmäßige Komponenten der Zufallsschwankungen unterschiedlicher Intensität, der Ausreißer sowie durch Strukturbrüche und Strukturveränderungen der Komponenten überlagert, welche weiter zur Instationarität der Zeitreihen beitragen. Die Vernachlässigung regelmäßiger sowie unregelmäßiger Komponenten in der Prognose führt zu einem Ansteigen der Varianz der Bedarfsverteilung im Vergleich zur tatsächlichen Streuung und somit zu systematisch überhöhten Sicherheitsbeständen und Bestellmengen. Somit verbieten sowohl regelmäßige als auch unregelmäßige Komponenten von Zeitreihen eine valide Bedarfsschätzung anhand von Verteilungsannahmen. Vielmehr ist eine Berücksichtigung der jeweiligen Zeitreihenkomponenten durch die Anwendung statistischer Prognoseverfahren erforderlich, die für die Vorhersage der jeweiligen Form der

¹ Die Kombination der Komponenten bildet typische Zeitreihenverläufe, beispielsweise einen saisonalen Verlauf aus Niveau und Saison, einen trendbehafteten Verlauf aus Trend und Niveau oder einen trendsaisonalen Verlaufs aus Trend, Saison und Niveau (vgl. Abschnitt 2.3.2.1).

instationären Prozesse entwickelt wurden.¹ Unter der Annahme, dass sich die zukünftige Bedarfsentwicklung analog zu der vergangenen gestaltet, kommen konkurrierende objektiv-statistische, subjektiv-menschliche sowie hybride Modelle der Bedarfsprognose zum Einsatz.

Die objektiven Modelle umfassen sowohl zeitreihenanalytische Prognosemodelle unter Verwendung ausschließlich vergangener Realisationen der abhängigen Variable als auch kausalanalytische Prognosen unter Verwendung exogener Erklärungsvariablen zur Berücksichtigung kausaler Ursache-Wirkungs-Zusammenhänge, zum Beispiel durch Aktivitäten der Preis-, Produkt-, Sortiments- oder Kommunikationspolitik des Unternehmens und der Mitbewerber am Markt. Zur Lösung der jeweiligen Modelle kommen unterschiedliche konventionelle statistische Prognoseverfahren zur Anwendung, insbesondere die zeitreihenanalytischen Verfahren der Exponentiellen Glättung und der Autoregressiven Integrierten Moving-Average-(ARIMA-)Prozesse, die für die jeweiligen Modelle der stationären, trendbehafteten, saisonalen oder trend-saisonalen Zeitreihenverläufe zu spezifizieren und zu parametrisieren sind. Für kausalanalytische Modelle werden Verfahren der einfachen, multiplen sowie dynamischen Regression verwendet, ebenfalls unter A-priori-Spezifikation des Prognosemodells. Als Erweiterung konventioneller Verfahren wurden nichtparametrische, objektive Prognoseverfahren der Neuronalen Netze entwickelt, die ohne A-priori-Modellannahmen zur Abbildung der regelmäßigen Zeitreihenkomponenten oder der Verknüpfung kausaler Erklärungsvariablen zur stochastischen Bedarfsprognose eingesetzt werden können. Sie werden in späteren Kapiteln eingehend hinsichtlich ihrer Prognose- und Dispositionsgüte analysiert.

Darüber hinaus werden in der Praxis vielfach subjektive Prognosen durch menschliche Experten durchgeführt. In Abweichung zu materialwirtschaftlichen Veröffentlichungen [Schulte 1991, S. 156; Trux 1972; Hartmann 2002, S. 170] werden subjektive Prognosen durch ihre intuitive, informale Modellbildung ebenfalls zu den stochastischen Verfahren der Bedarfsrechnung gezählt und nicht als eigenständige Verfahrensklasse betrachtet. Die subjektiven Prognoseverfahren umfassen eine Vielzahl von Verfahren, die trotz irreführender Bezeichnungen eine ähnliche Form der Bedarfsprognose zur Warendisposition angeben: die „manuellen Sichtdisposition“ erfolgt durch die Bedarfsermittlung durch unregelmäßige visuelle Überprüfung der Lagerbestände in einzelnen Lagerfächern [Meuser 1998, S. 46], die „Ermessensdisposition“ erfolgt durch eine subjektive Bestandsfortschreibung [Lensing und Sonnemann 1995, S. 78], ebenso die „personellen Disposition“ und „Abverkaufsdisposition“ [Toporowski 1996]. Alle Verfahren sind als Sonderformen der subjektiven, stochastischen Prognoseverfahren zu interpretieren, die in Abschnitt 3.2.1 dargestellt werden.²

Die Güte der Prognosemodelle und Verfahren wird in der empirischen Anwendung dabei sowohl durch ihre Fähigkeit zur Abbildung der regelmäßigen Zeitreihenkomponenten als auch durch ihre Robustheit gegen unregelmäßige Komponenten der Ausreißer, Niveauverschiebungen und Strukturveränderungen bestimmt. Sie ist in Abhängigkeit von der Bedarfsstruktur der Anwendung zu bestimmen, was für die Verfahren der Neuronalen Netze in Abschnitt 5.4 auf synthetischen Zeitreihen sowie in Abschnitt 6 auf empirischen Zeitreihen überprüft wird. Die Unterteilung der Bedarfsarten in deterministischen oder stochastischen, stationären oder instationären Bedarf wird in den Modellen der Bestellmengenrechnung

¹ Die Komponenten des Bedarfs müssen dazu zunächst anhand der Zeitreihenanalyse erkannt und durch die Selektion geeigneter Prognosemodelle bei der Berechnung von Bestellmengen berücksichtigt werden. Alternativ kann die Bedarfszeitreihe durch geeignete Transformationen in einen stationären Prozess überführt werden, im Sinne einer Dekomposition mit anschließender Extrapolation der stationären Zeitreihe und Rekomposition der systematischen Komponenten (siehe Abschnitt 3).

² Durch eine subjektive Prognose und simultane Bestellmengenrechnung erfolgt mitunter keine formale Trennung der Modelle. Daher treten die Prognosemengen nicht explizit in Erscheinung sondern werden unmittelbar in Bestellmengen überführt. Dies wird nachfolgend jedoch vernachlässigt.

explizit berücksichtigt, die nachfolgend dargestellt und hinsichtlich ihrer Interaktion mit den Modellen der Bedarfsrechnung und Bedarfsprognose untersucht werden.

2.3 Modelle der Bestellmengenrechnung

2.3.1 Determinanten

2.3.1.1 Entscheidungsvariablen

Nach NADDOR [1971, S. 29 ff.] werden Bestellmengen- und Lagerhaltungssysteme durch vier Einflussfaktoren hinreichend charakterisiert: Bedarf, Bestellung, Kosten sowie Nebenbedingungen der Einflussfaktoren.¹ Ihre Bedeutung wurde durch empirische Studien und analytische Modellierung vielfach bestätigt [Silver und Peterson 1985, S. 61]. Bedarf, Kosten und Nebenbedingungen repräsentieren dabei nicht direkt beeinflussbare Zustandsparameter, während die Entscheidungsvariablen der Bestellung in Form von Bestellmengen und -zeitpunkten durch den Disponenten beeinflusst werden können [Naddor 1971, S. 33]. Ziel der Bestellmengenrechnung ist somit die Festlegung der wirtschaftlichen Beschaffungsmengen und Zeitpunkte in Abhängigkeit des ermittelten Bedarfs, der Kosten und der Nebenbedingungen.²

Nach der eingehenden Untersuchung des Bedarfs im vorangehenden Abschnitt 2.2 sind die Einflussfaktoren der Bestellung in ihren Elementen Bestellmenge, Bestellzyklus sowie Beschaffungszeit zu untersuchen [Naddor 1971, S. 34].³ Die Bestellmenge q_i [ME]⁴ ist jene Menge, die im Fall einer Bestellauslösung zur Auffüllung des Lagerbestands in Planungsperiode i bestellt wird [Fratzl 1992].⁵ Ist q_i für jede Bestellperiode i identisch, wird sie vereinfacht als Losgröße q bezeichnet [Naddor 1971, S. 39]. Die Höhe der Bestellmenge ist in der Praxis vielfach durch Nebenbedingungen in Form von Mindestbestellmengen oder Mindestbestellwerten für die Auslösung der Bestellung einer Ware beschränkt [Götz 1999, S. 65]. Damit verbunden sind auch Probleme der Verbunddisposition in Mehrproduktmodellen zu berücksichtigen, in der aufgrund von Mindestbestellmengen eines Lieferanten oder durch ein gekoppeltes Nachfrageverhalten komplementärer Waren mehrere Artikel gleichzeitig bestellt werden müssen [Götz 1999, S. 66 f.; Wahl 1999].

Üblicherweise wird angenommen, dass die bestellte Menge q_i vollständig geliefert und zum Bestand hinzuaddiert wird. Ist dies nicht der Fall, sind abweichende Liefermengen zu

¹ Die Nebenbedingungen legen in Form von administrativen, physikalischen Faktoren die Grenzen für Bedarf, Beschaffung und Kosten fest und müssen bei der Modellbildung und Entscheidungsfindung berücksichtigt werden [Naddor 1971]. Nachfolgend werden relevante Nebenbedingungen der Entscheidungsvariablen und Zustandsparameter für die Modellauswahl im Rahmen der Untersuchung der einzelnen Modelle betrachtet.

² Sie wird auch als Bestellmengenplanung [Haderl 2000] oder Bestellmengenpolitik bezeichnet [Toporowski 1998], im Fall der Produktionsplanung als Losgrößenplanung [Domschke et al. 1997]. Der Begriff des „Problems“ wird vielfach synonym mit dem des „Modells“ verwendet und so weitere Begriffsstrukturen gebildet [Domschke et al. 1997, S. 36]. NADDOR [1971, S. 34] bezeichnet Lagerhaltungsprobleme auch als Beschaffungsprobleme, da diese die Entscheidungsvariablen bilden und so die Modellbildung bestimmen. DOMSCHKE et al. [1997, S. 69] grenzen die Begriffe des Modells bzw. Problems und der Politik voneinander ab. Die Bestellmengenpolitik bezeichnet eine spezifizierte Lösung des Bestellmengenmodells durch die Festlegung der Entscheidungsvariablen, z. B. der Zyklusdauern und aller zugehörigen Bestellmengen für den betrachteten Planungszeitraum. Eine Bezeichnung als „Dispositionsregeln“ ist nachfolgend zu vermeiden, um gemäß der prozessorientierten Betrachtung einer Entscheidung der Disposition in die konstituierenden Modelle der Bedarfsrechnung und der Bestellmengenrechnung gerecht zu werden.

³ Produktionswirtschaftliche Determinanten der Lieferzeiträume und Lieferraten zur Auffüllung der Lager werden hier vernachlässigt. Siehe NADDOR [1971, S. 39 ff.] für unterschiedliche Lieferformen.

⁴ In Anlehnung an DOMSCHKE et al. [1997] wird die Dimension der Variablen und Parameter in Mengeneinheiten [ME], Zeiteinheiten [ZE], Geldeinheiten [GE] oder Kombinationen hiervon, zum Beispiel [GE]/[ME], im Text mit angegeben, um die Interpretation der Formeln zu erleichtern.

⁵ In der Literatur zur Bestellmengenrechnung existieren diverse abweichende Notationen. Beispielsweise wird die Bestellmenge sowohl als Q [Assfalg 1976; Hillier und Lieberman 1997; Hochstädter 1969; Lee und Nahmias 1993; Silver und Peterson 1985; Silver et al. 1998; Toporowski 1996], als q [Schneeweiß 1981; Domschke et al. 1997; Götz 1999; Naddor 1971] oder als x bezeichnet [Grochla 1992; Trux 1972]. Nach NADDOR [1971] wird nachfolgend die Bestellmenge als q bezeichnet.

berücksichtigen, die ebenso wie der Bedarf eine deterministische oder stochastische sowie stationäre oder instationäre Form annehmen können.

Der Bestelltermin t_i [-] bezeichnet jenen Zeitpunkt innerhalb der i -ten Planungsperiode, zu dem die Bestellung in Höhe von q ausgelöst wird. Die Zeit zwischen zwei aufeinanderfolgenden Bestellterminen t_i und t_{i+1} bezeichnet die Dauer des Bestellzyklus τ [ZE] [Naddor 1971, S. 34 ff.], mit

$$\tau_i = t_{i+1} - t_i \quad . \quad (2.4)$$

Bestellzyklen können von variierender oder gleicher Dauer sein; im Fall konstant gewählter Bestellzyklen wird von einem Bestellrhythmus r gesprochen, mit $\tau_i = \tau_{i+1} = \dots = r$ [Naddor 1971, S. 34 ff.].¹ Sind Bestelltermine und Bestellzyklen durch den Disponenten bestimmbar, gelten sie als Entscheidungsvariable des Modells. Sind sie hingegen als Zustandsparameter exogen vorgegeben, stellen sie keine Entscheidungsvariable im eigentlichen Sinne dar, und die Bestellmenge q repräsentiert die einzige Entscheidungsvariable des Modells.²

Die Bestellmenge q ist nach der Wiederbeschaffungszeit Δt_W [ZE] am Lager verfügbar, die den Zeitraum zwischen dem Erkennen einer Bestellnotwendigkeit und der anschließenden Verfügbarkeit im Lager angibt [Götz 1999, S. 53]. Die Wiederbeschaffungszeit besteht bei kontinuierlich inspizierten Systemen aus der Auftragsvorbereitungszeit Δt_A [ZE], der Lieferzeit Δt_L [ZE] und der Einlagerungszeit Δt_E [ZE] [Trux 1972, S. 232 f.] sowie bei nicht kontinuierlich überwachten Systemen zusätzlich aus dem Überprüfungsintervall der mittleren Zeit bis zum Erkennen des Bestellpunktes Δt_U [ZE] [Götz 1999, S. 58; Oeldorf und Olfert 2002, S. 226].³ Somit errechnet sich der Termin des Wareneingangs t'_i [-] als

$$t'_i = t_i + \Delta t_W = t_i + (\Delta t_A + \Delta t_L + \Delta t_E + \Delta t_U) \quad , \quad (2.5)$$

mit einer Wiederbeschaffungszeit von

$$\Delta t_W = t'_i - t_i \quad . \quad (2.6)$$

Da die Wiederbeschaffungszeit im Allgemeinen exogen vorgegeben ist [Naddor 1971, S. 34], bildet sie im Gegensatz zu den Entscheidungsvariablen von Bestellzyklus und Bestellmenge einen Parameter der Bestellmengenrechnung. Die Wiederbeschaffungszeit kann ebenso wie Bedarf und Liefermengen im Zeitablauf konstant oder variabel sowie deterministischer oder stochastischer Art sein. Bei geringen Beschaffungszeiten wird Δt_W oftmals vernachlässigt.

¹ DOMSCHKE et al. [1997, S. 69] bezeichnen den Bestellzyklus auch als Zyklusdauer oder Reichweite einer Bestellmenge q , da im klassischen Bestellmengenmodell das Lager nach τ Zeiteinheiten auf q Mengeneinheiten aufgefüllt wird. NADDOR [1971, S. 35] unterscheidet zudem deterministisch schwankende von stochastisch variierenden Bestellzyklen, auf die nachfolgend nicht eingegangen wird. Ist die Bestellmenge q oder die Zyklusdauer τ über alle Perioden des Planungszeitraums zeitinvariant, bezeichnen DOMSCHKE et al. [1997, S. 69 f.] die Bestellmengenpolitik als q -stationär bzw. τ -stationär. Bestellmengenpolitiken, die sowohl q -stationär als auch τ -stationär sind, heißen „stationär“, solche, in denen eine der Größen nicht stationär ist, repräsentieren „instationäre“ Bestellmengenpolitiken.¹ Um Verwechslungen mit der ähnlich lautenden Charakterisierung der Bedarfsarten zu vermeiden wird diese Bezeichnung nachfolgend nicht verwendet, auch wenn beispielsweise ein τ -stationären Bestellzyklus [Domschke et al. 1997, S. 161] auf einen stationären Bedarfsverlauf zurückgeführt werden kann.

² Darüber hinaus können einzelne Entscheidungsvariablen der Bestellzyklen, Überprüfungsperioden, Bestandshöhen oder sogar eine spezifische Bestellmengenpolitik durch die Betriebsleitung vorgeschrieben sein, was ebenso als Nebenbedingung in der Modellbildung zu berücksichtigen ist [Naddor 1971, S. 48].

³ Sie geht somit über die Zeit zwischen dem Treffen der Bestellentscheidung und der Verfügbarkeit der Ware am Lager im Sinne der Definition nach NADDOR [1971, S. 36] hinaus.

2.3.1.2 Kostenparameter

Neben den Einflussfaktoren der Bestellung sind nach NADDOR [1971, S. 13, 29, 42 ff.] die Zustandsparameter der Kosten zu berücksichtigen, die von den getroffenen Entscheidungen der Bestellmengen und -zeitpunkte sowie den Nebenbedingungen abhängen. Nach SILVER et al. [1998, S. 44-48] sowie NADDOR [1971, S. 42-46] werden die Dispositionsentscheidungen vor allem durch die Kostenfaktoren Lagerungskosten, Beschaffungskosten und Fehlmengenkosten beeinflusst.¹

Die Lagerungskosten C_L [GE] umfassen alle Kosten, die bei der Lagerung von Waren durch die Bindung von Kapital und Liquidität, die Nutzung des Lagerraums sowie durch die Behandlung der Gegenstände im Lager und durch Wertminderungen aus Alterung, Beschädigung, Diebstahl, Schwund und Verderb und damit korrespondierende Versicherungen entstehen [Naddor 1971, S. 13; Götz 1999, S. 60; Grochla 1992, S. 75]. Als bedeutendste Kostenkomponente sind die Kapitalbindungskosten hervorzuheben, die aus dem zeitlichen Auseinanderfallen der Geldströme entstehen: einerseits der Auszahlungen zur Beschaffung der Waren auf der Versorgungsseite, andererseits der späteren Einzahlungen aus dem Verkauf der Ware auf der Nachfrageseite (siehe Abschnitt 2.1.3). Diese werden durch den Kapitalbedarf an Fremdkapitalzinsen, kalkulatorischen Eigenkapitalzinsen sowie Opportunitätskosten einer alternativen Verwendung bewertet [Assfalg 1976]. Die variablen Lagerhaltungskosten c_L [GE/ME*ZE] fallen proportional zum gebundenen Kapital im physischen Lagerbestand S [ME] an. Sie werden alternativ als Bruchteil des Produktwertes der gelagerten Ware pro Zeiteinheit ausgedrückt, und zwar durch das Produkt von Einstandspreis p [GE/ME] der Ware pro Stück und einen Lagerkostensatz l in Prozent des gebundenen Kapitals je Zeiteinheit [%/ZE] [Toporowski 1996, S. 198; Naddor 1971, S. 44; Lee und Nahmias 1993]. Aus Gründen der Vereinfachung wird vielfach ein konstanter Lagerabgang und Mittelrückfluss durch stationäre Bedarfs-, Waren- und Geldflüsse in einer Planungsperiode unterstellt. Somit werden die Lagerhaltungskosten c_L proportional zur Lagerzeit des Bestellzyklus τ [ZE] und zum mittleren Lagerbestand $S/2$ [ME] bewertet [Oeldorf und Olfert 2002, S. 297-300; Grochla 1992, S. 49; Hillier und Lieberman 1997, S. 606], mit

$$C_L = \tau \frac{S}{2} c_L = \tau \frac{S}{2} p \cdot l \quad . \quad (2.7)$$

Bestimmend für die Höhe der Kapitalbindungskosten ist jedoch die tatsächliche zeitliche Verteilung von Lagerbestandsmenge, -wert und -dauer in der Planungsperiode [Naddor 1971, S. 29-34] sowie der anzusetzende kalkulatorische Zinssatz [Götz 1999, S. 61]. In der Bestimmung der Lagerhaltungskosten können auch Zusatzrestriktionen der effektiven und verfügbaren Lagerkapazität des physischen Lagerraums [Brink 1988] sowie der prozessorientierten Leistungsfähigkeit und der Kapazität an Ein-, Um- und Auslagerungsvorgängen berücksichtigt werden [Götz 1999, S. 64-67]. DOMSCHKE et al. [1997, S. 65 f.] betrachten neben den zeitlichen und mengenmäßigen auch monetäre Kapazitätsrestriktionen und erweitern so die Modellierung zur Berücksichtigung der Finanzierungsseite. Diese können durch die Berücksichtigung erhöhter Handhabungskosten oder einer zusätzlichen Anmietung von Lagerräumen durch nichtlineare, sprungfixe Lagerungskosten abgebildet und einheitlich in der Modellierung berücksichtigt werden [Domschke et al. 1997, S. 66].

¹ Eine umfassende Beschreibung aller Kostengrößen findet sich bei GROCHLA [1992, S. 75 ff.]. SILVER et al. geben darüber hinaus eine umfangreiche Auflistung potenzieller Variablen [1998, S. 48 ff.]. OELDFORT und OLFERT [2002, S. 295-302] nennen im Kontext der Materialwirtschaft zudem die Faktoren der Losgrößen-einheiten und des Finanzvolumens, die jedoch im Kontext der Handelsbetriebe vernachlässigbar erscheinen. Eine umfangreiche Diskussion der einzelnen Kostenarten und -größen für Bestellmengenmodelle des Handels findet sich in [Götz 1999, S. 56-64; Toporowski 1996].

Die Bestellkosten C_B [GE] umfassen alle fixen und variablen Kostenkomponenten, die beim Auslösen einer zusätzlichen Bestellung zum Fremdbezug von Waren oder Material entstehen [Assfalg 1976, S. 19 ff.; Oeldorf und Olfert 2002, S. 295].¹ Die fixen Kosten je Bestellung K_B [GE] werden durch die Häufigkeit der Bestellungen in der Planungsperiode bestimmt und sind unabhängig von der Bestellmenge [Grochla 1992, S. 74]. Sie umfassen die Sach- und Personalkosten der Warendisposition durch Bedarfsmeldung, Bestellabwicklung, Liefertermin- und Liefermengenüberwachung, Warenannahme, -prüfung und -einlagerung sowie die administrativen Tätigkeiten des Rechnungswesens je Bestellvorgang [Götz 1999, S. 57].² Zusätzlich entstehen durch die Bestellung variable Stückkosten c_B [GE/ME] proportional zu der Menge der bestellten Waren q [ME]. Die variablen Stückkosten werden mitunter vereinfachend mit den Beschaffungspreisen p [GE/ME] gleichgesetzt. Sie können jedoch auch komplexe Kostenstrukturen durch die Zulassung von Mengenrabatten sowie Mindermengenaufpreisen und zusätzlichen Transportkosten, Steuern und Zöllen abbilden [Grochla 1992, S. 74], und werden im Folgenden je nach Modellprämisse spezifiziert. Vereinfachend wird zunächst von linearen Stückkosten ausgegangen, wobei die Bestellkosten nur dann auftreten, wenn die Bestellmenge $q > 0$ ist, mit

$$C_B(q) = \begin{cases} K_B + c_B q & \text{für } q > 0 \\ 0 & \text{für } q \leq 0 \end{cases} \quad (2.8)$$

In der Handelspraxis verursachen Mengenrabatte und diskrete Rabattstaffeln unterschiedliche nichtlineare, konvexe und konkave Kostenverläufe, die in Modellen zunehmender Komplexität ebenfalls berücksichtigt werden können [Hochstädter 1969, S. 10].³ Die fixen Bestellkosten können zudem zur Berücksichtigung von Verbundeffekten zwischen einzelnen Produkten bei Sammelbestellungen erweitert werden [Domschke et al. 1997, S. 66]. MÜLLER-HAGEDORN [1998, S. 527] untersucht die variablen Kosten aus $c_B q$ als eigenständige Kostenart und bezeichnet sie als Warenkosten. NADDOR [1971, S. 13] vernachlässigt die Warenkosten der Bestellung in seiner Analyse, da diese unabhängig von den Entscheidungsvariablen Bestellmenge und -zeitpunkt entstehen (vgl. Abschnitt 2.3.2.1). Demnach sind für die Entscheidungsfindung in Lagerhaltungssystemen nur die Kostenarten Lagerung, Beschaffung und Fehlmengen relevant.⁴

Werden mehr Wareneinheiten nachgefragt, als auf einer Lagerstufe verfügbar sind, entstehen Fehlmengen [Assfalg 1976, S. 17]. Die Einflussfaktoren der Bestellung können ebenso wie der Bedarf stochastischer Art sein, sofern Unsicherheit über die Mengen und Zeitpunkte ihrer Eingangs- und Ausgangsströme im Lagerhaltungssystem besteht (siehe Abschnitt 2.1.3 und Abbildung 2.3).⁵ Sind die Wiederbeschaffungszeiten, die Lagerzugänge von Lieferanten, die

¹ GROCHLA [1992, S. 74 f.] bezeichnet die fixen Bestellkosten auch als mittelbare Kosten oder feste Bezugskosten, die variablen Bestellkosten als unmittelbare Kosten der Bestellung.

² In den fixen Bestellkosten werden vereinzelt auch die Kosten der Einkaufsfunktion durch die Beschaffungsmarktforschung und die Lieferantenauswahl berücksichtigt und proportional auf die Anzahl der Bestellungen umgelegt, während diese hier der vorgelagerten Prozessstufe des Einkaufs zuzuordnen sind.

³ ASSFALG [1976, S. 21 f.] sowie HADLEY und WHITIN [1963, S. 39] erweitern darüber hinaus auch die Annahme der fixen Bestellkosten zur realistischeren Form der sprungfixen Kosten.

⁴ NADDOR [Naddor 1971] vernachlässigt somit jedoch auch die Relevanz von Mengenrabatten in der Bestellung, welche SILVER et al. [1998, S. 159-164] intensiv betrachten. Dabei sind nach NADDOR in Lagerhaltungssystemen mindestens zwei der Kostenarten durch den Entscheidungsträger beeinflussbar, während in Produktionssystemen alle Kosten beeinflussbar sind [Naddor 1971, S. 13]. Die Beeinflussbarkeit bestimmt darüber hinaus die Typen von Lagerhaltungssystemen, auf die hier nicht explizit eingegangen wird.

⁵ Dabei berücksichtigt die Bedarfsrechnung ausschließlich die Unsicherheiten der Bedarfskomponente in Menge und Zeit. Alle weiteren Komponenten der Lagerabgänge und Lagerzugänge, insbesondere die vom Lieferanten befriedigte Bestellmenge, Komponenten der Wiederbeschaffungszeit sowie die realisierten Bestelltermine, werden in Menge und Zeitpunkt durch die Bestellmengenrechnung berücksichtigt. Die Dekomposition der Warendisposition zur Lagersteuerung in disjunkte Teilprobleme der Bedarfs- und Be-

Lagerabgänge an Kunden oder der Bedarf stochastischer Art,¹ kann keine unmittelbare und vollständige Bedarfsdeckung aus dem Lagerbestand garantiert werden, wodurch Fehlmengen entstehen [Götz 1999, S. 54]. Eine Fehlmenge bezeichnet jenen Bedarf, der nicht durch den vorhandenen Lagerbestand gedeckt werden kann [Hochstädter 1969].² Das Auftreten von Fehlmengen verursacht Fehlmengenkosten C_F [ME], die in der stochastischen Warendisposition und Bestellmengenrechnung berücksichtigt werden müssen [Toporowski 1998; Bichler und Lörsch 1985, S. 13]. Anhand des Verhaltens der Kunden werden zwei Fehlmengenszenarien mit unterschiedlichen Kostenstrukturen unterschieden: der Vormerkfall unter Nachlieferung der Fehlmengen und der Verlustfall, in dem der Bedarf in Höhe der Fehlmenge verloren geht [Domschke et al. 1997, S. 67].

Im Vormerkfall (engl. *back-order case*) wird der Kundenbedarf durch Vormerken der Fehlmenge und Nachlieferung zu einem späteren Zeitpunkt befriedigt [Götz 1999, S. 62], wodurch Fehlmengenkosten von $C_{F,V}$ [GE] durch zusätzliche Beschaffungskosten und kalkulatorische Kosten der verringerten Kundenzufriedenheit entstehen. Zusätzliche Beschaffungskosten entstehen durch Sonder- und Teillieferungen aus übergeordneten Lagerstufen, Zuschläge für Mindermengen und Eillieferungen der Beschaffung, die Lieferung höherwertiger Ersatzartikel, die Verwendung alternativer Transportformen und -wege, zusätzliche Produktionsaufträge und Änderungen der Produktions- und Ablaufplanung der Lieferanten sowie zusätzliche Verwaltungskosten, die direkt quantifizierbar sind [Götz 1999, S. 62]. Sie werden als einmalige Fehlmengenkosten $c_{F,V}$ [GE/ME] für jede Mengeneinheit der Fehlmenge q_F [ME] verursacht. Darüber hinaus entstehen aus der verminderten Lieferbereitschaft kalkulatorische Kosten durch eine verringerte Kundenzufriedenheit und eine kurz- und langfristige Verschlechterung des Ansehens und der Vertrauenswürdigkeit des Betriebes, was auch als Verlust der Gunst (engl. *goodwill*) bezeichnet wird [Naddor 1971]. Die beim Kunden gesunkene Erwartung hinsichtlich der zukünftigen Lieferfähigkeit äußert sich unmittelbar oder verzögert in einer verringerten Bereitschaft zur Pflege der zukünftigen Geschäftsbeziehungen [Assfalg 1976]. Dies kann je Kunde zu einer in Menge und Häufigkeit reduzierten Nachfrage und einer Veränderung der Beschaffungsbeziehungen in der Zukunft führen [Götz 1999, S. 62; Hochstädter 1969] und bedeutet für den Handelsbetrieb den möglichen Verlust zukünftiger Gewinne. Diese Kostengröße ist durch kundenindividuelle Einstellungen, die Dringlichkeit der Bedarfsbefriedigung beim Kunden, die zeitliche Dauer und den Grad der Lieferunfähigkeit sowie die dynamische Entwicklung aller Faktoren jedoch nur schwer quantifizierbar [Götz 1999, S. 63]. Anstelle einer Bewertung der betroffenen Kunden, Aufträge oder Nachlieferungen erfassen HADLEY und WHITIN [1963, S. 19] die Fehlmengenkosten proportional zur Höhe des Menge der vorgemerkten Waren. Dabei wird vereinfachend angenommen, dass der Verlust an Goodwill, bewertet mit einem Kostensatz $c_{G,V}$ [GE/ME] * [ZE], proportional zur zeitlichen Dauer der Fehlmengensituation Δt_F [ZE] und zur

stellmengenrechnung trennt somit lediglich den Output-Fluss des Bedarfs von allen anderen Input- und Output-Flüssen sowie von den Kostenkomponenten der Lagerhaltung.

¹ Sie erfordern in diesem Fall ebenfalls eine entsprechende Prüfung auf Stationarität sowie eine Prognose durch ein geeignetes Verfahren [Hochstädter 1969], was nachfolgend jedoch vernachlässigt wird.

² Fehlmengen werden im Handel uneinheitlich als Fehlbestand, Fehlbedarf, Versorgungs- oder Präsenzlücke bezeichnet [siehe u.a. die Beiträge in Atzberger 1998b]. Dabei ist zu unterscheiden, ob im Fall einer Fehlmenge ein Bedarf nicht gedeckt werden konnte oder ob eine Präsenzlücke (engl. *stockout* oder *out-of-stock*, *OOS*) mit einem Warenbestand von null Mengeneinheiten ohne Bedarfsverlust aufgetreten ist. Letzteres stellt einen optimalen Zustand in der stochastischen Warendisposition dar, sofern aller Bedarf befriedigt wurde, während im Fall einer Fehlmenge nicht der gesamte Bedarf befriedigt werden konnte. In Lagerstufen des Verkaufsräumens bzw. im Regal des Selbstbedienungshandels kann lediglich das Auftreten einer Präsenzlücke erhoben werden, aufgrund der möglichen Substitution durch andere Produkte oder Packungsgrößen jedoch nicht der tatsächlich eingetretene Umsatz- und Goodwill-Verlust. Da hierbei keine Unterscheidung zwischen einer Präsenzlücke und einem Bedarfsverlust erfolgen kann, entstehen besondere Probleme bei der Erhebung der tatsächlichen Höhe von Fehlmengen, die hier in Übereinstimmung mit der Literatur vernachlässigt werden. Nachfolgend werden Präsenzlücken eines OOS synonym mit einer Fehlmenge behandelt.

Höhe der Fehlmenge q_F [ME] entsteht [Götz 1999, S. 62]. Die Fehlmengenkosten im Vormerkfall ergeben sich somit durch

$$C_{F,V} = (c_{F,V} + c_{G,V} \Delta t_F) q_F \quad . \quad (2.9)$$

Im Verlustfall (engl. *lost-sales case*) resultiert ein nicht unmittelbar befriedigter Bedarf dagegen in Kosten $C_{F,L}$ [GE] aus dem entgangenen Gewinn der verlorenen Nachfrage sowie zukünftigen Gewinnausfällen durch den Verlust an Goodwill [Assfalg 1976, S. 26]. Der Verlust der Nachfrage verursacht unmittelbar Opportunitätskosten der Fehlmengen $c_{F,L}$ [GE/ME] aus dem entgangenen Gewinn bzw. Deckungsbeitrag pro Stück proportional zur Höhe der Fehlmenge q_F [Assfalg 1976, S. 26 f.; Hadley und Whitin 1963, S. 20 f.].¹ Zudem verursacht eine Fehlmenge im Verlustfall einen zeitunabhängigen Verlust des Goodwill $c_{G,L}$ [GE/ME] und somit zukünftiger Gewinne [Götz 1999, S. 55, 63]. Die Fehlmengenkosten im Verlustfall errechnen sich somit als

$$C_{F,L} = (c_{F,L} + c_{G,L}) q_F \quad .^2 \quad (2.10)$$

ASSFALG [1976, S. 27 f.] führt darüber hinaus auch die Kosten für die Anwendung der Modelle und Verfahren der Bedarfs- und Bestellmengenrechnung an, die auch als Kosten der Systemkontrolle [Silver et al. 1998, S. 48] oder Optimierung [Götz 1999, S. 64] bezeichnet werden. Unterschiedliche Verfahren verursachen unterschiedlich hohe Kosten in der Güte der ermittelten Warenbestände, aber auch in der Informationsbeschaffung und -verarbeitung sowie der Installation und Handhabung durch das Personal. Obwohl diese Kosten der Verfahrensauswahl in der Einführung und regelmäßigen Durchführung nicht vernachlässigt werden sollten, werden sie in der Lagerhaltungsliteratur kaum diskutiert [Assfalg 1976, S. 27]. Sie sind nachfolgend zu berücksichtigen, da die unterschiedlichen Verfahrensansätze statistischer Prognoseverfahren und Neuronaler Netze unterschiedliche Anforderungen an die Ressourcen stellen.

Die drei Kostenarten Lagerungskosten C_L , Bestellkosten C_B und Fehlmengenkosten C_F sind interdependent miteinander verknüpft, so dass die Verringerung einer Kostenkomponente zur Erhöhung anderer Kostenarten führt [Naddor 1971, S. 14]. Ein positiver Lagerbestand verursacht Lagerungskosten, die mit der Bestandshöhe wachsen. Daher ist der Lagerbestand möglichst niedrig zu halten. Ein geringer Lagerbestand erhöht aber für einen gegebene Bedarf der Planungsperiode die Häufigkeit der Bestellungen mit kleineren Bestellmengen und gleichzeitig die Wahrscheinlichkeit von Fehlmengen, die zu Kosten führen, die mit der Häufigkeit der Bestellungen und der erwarteten Höhe der Fehlbestände steigen [Assfalg 1976, S. 24]. Ziel des Entscheidungsproblems der Lagerhaltung ist somit die Bestimmung optimaler Entscheidungsvariablen zur Minimierung der Gesamtkosten C [GE] mit

¹ Die Auswirkungen der entgangenen Umsatz- und Ertragssteigerungen durch Mehrverkäufe auf der Beschaffungsseite werden weitgehend vernachlässigt, obwohl diese mit den Bestellmengen interagieren.

² HOCHSTÄDTER unterscheidet darüber hinaus ein drittes Fehlmengen-Szenario des Kundenverlustes (engl. *lost consumer case*), in dem die Nachfrage des Kunden verloren geht und zudem der Kunde nicht erneut von der Lagerstufe kaufen wird, so dass alle zukünftigen Bedarfe unmittelbar verloren gehen [Hochstädter 1969]. Nach MINNER [2000, S. 13] ist die Modellierung verllorener Kunden in Bestellmengenmodellen besonders schwierig, da sie eine Veränderung des Bedarfsniveaus impliziert und so ebenfalls durch eine Anpassung in der Bedarfsprognose berücksichtigt werden muss. Zudem implizieren die hier dargestellten Alternativen eine grundsätzliche Vereinfachung, da Fehlmengen in der Praxis sowohl zu einer verlorenen Nachfrage als auch zu Vormerkungen führen können. Zudem können wiederholte Fehlmengen zeitlich verzögert zu einem Kundenverlust führen, so dass auch zeitlich veränderbare, gemischte Effekte zu berücksichtigen wären [Minner 2000, S. 13]. Dies übersteigt jedoch den Rahmen dieser Arbeit und wird daher vernachlässigt.

$$C = C_L + C_B + C_F \quad . \quad (2.11)$$

Aufgrund der hohen Sach- und Personalkosten aller mit der Warendisposition verbundenen Prozesse erhalten die fixen Bestell- und Rüstkosten in Lagerhaltungssystemen der Produktion vielfach eine besondere Bedeutung [Götz 1999, S. 59-60; Hillier und Lieberman 1997, S. 603]. In der Warendisposition des Handels überwiegt dagegen eine hohe Bestellfrequenz mit häufigen, bis zu täglichen Bestellterminen [Atzberger 1998b], was auf eine größere Bedeutung der Kostenkomponenten der Lager- und Fehlmengenkosten hinweist. Somit sind die Entscheidungsvariablen der Warendisposition in Abhängigkeit von den Kosten situativ zu bestimmen.

Können die Fehlmengenkosten C_F nicht valide bestimmt werden, wird das Auftreten von Fehlmengen nicht durch Kosten, sondern durch eine Beschränkung der Häufigkeit oder der Höhe ihres Auftretens in Form von Servicegraden berücksichtigt [Assfalg 1976, S. 27].¹ Der Servicegrad bezeichnet das Verhältnis von befriedigter zu unbefriedigter Nachfrage und dient insbesondere bei einer stochastischen Nachfragestruktur als eine Maßzahl zur Beurteilung der Lieferbereitschaft eines Lagers [Müller-Hagedorn 1998, S. 545].² Nach der Berücksichtigung der Häufigkeit des Auftretens oder der mengenmäßigen Höhe der Fehlmengen werden α - und β -Servicegrade unterschieden [Toporowski 1998, S. 1150; Götz 1999; Oeldorf und Olfert 2002, S. 219]. Der α -Servicegrad [%] ist eine ereignisorientierte Kennziffer, welche die relative Häufigkeit von Fehlmengenereignissen einer Planungsperiode ohne Berücksichtigung der Höhe oder Dauer der Fehlbedarfe angibt [Domschke et al. 1997, S. 159]:³

$$\alpha = \frac{\text{Anzahl der Perioden mit Bedarfsdeckung}}{\text{Anzahl aller Perioden}} \quad . \quad (2.12)$$

Der β -Servicegrad [%] stellt dagegen eine mengenorientierte Kennziffer dar, die den Anteil der unmittelbar aus dem Lagerbestand befriedigten Nachfrage an der Gesamtnachfrage aller Bedarfe einer Planungsperiode angibt:

$$\beta = \frac{\text{Befriedigte Nachfrage pro Periode}}{\text{Gesamtnachfrage pro Periode}} \quad . \quad (2.13)$$

Die Servicegrade können in der Modellbildung der Bestellmengenrechnung anstelle der Fehlmengenkosten als zusätzliche Komponente in der Zielfunktion oder als Nebenbedingung neben den Lagerungs- und Bestellkosten berücksichtigt werden [Schneeweiß 1981].⁴

¹ Nach SILVER et al. [1998] werden Lagerhaltungssysteme neben den Einflussfaktoren des Bedarfs und der Kostenfaktoren auch durch die Definition der Servicegrade determiniert. Da diese jedoch durch die Höhe der Fehlmengenkosten bestimmt werden, sind sie der Diskussion der Kostenkomponenten unterzuordnen.

² Beispielsweise wird für unterschiedliche Warengruppen und Lagerstufen des Handels das Einhalten eines vorgegebenen Servicegrades von 98% oder 99,5% gefordert und so die Wahrscheinlichkeit von Fehlmengen auf 2% bzw. 0,5% begrenzt [Toporowski 1998, S. 1150; Götz 1999]. Der Servicegrad wird auch als Grad der Lieferbereitschaft zur sofortigen Erfüllung der externen Warenanforderungen bezeichnet [Brink 1988].

³ SCHNEEWEIß [1981, S. 100 ff.], MÜLLER-HAGEDORN [1998, S. 543 f.] sowie TOPOROWSKI [1998] definieren dagegen den Servicegrad als die Wahrscheinlichkeit des Eintritts einer Fehlmengensituation in einem Bestellzyklus und bestimmen ihn daher aus dem Quotienten der Erwartungswerte der jeweiligen Größen. Dann gibt der α -Servicegrad die Wahrscheinlichkeit dafür an, dass ein Bedarf mit beliebiger Menge in einem Bestellzyklus vollständig aus dem Lagerbestand befriedigt werden kann, während der β -Servicegrad die Wahrscheinlichkeit für den Anteil der befriedigten Nachfragemenge aus dem Lagerbestand ohne Berücksichtigung der Verteilung über die Bedarfe einzelner Kunden angibt.

⁴ Nach dem ökonomischen Prinzip kann ebenso ein vorbestimmter Servicegrad zu minimalen Kosten oder ein maximaler Servicegrad zu festgelegten Kosten berechnet werden. In Theorie und Praxis sowie der Literatur zur Bestellmengenrechnung dominiert das Ziel der Kostenminimierung unter einem gegebenen Servicegrad [Toporowski 1996], obwohl dies keine Voraussetzung ist. Dabei muss die Höhe des Servicegrades jedoch

In der Vergangenheit bestanden erhebliche Probleme bei der Erhebung und monetären Bewertung von Fehlmengensituationen in operativen Prozessen [Granger 1980], so dass vielfach auf eine explizite Modellierung von Fehlmengenkosten zugunsten einer Abbildung durch Servicegrade verzichtet wurde. Die Verbreitung der Konzepte der Deckungsbeitrags- und der Prozesskostenrechnung haben jedoch wesentlich zur Transparenz der Kosten beigetragen. Die valide Erhebung und Zurechnung der Kostenkomponenten wird im Handel durch den Fortfall der Produktionsprozesse erleichtert. Der Kostenrechnungsansatz der direkten Produktrentabilität [Seifert 2001; 2004] belegt die Möglichkeit ihrer Quantifizierung insbesondere im Einzelhandel.

2.3.1.3 Modellauswahl

Modelle der Bestellmengenrechnung sind vereinfachte Abbildungen empirischer Lagerhaltungsprobleme, die je nach gegebener Situation und getroffenen Annahmen über die Entscheidungsvariablen und Parameter unterschiedliche Problem- und Modellkomponenten umfassen können. Durch die schrittweise Annäherung der Modellkomplexität an die Realität werden unterschiedliche Modellansätze gebildet, die anhand der einbezogenen Determinanten kategorisiert werden können.

Aufgrund der Bedeutung der Bedarfssituation für diese Untersuchung werden die Modelle der Bestellmengenrechnung ebenfalls nach den Determinanten des Bedarfs unterschieden (siehe Abschnitt 2.2.1 sowie Abbildung 2.6).¹ Die besondere Bedeutung der Bedarfsstruktur für die Bestellmengenrechnung wird vielfach hervorgehoben: „It is generally the assumption about demand that is most important in determining the complexity and structure of an inventory model“ [Lee und Nahmias 1993, S. 18].² NADDOR führt dazu weiter aus: „Der Bedarf wird vor allen anderen Komponenten der Lagerhaltungssysteme diskutiert, weil er die wesentlichste Komponente unter den Einflußfaktoren eines Lagerhaltungssystems ist. Lagerbestände werden gehalten, damit Bedarf gedeckt, Bestellungen erfüllt, Anforderungen befriedigt werden können. Lagerhaltungsprobleme existieren nur, weil Bedarf auftritt; andernfalls gibt es keine Lagerhaltungsprobleme“ [Naddor 1971, S. 29].

Somit wird zunächst je nach dem Grad der Unsicherheit zwischen deterministischem oder stochastischem Bedarf unterschieden; anschließend wird je nach der Konstanz des Bedarfs im Zeitablauf zwischen stationärem und instationärem Bedarf unterschieden. Durch Kombination der Determinanten des Bedarfs ergeben sich folgende Klassen von Modellen zur Bestellmengenrechnung:

- (1) Stationär-deterministische Modelle
- (2) Instationär-deterministische Modelle
- (3) Stationär-stochastische Modelle
- (4) Instationär-stochastische Modelle

Innerhalb dieser Modellklassen werden spezifische Modelle gemäß der berücksichtigten Nebenbedingungen der Entscheidungsvariablen gebildet: periodische oder kontinuierliche Bestellzeitpunkte, fixe oder variable Bestellmengen und die Berücksichtigung oder Vernach-

ebenfalls anhand der verbundenen Kosten der Disposition, Lagerung und Distribution bestimmt werden, so dass ein beliebig festgesetzter Servicegrad nicht zu effizienten und effektiven Bestellmengen führt.

¹ Diese entscheidungsorientierte Kategorisierung der Modelle der Bestellmengenrechnung nach den Determinanten des Bedarfsverlaufs erfolgt in Übereinstimmung mit ASSFALG [1976, S. 28 ff.], DOMSCHKE et al. [1997, S. 64 ff.] sowie GÖTZ [1999, S. 67 f.], da die Nachfrage das bedeutendste Einteilungskriterium bildet und sich das auszuwählende Bestellverfahren erst als Lösung des Bestellmengenproblems ergibt. Die Möglichkeit der Berücksichtigung eines negativen Bedarfs durch Warenrückgabe der Verbraucher wird hier nicht berücksichtigt [Naddor 1971, S. 48].

² Dennoch berücksichtigen LEE und NAHMIAAS [1993, S. 6] keinen instationären stochastischen Bedarf in ihren Ausführungen, sondern gehen grundsätzlich von stationären Verteilungen aus.

lässigung von Lieferzeiten. Ferner werden unterschiedliche Modelle je nach Berücksichtigung der Zustandsparameter gebildet: Behandlung von Fehlmengen durch Vormerkung oder Nachfrageverlust, Ausprägung der Liefermengen und -zeiten, Berücksichtigung der Kostenstrukturen aus variablen oder fixen, linearen oder nichtlinearen sowie im Zeitablauf veränderbarer Kosten. Unter Berücksichtigung aller Nebenbedingungen können unterschiedliche Modelle mit zunehmender Realitätsnähe und Komplexität gebildet werden [Assfalg 1976].¹ DOMSCHKE et al. [1997, S. 68] unterscheiden Modelle zudem nach der Anzahl simultan geplanter Produkte, der Anzahl der Dispositionsstufen und der Berücksichtigung von Kapazitätsrestriktionen. Innerhalb dieser Taxonomie von Modellklassen unterscheiden sie weitere Modellvarianten nach dem Umfang des Planungszeitraums, der Konstanz der Parameter sowie der Berücksichtigung von Fehlmengen. In Tabelle 2.2 ist eine Systematik der Strukturdimensionen von Modellen zur Bestellmengenrechnung dargestellt.

Tabelle 2.2: Systematik der Strukturdimensionen von Bestellmengenmodellen²

Unsicherheit des Bedarfs	Deterministische Bestellmengenmodelle		Stochastische Bestellmengenmodelle		
	Modelle für stationären Bedarf	Modelle für instationären Bedarf	Modelle für stationären Bedarf	Modelle für instationären Bedarf	
Unsicherheit der Wiederbeschaffungszeit	Keine Lieferzeiten	Deterministische Wiederbeschaffungszeit		Stochastische Wiederbeschaffungszeit	
Schwankung der Wiederbeschaffungszeit	Keine Schwankungen	Stationäre Lieferzeit	Instationäre Lieferzeit	Stationäre Lieferzeit	Instationäre Lieferzeit
Berücksichtigung von (Absatz-)Fehlmengen	Keine Fehlmengen	Berücksichtigung von Fehlmengenkosten		Berücksichtigung von Servicegraden	
Art der Fehlmengen	Keine Fehlmengen	Fehlmengen im Vormerkfall	Fehlmengen im Verlustfall	α -Servicegrade	β -Servicegrade
Berücksichtigung von Liefer-Fehlmengen	Keine Liefer-Fehlmengen	Berücksichtigung von Liefer-Fehlmengenkosten		Berücksichtigung von Liefer-Servicegraden	
Anzahl der Dispositionsstufen	Einstufige Modelle		Zweistufige Modelle	Mehrstufige Modelle	
Anzahl der Produkte	Einprodukt-Modelle			Mehrprodukt-Modelle	
Kapazitätsrestriktionen	keine Ressourcen- und Kapazitätsrestriktionen			Ressourcen- und Kapazitätsrestriktionen	
Konstanz der Parameter	konstante Parameter			variable Parameter	
Umfang des Planungszeitraums	Einperiodiger, endlicher Planungszeitraum		Mehrperiodiger, endlicher Planungszeitraum	Unendlicher Planungszeitraum	

Nach ASSFALG [1976, S. 29 f.] erlaubt die Berücksichtigung zusätzlicher Determinanten eine fortschreitende Annäherung an die Lagerhaltungswirklichkeit unter sukzessiver Aufgabe realitätsfremder Annahmen. Ausgehend von einem stationär-deterministischen Modell ohne Berücksichtigung von Nebenbedingungen können so schrittweise die Wahrscheinlichkeit der einzelnen Elemente, die Dynamik des Bestellmengenproblems und seiner Determinanten sowie die Besonderheiten der Nebenbedingungen in die Modellformulierung einbezogen werden. Im Anschluss kann ein solches einstufiges Einprodukt-Modell auf mehrere Produkte und mehrerer Lagerstufen ausgeweitet werden. Die vorangehenden Ausführungen und die Systematik in Tabelle 2.2 verdeutlichen die Vielfalt möglicher Modelle und die erhebliche

¹ Neben dem Bedarf kann auch die Unsicherheit und Stationarität der anderen Eingangs- und Ausgangsströme der Waren in der Modellbildung berücksichtigt werden (vgl. Abschnitt 2.1.3). Die Berücksichtigung der Fehlmengen im Vormerk- oder Verlustfall entspricht der Unsicherheit der unbefriedigten Nachfrage durch stochastische Warenausgangsmengen und -zeiten. Unsicherheiten bei der Menge und dem Zeitpunkt von Wareneinzügen werden durch stochastische Lieferzeiten und Liefermengen abgebildet. Fehlmengen des Absatzes sowie der Liefermengen können sowohl durch die verursachten Kosten als auch durch Servicegrade berücksichtigt werden.

² Eigene Darstellung in Anlehnung an [Toporowski 1996, S. 113].

Modellkomplexität bei einer Berücksichtigung aller Annahmen und Zusatzrestriktionen selbst für ein einzelnes Produkt und Lager. Aus diesem Grund sind der Wirklichkeitsnähe von Modellen der Bestellmengenrechnung Grenzen gesetzt, die zu einer Weiterentwicklung der Modelle unter Betrachtung ausschließlich ausgewählter Determinanten geführt haben.

ASSFALG [Assfalg 1976, S. 59] bezeichnet die Theorie der Bestellmengenmodelle eines Produktes mit deterministischem Bedarf als abgeschlossen. Seit Mitte der sechziger Jahre befassten sich Theoretiker zunächst verstärkt mit deterministischen Modellen mehrerer Produkte in einem Lager [Assfalg 1976] sowie mit Modellen mit einem Produkt in mehreren verbundenen Lagerstufen unter Berücksichtigung von Kapazitätsrestriktionen [Silver 1981]. Seitdem gewinnt die Analyse stochastischer Systeme an Bedeutung. Die Untersuchung stochastischer Modelle geht nach NADDOR [1971, S. 25] bis auf FRY [1928] zurück. Bereits die Monographie von WHITIN [1953] umfasst eine Bibliographie von 180 Literaturquellen, so dass der Umfang des Forschungsfeldes deutlich wird. Trotz der Menge an Veröffentlichungen existieren nur wenige aktuelle Literaturübersichten. LEE und NAHMIAS [1993] geben einen Überblick über stationär-stochastische, einstufige Einprodukt-Modelle, ebenso TOPOROWSKI [1998; 1999a; 1999b] unter Beschränkung auf stochastisch-stationäre Modelle im Handel. Eine analytische Herleitung und Diskussion dieser grundlegenden Modelle liefern [Naddor 1971; Silver und Peterson 1985; Hochstädter 1969; Klemm und Mikut 1972]. MUCKSTADT und ROUNDY [1993] betrachten die Entwicklung von Bestellmengenmodellen in mehrstufigen Distributions- und Produktionssystemen bei stationär-deterministischem Bedarf. FEDERGRUEN [1993] fasst die Entwicklungen mehrstufiger Systeme mit stochastischem Bedarf und zentraler Koordination zusammen. AXSÄTER [1993] betrachtet ferner mehrstufige Systeme mit stochastischem Bedarf bei dezentraler Kontrolle (ebenso Lee und Billington {, 1993, ISI:A1993MD43400002}).¹ Die Untersuchungen stochastischer Modelle beschränken sich vielfach auf die Modellierung stationär-stochastischer Bedarfe unter weitgehender Vernachlässigung instationärer Bedarfsverläufe aufgrund der besonderen Komplexität ihrer Lösung. In Anbetracht des Untersuchungsgegenstandes der Bedarfsprognose dieser Untersuchung ist die nachfolgende Analyse auf die Grundlagen einstufiger Einprodukt-Modelle der ein- und mehrperiodigen Bestellmengenrechnung beschränkt.²

Nachfolgend werden zunächst die Modellklassen der Bestellmengenrechnung für stationären und instationären deterministischen sowie stochastischen Bedarf kurz dargestellt, um den Untersuchungsgegenstand einordnen zu können. Innerhalb dieser Klassen werden spezifische Modelle unter Variation der Determinanten Bestellzeitpunkt, Bestellmenge, Lieferzeit, Behandlung von Fehlmengen und Lagerhaltungskosten diskutiert [Toporowski 1998, S. 1146; Naddor 1971]³, die in der Filialdisposition des Handels durch den Disponenten beeinflussbar sind [Müller-Hagedorn 1998, S. 526].

¹ Aufgrund der erheblichen Forschungsanstrengungen auf diesen Gebieten seit ASSFALG [1976] kann die Feststellung von DOMSCHKE et al. [1997, S. 158], dass Weiterentwicklungen von Modellen für stationär-stochastischen Bedarf sich überwiegend auf die Bestellmengenrechnung eines Produktes in einem Lager beschränkt hätten, nicht bestätigt werden.

² Insbesondere die zentrale bzw. die dezentrale Bestellmengenrechnung in mehrstufigen Lagerhaltungssystemen erfordert erhebliche Modellerweiterungen [Assfalg 1976, S. 29; Toporowski 1996]. Realitätsnahe Mehrprodukt-Modelle von mehreren simultan zu disponierenden Waren können häufig durch die Zerlegung in eine entsprechende Anzahl von Einprodukt-Modellen ohne Verlust der Optimalität gelöst werden [Hillier und Lieberman 1997, S. 648]. Da diese praxisrelevanten Erweiterungen hier nicht zum Verständnis der Interaktion zwischen Prognoseverfahren für instationären Bedarf und Modellen der Bestellmengenrechnung beitragen, wird auf ihre Darstellung verzichtet.

³ Dabei ist zu beachten, dass im operativen Kontext der Filialdisposition im Handel die Problemstellungen der stochastischen Bedarfe und Ausgangsströme besondere Bedeutung besitzen, während Bestelltermine, Wiederbeschaffungszeiten und -mengen vielfach extern vorgegebene Parameter sind. Dies schränkt die Auswahl möglicher Modelle der Bestellmengenrechnung im Handel sinnvoll ein.

2.3.2 Deterministische Modelle

2.3.2.1 Modelle für stationär-deterministischen Bedarf

Deterministische Bestellmengenmodelle gehen von einem bekannten zukünftigen Bedarf aus und bestimmen optimale Bestellmengen und Bestellzeitpunkte unter Kostenminimierung. Je nach der Bedarfsentwicklung im Zeitablauf wird zwischen deterministischen Modellen mit stationärem und instationärem Bedarfsverlauf unterschieden, wovon hier zunächst Modelle mit stationärem Bedarf untersucht werden.¹

Das Modell der klassischen Losgrößenformel nach HARRIS [1915, S. 48-52] ermittelt die optimale Bestellmenge q [ME] für deterministisch bekannten, stationären Bedarf unter Minimierung der mittleren Gesamtkosten je Zeiteinheit.² Das Modell geht von den Prämissen aus, dass (1) der Lagerbestand einer Ware bei deterministischem, konstantem Bedarf y [ME/ZE] kontinuierlich überwacht wird, dass (2) die Wiederbeschaffungszeit Null ist, $\Delta t_W = 0$, dass (3) keine Fehlmengen entstehen und dass (4) keine Restriktionen hinsichtlich Lagerkapazität und Höhe des eingesetzten Kapitals bestehen. Bei kontinuierlicher Überprüfung der Lagerbestände wird genau zu dem Zeitpunkt eine Bestellung der Höhe q ausgelöst, in dem der Lagerbestand auf Null abgesunken ist, $S = 0$. Die Ware trifft ohne Lieferzeit unmittelbar ein; da keine Fehlmengen auftreten können, entspricht der Lagerbestand S [ME] unmittelbar nach dem Wareneingang der Bestellmenge q , $S = q$.

Die Lagerungskosten C_L [GE] entstehen proportional zum durchschnittlichen Lagerbestand $(q-0) / 2$ [ME] und zur Länge der Lagerzeit von eines Bestellzyklus τ [ZE], bewertet mit den Lagerungskosten c_L [GE/ME*ZE] aus (2.7):

$$C_L(q) = \tau \frac{q}{2} c_L \quad . \quad (2.14)$$

Die Gesamtkosten $C(q)$ [GE] ergeben sich als Summe der Lagerungskosten (2.14) und der Bestellkosten einschließlich der Warenkosten (2.8) in Abhängigkeit von der Bestellmenge q :³

$$\begin{aligned} C(q) &= C_L(q) + C_B(q) \\ &= \tau \frac{q}{2} c_L + K_B + c_B q \end{aligned} \quad . \quad (2.15)$$

Zur Ermittlung der optimalen Bestellmenge werden die mittleren Kosten $C(q)$ pro Zeiteinheit [ZE] ermittelt:

$$C(q) = \frac{1}{\tau} \left(\tau \frac{q}{2} c_L + K_B + c_B q \right) \quad . \quad (2.16)$$

Da die Menge der Wareneingänge der Bestellung q [ME] ohne Fehlmengen gleich der Bedarfsmenge der Warenausgänge y [ME/ZE] eines Bestellzyklus der Länge τ [ZE] ist, gilt

$$q = y \tau \quad . \quad (2.17)$$

¹ Da für deterministische Modelle der stationäre Bedarf im Zeitablauf unverändert bleibt, erfolgt keine explizite Unterscheidung in einperiodige oder mehrperiodige Modelle, sondern es wird von einem unendlichen Planungshorizont ausgegangen.

² Das Modell ist auch als Bestellmengenformel (engl. *economic order quantity*, EOQ) oder als HARRIS-Modell bekannt [Toporowski 1996, S. 114; Müller-Hagedorn 1998, S. 527]. Ebenso wird es als WILSON-Modell [Wilson 1934] oder als ANDLER-Modell [Anderl 1929] bezeichnet.

³ Aufgrund des deterministischen Modells werden keine stochastischen Fehlmengenkosten berücksichtigt.

Der deterministische Bedarf y tritt als konstante und kontinuierliche Rate auf und wird ohne zeitliche Verzögerungen unmittelbar aus dem Lagerbestand befriedigt. Daraus folgt, dass die Länge eines Bestellzyklus τ durch den Quotienten aus Bestellmenge q und Bedarf y gegeben ist, mit

$$\tau = \frac{q}{y} \quad (2.18)$$

Durch Einsetzen von (2.18) in (2.16) werden die Gesamtkosten pro Zeiteinheit berechnet als

$$C(q) = \frac{y}{q} \left(\frac{q^2}{2} c_L + K_B + c_B q \right) = \frac{q}{2} c_L + \frac{y}{q} K_B + y c_B \quad (2.19)$$

Somit ergeben sich die mittleren Bestellkosten pro Zeiteinheit $C_B(q)$ als

$$C_B(q) = \frac{y}{q} K_B \quad (2.20)$$

wobei der Quotient y/q aus Bedarf und Bestellmenge die Anzahl der Bestellungen pro Zeiteinheit in Abhängigkeit von der Bestellmenge angibt. Die mittleren Lagerungskosten $C_L(q)$ pro Zeiteinheit werden durch den mittleren Lagerbestand $q/2$ in Abhängigkeit von der Bestellmenge bestimmt und lauten

$$C_L(q) = \frac{q}{2} c_L \quad (2.21)$$

Die Auswirkung der Bestellmenge q auf die einzelnen Kostengrößen und die Gesamtkosten pro Zeiteinheit wird grafisch in Abbildung 2.9 veranschaulicht:

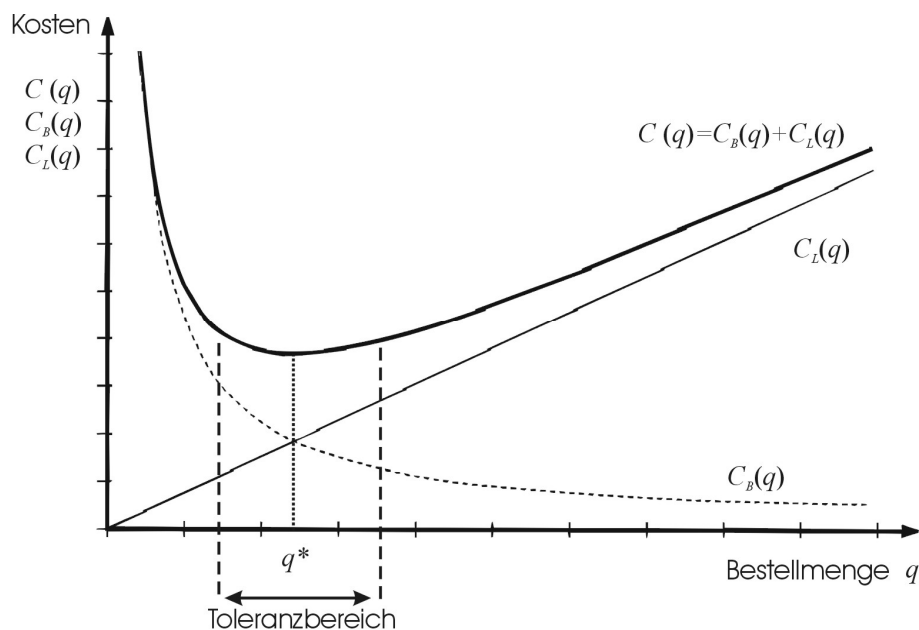


Abbildung 2.9: Bestellkosten, Lagerungskosten und Gesamtkosten in Abhängigkeit von der Bestellmenge¹

¹ Aus (2.19) wird deutlich, dass die variablen Bestellkosten c_B der Waren nicht von der Höhe der Bestellmenge abhängen und somit für die Bestimmung optimaler Bestellmengen nicht entscheidungsrelevant sind [Assfalg 1976; Bichler und Lörsch 1985].

Die Lagerungskosten steigen durch das gebundene Kapital linear mit der Bestellmenge, während die Bestellkosten durch die sinkende Anzahl an Bestellungen mit der Höhe der Bestellmenge hyperbolisch sinken. Das Minimum der Gesamtkosten bestimmt die optimale Bestellmenge q^* , die analytisch durch Differenzierung der Zielfunktion $C(q)$ in (2.19) nach der Bestellmenge q ermittelt wird:

$$\frac{dC(q)}{dq} = \frac{1}{2}c_L - \frac{y}{q^2}K_B = 0 \quad (2.22)$$

Nach (2.22) nehmen die Gesamtkosten für jene Bestellmenge q^* ihr Minimum an, wo die Bestell- und Lagerungskosten gleich sind, mit

$$q^* = \sqrt{\frac{2yK_B}{c_L}} \quad (2.23)$$

Aufgrund konstanter Nachfrage und Absatzrate im Planungszeitraum determiniert die statische, optimale Bestellmenge unmittelbar die optimale Länge der Bestellzyklen τ^* durch

$$\tau^* = \frac{q^*}{y} = \sqrt{\frac{2K_B}{yc_L}} \quad (2.24)$$

und durch den Bestellzyklus die optimale zeitliche Verteilung der Bestelltermine t_i über die Planungsperiode sowie die optimale Anzahl der Bestellungen im Planungszeitraum [Götz 1999, S. 52].² Da die Bestellzyklen konstant sind, bestimmen q^* sowie τ^* auch den stationären Bestellrhythmus r [ZE] [Toporowski 1996, S. 117; Schulte 1991].

Die optimale Bestellmenge wird unter restriktiven Prämissen hergeleitet, die in dieser Strenge selbst für Bedarfsituationen eines deterministischen Bedarfs nur in wenigen Ausnahmefällen erfüllt sind [Müller-Hagedorn 1998, S. 130; Silver und Peterson 1985; Bichler und Lörsh 1985].³ Die Entscheidungsvariablen ergeben sich unter der Prämisse, dass jede Lieferung aufgrund bekannter und konstanter Wiederbeschaffungszeiten in exakter Höhe präzise zu dem Zeitpunkt eintrifft, wenn der Bestand auf Null abgesunken ist. Ferner gelten die Kostenparameter Lagerhaltungskostensätze, Bestellkostensätze und Stückpreise in der Planungsperiode als bekannt, konstant und unabhängig von anderen Determinanten. Zudem existieren keine Mindestliefermengen, Verbundeffekte zwischen Bestellungen von Artikeln oder Kapital- und Lagerraumrestriktionen. Trotz der eingeschränkten Realitätsnähe des klassischen Losgrößenmodells [Günther 1991] können die grundlegenden Determinanten und Interdependenzen der Bestellmengenrechnung an ihm dargestellt und anschließend unter sukzessiver Aufhebung einzelner Prämissen verfeinerte Modelle höherer Komplexität und Realitätsnähe daraus abgeleitet werden [Müller-Hagedorn 1998, S. 527].⁴

¹ Erweiterte Darstellung nach [Toporowski 1996, S. 117; Bichler und Lörsh 1985].

² Entsprechend sinkt die Anzahl der Bestellungen zur Deckung des deterministischen Bedarfs einer Planungsperiode mit wachsender Bestellmenge [Arnold et al. 2004, S. A3-65].

³ Allerdings lassen sich aus dem Zusammenhang des einfachen Losgrößenmodells bereits die relevanten Kernaussagen der Interdependenz zwischen Bedarf, Bestellkosten, Lagerkosten, Bestellmenge und Anzahl der Bestellungen ableiten. Einen Überblick gibt [Arnold et al. 2004].

⁴ Dabei ist anzumerken, dass das klassische Losgrößenmodell relativ robust gegen Abweichungen der Entscheidungsvariablen der Bestellmenge und der Bestellzyklen sowie der Parameter der Kosten ist, die einen unterproportionalen Kostenanstieg bewirken [Waters 2003, S. 48-64; Domschke et al. 1997, S. 73; Naddor 1971, S. 57 f.]. Dies gilt auch für die Abweichungen des Bedarfs durch stochastische, stationäre sowie instationäre Bedarfsverläufe und erklärt die häufige Anwendung der klassischen Losgrößenformel für deterministischen Bedarf in der Praxis, trotz offensichtlicher Fehlspezifikation des Bestellmengenmodells.

Unter Aufhebung der Prämissen wird das Modell zur Berücksichtigung weiterer Parameter ohne Änderung der grundlegenden Modellstruktur erweitert [Müller-Hagedorn 1998, S. 534]. Das Grundmodell der Losgrößenformel kann zur Berücksichtigung deterministischer Wiederbeschaffungs- bzw. Lieferzeiten Δt_L [ZE] von endlicher Dauer erweitert werden. Dabei bleiben die optimale Bestellmenge q^* sowie der Bestellzyklus τ^* unverändert. Es wird lediglich der Bestellzeitpunkt t_i [-] um Δt_L [ZE] vor den Lieferzeitpunkt vorgezogen, an dem das Lager zur Vermeidung von Fehlmengen aufgefüllt werden muss. Eine Bestellung wird ausgelöst, wenn der Lagerbestand S [ME] auf den kritischen Bestellbestand oder Bestellpunkt s [ME] abgesunken ist, mit

$$s = \Delta t_L \cdot y \quad , \quad (2.25)$$

der gerade ausreicht, um den kontinuierlichen Bedarf in der Wiederbeschaffungszeit zu decken [Arnold et al. 2004, S. A3-65]. Bei Wareneingang wird das Lager auf den optimalen Lagerbestand S^* aufgefüllt, welcher ohne Auftreten von Fehlmengen wie zuvor der Bestellmenge q^* entspricht, $q^* = S^*$. Der typische Verlauf des Lagerbestands deterministischer Modelle sowie seine Entscheidungsparameter sind in Abbildung 2.10 grafisch dargestellt.

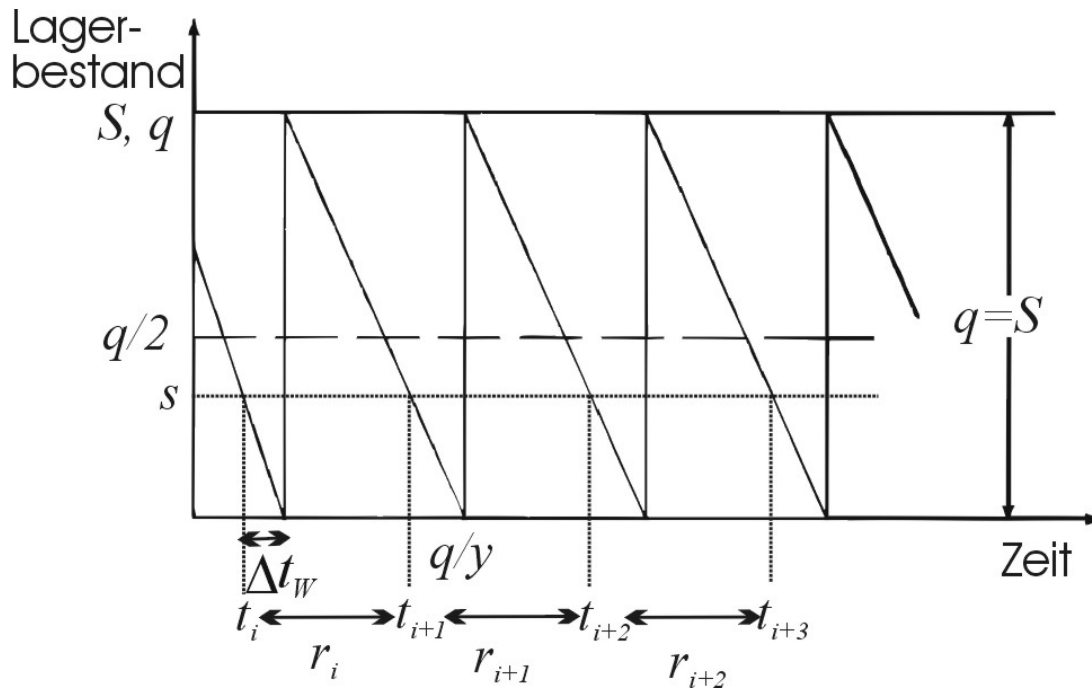


Abbildung 2.10: Lagerbestandsentwicklung im deterministischen Modell nach HARRIS¹

Darüber hinaus existieren vielfältige handelsrelevante² Erweiterungen der klassischen Losgrößenformel, insbesondere die Berücksichtigung verringerter Preis- und Transportkosten durch Rabattstufen oder Rabattgrenzen bei höheren Bestellmengen [Müller-Hagedorn 1998, S. 529-534; Toporowski 1996, S. 84; Grochla 1992] sowie durch Lieferantenkredite und Zahlungsziele [Silver und Peterson 1985, S. 534-536; Müller-Hagedorn 1998]. BICHLER und LÖRSCH [1985, S. 78 f.] betrachten Preisänderungen der Beschaffungsmärkte im Zeitablauf in Form von Preiserhöhungen und zeitlich begrenzten Sonderangeboten [ebenso Grochla 1992, S. 85 ff.; Waters 2003, S. 112 f.]. Darüber hinaus werden die Modelle zur Berücksichtigung von zeitabhängigen Fehlmengenkosten im Vormerkfall [Assfalg 1976, S. 34 ff.; Hillier und

¹ Erweiterte Darstellung nach TOPOROWSKI [1996, S. 115-116].

² Darüber hinaus existiert eine Vielzahl relevanter Erweiterungen für Produktionsbetriebe. DOMSCHKE et al. [1997, S. 74 f.] sowie NADDOR [1971, S. 60 ff.] zeigen Erweiterungen des Modells um endliche Fertigungsgeschwindigkeit bei kontinuierlichem Lagerzugang und Verzugs Mengen.

Lieberman 1997, S. 608 ff.; Waters 2003, S. 104 ff.] sowie Fehlmengen im Verlustfall erweitert [Assfalg 1976, S. 37 f.; Waters 2003, S. 108 ff.]. WATERS [2003, S. 90 ff.] untersucht steigende Bestellkosten mit sprungfixem Kostenverlauf. Zudem können die Modelle zur Berücksichtigung von Verbundeffekten zwischen den Waren in der Beschaffung sowie der tatsächlichen Zahlungsströme und der Kapitalbindung erweitert werden [Waters 2003, S. 115-125; Assfalg 1976, S. 38 ff.]. DOMSCHKE et al. [1997, S. 78-97] untersuchen einstufige Mehrproduktmodelle mit Sammelbestellungen und Kapazitätsbeschränkungen sowie Erweiterungen auf mehrstufige Mehrproduktmodelle [Domschke et al. 1997, S. 98-109]. Diese Erweiterungen führen zu einer gesteigerten Realitätsnähe für Probleme mit stationärem, deterministischem Bedarf. Für eine umfassende Diskussion deterministischer Modelle und weiterführende Literaturempfehlungen für den Handel siehe [Toporowski 1996], im Hinblick auf Materialwirtschaft und Produktionsplanung siehe [Tempelmeier 2003; Silver und Peterson 1985; Schneeweiß 1981; Naddor 1971].

2.3.2.2 Modelle für instationär-deterministischen Bedarf

Ein deterministischer Bedarf mit bekanntem, aber instationärem Nachfrageverlauf durch schwankenden und veränderlichen Absatz über mehrere Planungsperioden erfordert alternative Bestellmengenmodelle für die Bestimmung optimaler Bestellmengen.

Für instationär-deterministischen Bedarf wurden spezifische Algorithmen zur Lösung dynamischer Entscheidungsprobleme der Bestellmengenrechnung (engl. *dynamic lot sizing problems*) entwickelt. RESH et al. [1976] erweitern die klassische Losgrößenformel zur Berücksichtigung eines Bedarfs mit konstantem, linearem Trend, was jedoch die Bestimmung der optimalen Bestellmengenpolitik erheblich verkompliziert [Lee und Nahmias 1993, S. 17]. NADDOR untersucht wachsenden Bedarf mit deterministischen Trends für unterschiedliche Bestellmengenpolitiken [Naddor 1971, S. 113-121] sowie ein Modell mit veränderlichem Bedarf zum Periodenbeginn [Naddor 1971, S. 121-124].

Unter den dynamischen Modellen kommt jedoch den Varianten des WAGNER-WITHIN-Algorithmus [Wagner und Whitin 1958] eine besondere Bedeutung zu [Silver und Peterson 1985; Tempelmeier 2003; Müller-Hagedorn 1998, S. 538 ff.]. Das Grundmodell ermittelt optimale Bestellmengen in einem einstufigen Einprodukt-Modell mit endlichem Planungshorizont unter der Annahme ähnlicher Restriktionen wie die klassische Losgrößenformel: keine Wiederbeschaffungszeiten, keine Kapazitätsbeschränkungen, keine Fehlmengen und zeitinvariante variable Bestellmengenkosten, die somit ebenfalls entscheidungsirrelevant sind [Domschke et al. 1997, S. 110]. Das Modell lässt dagegen zeitlich veränderliche Bedarfsmengen y_t [ME/ZE] am Periodenbeginn und zeitlich veränderliche fixe Bestellkosten und lineare Lagerungskosten zu. Ziel des Modells ist die Ermittlung der im Zeitablauf veränderlichen Bestellmengen, die zu minimalen Gesamtkosten führen [Domschke et al. 1997, S. 110]. Das Modell ist mit Hilfe der Dynamischen Optimierung exakt lösbar und berücksichtigt unterschiedliche Kombinationen von Bestellauslösung und Bestellmengen mittels einer iterativen Rückwärtsrekursion [siehe u.a. Domschke et al. 1997, S. 115-121].¹ Aufgrund der hohen Komplexität und der langen Rechenzeiten des Algorithmus, selbst bei einer begrenzten Anzahl von Teilperioden, wurden unterschiedliche heuristische Ansätze entwickelt. Eine umfassende Darstellung der Modelle der Part-Period-Verfahren, des Verfahrens von GROFF, der SILVER-MEAL-Heuristik sowie der Incremental-Order-Quantity-Regel und der Least-Unit-Cost-Regel findet sich in [Tempelmeier 2003; Schneeweiß 1981; Götz 1999, S. 75; Lee und Nahmias 1993, S. 21 ff.]; eine Analyse mit Simulationsergebnissen geben [Robrade und Zoller 1987; Robrade 1991].

¹ DOMSCHKE et al. [1997, S. 112 ff.] stellen darüber hinaus exakte Lösungsansätze durch eine Formulierung des Modells als Kürzeste-Wege-Problem oder als unkapazitiertes Warehouse-Location-Problem vor.

DOMSCHKE et al. [1997, S. 123 ff.] untersuchen Erweiterungen des einstufigen Einprodukt-Grundmodells für instationär-deterministischen Bedarf durch rollierende Planungshorizonte, variable Kostenparameter, Rabatte auf Einstandspreise und Kapazitätsrestriktionen. Darüber hinaus erweitern sie ihre Betrachtung auf einstufige Mehrproduktmodelle (engl. *capacitated lot-sizing problems*) und mehrstufige Mehrproduktmodelle unter Modifikationen und Verallgemeinerungen [Domschke et al. 1997, S. 127-158].

Alle hier genannten Verfahren gehen jedoch von einem deterministischen Bedarf aus. Die Problemstellungen der Bestellmengenrechnung im Handel erfordern hingegen alternative Verfahren zur Abbildung der stochastischen Realität, die im nachfolgenden Abschnitt vorgestellt werden.

2.3.3 Stochastische Modelle

2.3.3.1 Modelle für stationär-stochastischen Bedarf

2.3.3.1.1 Einperiodige stationär-stochastische Modelle

In stochastischen Modellen wird die Annahme einer deterministischen Bedarfsstruktur mit bekannten Bedarfsmengen und -zeitpunkten zugunsten der realitätsnäheren Annahme einer stochastischen Bedarfsstruktur aufgegeben [Götz 1999, S. 75].

Die Literatur zu stationär-stochastischen Einprodukt-Bestellmengenproblemen umfasst Tausende von Publikationen [Lee und Nahmias 1993]. Nachfolgend muss daher auf eine repräsentative Darstellung verzichtet werden. Es werden lediglich die für die weitere Untersuchung relevanten Strukturdimensionen und Entscheidungsvariablen analysiert, um eine Einordnung empirischer Problemstellungen in die Modellklassen zu ermöglichen. Nachfolgend wird zunächst ein einperiodiges Einprodukt-Modell bei stationär-stochastischem Bedarf betrachtet, das auch als Zeitungsjungen-Problem (engl. *newsboy* oder *news vendor problem*) bekannt ist [Morse und Kimball 1951; Hadley und Whitin 1963]. Da es die Grundlage weiterer stochastischer Bestellmengenmodelle bildet, wird es hier eingehend vorgestellt.¹

Bei deterministischen Warenflüssen der Nachfrage, Wiederbeschaffungszeit und Liefermenge wird das Bestellverhalten durch eine optimale Bestellmenge hinreichend kontrolliert, so dass die Ware immer genau dann eintrifft, wenn der Bestand auf Null abgesunken ist [Toporowski 1996, S. 125-126]. Im Fall stochastischer Schwankungen eines mittelwert- und varianzstationären Bedarfs y führt eine rhythmische Bestellung identischer Mengen hingegen wechselweise zu ungewollten Fehlmengen durch Bedarfsunterdeckung bei überdurchschnittlich hohem Bedarf oder zu Überbeständen durch Bedarfsüberdeckungen bei unterdurchschnittlich niedrigem Bedarf, wie in Abbildung 2.11 (a.) dargestellt.

Ebenso führen stochastische, stationäre Wiederbeschaffungszeiten wie in Abbildung 2.11 (b.) zu ungewollten Fehlmengen und Überbeständen durch über- und unterdurchschnittliche Wiederbeschaffungszeiten. Das Fehlmengenrisiko aus den stochastischen Eingangs- und Ausgangsflüssen muss bei der Bestimmung der Bestellmengen berücksichtigt werden. Hierzu wird das Optimierungskalkül der Bestellmengenrechnung um Sicherheitsbestände erweitert, die Fehlmengensituationen und Lieferausfälle vermeiden [Arrow et al. 1958; Lee und Nahmias 1993] und so einen bestimmten Servicegrad erzielen [Domschke et al. 1997, S. 160].

¹ SILVER et al. [1998, S. 385 ff.] und WATERS [2003, S. 127-153] unternehmen eine marginalanalytische Herleitung des Newsboy-Modells. Da diese mit der Herleitung der Problemstellung asymmetrischer Kostenfunktionen in der Entscheidungsfindung nach GRANGER [1969] korrespondiert, wird auf ihre Darstellung in Kapitel 5 verwiesen und hier lediglich eine einheitliche Herleitung unter Minimierung der Zielfunktion unternommen.

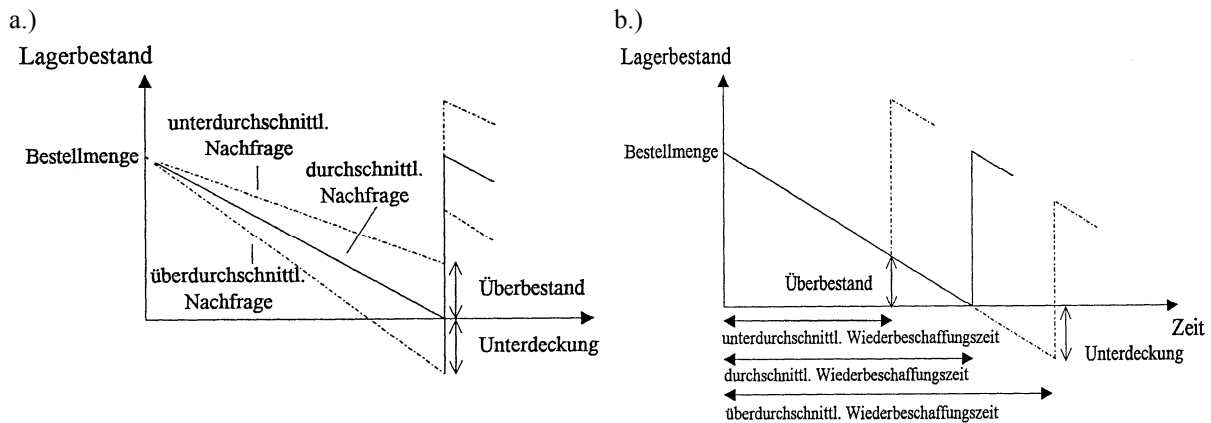


Abbildung 2.11: Lagerbestandsentwicklung bei stochastischem Bedarf (a.)
und bei stochastischer Wiederbeschaffungszeit (b.)¹

Das Newsboy-Modell bestimmt die Entscheidungsvariable der optimalen Bestellmenge q [ME] für einen einzelnen Bestellzyklus τ [ZE] mit $\tau = 1$ unter Minimierung der bekannten Bestellkosten, Lagerungskosten und Fehlmengenkosten (siehe 2.3.1.2) bei stationärstochastischem Bedarf. Der Bedarf (vgl. Abschnitt 2.2.3.1) wird als stationäre, diskrete Zufallsvariable Y betrachtet, die beliebige Werte y [ME] einer bekannten, vollständig spezifizierten Wahrscheinlichkeitsverteilung annehmen kann und zum Ende der Periode anfällt. Die Wahrscheinlichkeitsfunktion $P_Y(y)$ (siehe auch Abschnitt 3.2.21) ordnet dem Auftreten eines bestimmten, diskreten Bedarfs y eine Wahrscheinlichkeit p_i zu:

$$P_Y(y) = P\{Y = y\} \quad . \quad (2.26)$$

Aus der Wahrscheinlichkeitsfunktion $P_Y(y)$ wird die kumulative Verteilungsfunktion der Zufallsvariablen des Bedarfs $F_Y(b)$ gebildet, mit

$$F_Y(b) = \sum_{y=0}^b P_Y(y) \quad . \quad (2.27)$$

Ferner wird von einer Reihe vereinfachender Prämissen der Modellparameter ausgegangen, die später gelockert werden können.² Die Waren werden am Anfang der einzelnen Periode τ [ZE] bestellt, wobei variable Stückkosten c_B [GE/ME] proportional zur Bestellmenge q [ME] unter Vernachlässigung fixer Bestellkosten anfallen (2.8). Die variablen Lagerungskosten c_L [GE/ME] entstehen je Mengeneinheit des übrig gebliebenen Lagerbestandes am Ende der Periode (2.7); sie werden als Funktion des Überschusses über die geforderte Menge aufgefasst [Hillier und Lieberman 1997, S. 624]. Die Fehlmengenkosten im Verlustfall³ $c_{F,L}$ [GE/ME] bezeichnen die Opportunitätskosten der entgangenen Einnahmen proportional zum unbefriedigten Bedarf (2.10) unter der Nebenbedingung $c_{F,L} > c_B$.⁴ Es existieren weder ein Anfangsbestand des Lagers noch eine Wiederbeschaffungszeit. Somit bezeichnet q jenen

¹ Erweiterte Darstellung nach [Arnold et al. 2004, S. A3-66].

² In Anbetracht einer Vielzahl divergenter Notationen folgt die Darstellung der traditionellen Herleitung nach ARROW et al. [1958] unter Kostenminimierung, jedoch unter Anpassung an eine einheitliche Notation. Die Formulierung unter Kostenminimierung ist verbreitet [Lee und Nahmias 1993; Hillier und Lieberman 1997; Hochstädter 1969]; DOMSCHKE et al. [1997, S. 162 ff.] sowie SILVER et al. [1998] formulieren das Zeitungs-jungen-Problem dagegen unter dem Aspekt der Gewinnmaximierung, was zugunsten einer einheitlichen, kostenorientierten Betrachtung in Kapitel 5 und 6 hier unterbleibt.

³ Es wird vom Verlustfall ausgegangen, da im einperiodigen Modell keine offenen Bestellmengen durch Nachlieferungen in Folgeperioden berücksichtigt werden können.

⁴ Gemäß Nebenbedingung $c_{F,L} > c_B$ sind die Fehlmengenkosten höher als die Bestellmengkosten. Somit erscheint bei zufälligen, mittelwertstationären Bedarfsschwankungen eine Bestellmenge sinnvoll, die höher als der im Mittel erwartete Bedarf ist, aber geringer als der maximale Bedarf der Wahrscheinlichkeitsdichte.

Lagerbestand, der unmittelbar nach einer durchgeführten Bestellung und einem Lagereingang zur vollständigen oder anteiligen Deckung des Bedarfs y [ME] der Periode zur Verfügung steht [Hillier und Lieberman 1997, S. 625].¹

Die verkaufte Menge der Periode wird durch den niedrigeren Wert aus der bestellten Menge q und dem stochastischen Bedarf y angegeben, $\min\{y, q\}$, mit

$$\begin{cases} y & , \text{ wenn } y < q \\ q & , \text{ wenn } y \geq q \end{cases} = \min\{y, q\} \quad . \quad (2.28)$$

Ist die Bestellmenge höher als der Bedarf, $y < q$, so wird der gesamte Bedarf in Höhe von y befriedigt. Am Ende der Periode verbleibt so ein Restbestand an Waren W [ME] in Höhe von

$$W = q - y \quad . \quad (2.29)$$

Ist hingegen der Bedarf höher als die Bestellmenge, $y \geq q$, so wird nur der Bedarf in Höhe des vorhandenen Bestands durch die Bestellmenge q befriedigt, und es entstehen Fehlmengen q_F [ME] in Höhe von

$$q_F = y - q \quad . \quad (2.30)$$

Aufgrund des Zielkonfliktes der antinomischen Kostenverläufe ist die Bestellmenge q so zu bestimmen, dass ein Kompromiss zwischen (a.) dem Risiko von Fehlmengenkosten durch einen zu geringen Lagerbestand, und (b.) dem Risiko von überflüssigen Kosten der Bestellung und Lagerhaltung durch einen zu hohen Lagerbestand erreicht wird [Hillier und Lieberman 1997, S. 625]. Tritt bei einer Bestellmenge bzw. einem Lagerbestand q ein Bedarf von y auf, so fallen entweder Lagerungskosten c_L oder Fehlmengenkosten $c_{F,L}$ an, in Abhängigkeit davon, ob die Bestellmenge größer ist als der Bedarf, $y < q$ und somit $\max\{0, q - y\} > 0$, oder ob der Bedarf größer ist als die Bestellmenge, $y \geq q$ und somit $\max\{0, y - q\} > 0$, mit:

$$C(y, q) = c_B q + c_L \max\{0, q - y\} + c_{F,L} \max\{0, y - q\} \quad . \quad (2.31)$$

Da der Bedarf y eine Zufallsvariable mit der Wahrscheinlichkeitsfunktion $P_Y(y)$ ist, sind auch die anfallenden Gesamtkosten $C(q)$ in Abhängigkeit von der Höhe der Bestellmenge eine Zufallsvariable. Die Bestellmenge q ist daher so zu wählen, dass der Erwartungswert der Summe aus Bestell-, Lagerhaltungs- und Fehlmengenkosten minimiert wird [Arnold et al. 2004, S. A3-66]

$$\begin{aligned} C(q) &= E[C(y, q)] \\ &= \sum_{y=0}^{\infty} [c_B q + c_L \max\{0, q - y\} + c_{F,L} \max\{0, y - q\}] P_Y(y) \quad . \\ &= c_B q + \sum_{y=0}^{q-1} c_L (q - y) P_Y(y) + \sum_{y=q}^{\infty} c_{F,L} (y - q) P_Y(y) \end{aligned} \quad (2.32)$$

Gesucht wird die optimale Bestellmenge q^* die $C(q)$ minimiert. Leitet man (2.24) nach q ab,

¹ Durch die einzelne Zeitperiode und unter Vernachlässigung stochastischer Wiederbeschaffungszeiten entfällt das Problem der Bestimmung optimaler Bestellzeitpunkte, und die Modellbildung vereinfacht sich zur Bestimmung einer optimalen Bestellmenge q^* zur Bedarfsbefriedigung [Assfalg 1976, S. 47].

$$\frac{dC(q)}{dq} = c_B q + \sum_{y=0}^{q-1} c_L (q-y) P_Y(y) + \sum_{y=q}^{\infty} c_{F,L} (y-S) P_Y(y) = 0 \quad , \quad (2.33)$$

so folgt für die diskrete Zufallsvariable Y mit der kumulierten Verteilungsfunktion $F_Y(b)$ aus (2.20) die Vereinfachung:

$$\frac{dC(q)}{dq} = c_B + c_L F_Y(q^*) + c_{F,L} [1 - F_Y(q^*)] \quad . \quad (2.34)$$

Durch Auflösen ergibt dies die optimale Bestellmenge q^* als die kleinste ganze Zahl, für die gilt, dass die kumulierte Wahrscheinlichkeit $F_Y(q^*)$ dem Quotienten der Kosten entspricht:

$$F_Y(q^*) \geq \frac{c_{F,L} - c_B}{c_{F,L} + c_L} \quad . \quad (2.35)$$

Da $F_Y(q^*)$ die Wahrscheinlichkeit bezeichnet, dass der Bedarf y die Bestellmenge q^* nicht überschreitet, tritt die optimale Lösung auf, wenn die Wahrscheinlichkeit gleich dem kritischen Quotienten ist [Lee und Nahmias 1993, S. 27]. Der Quotient der Dispositionskosten determiniert somit die optimale Bestellmengenpolitik.² „Er liefert – mit einhundert multipliziert – den prozentualen Anteil der wahrscheinlichen Nachfrage, dessen Befriedigung für die vorgegebenen Werte von c_L , $c_{F,V}$ und c_B optimal ist“ [Hochstädter 1969, S. 36]. Der Disponent muss die Bestellmenge so bestimmen, dass die Wahrscheinlichkeit, dass der Bedarf y geringer ist als die Bestellmenge q , dem kritischen Quotienten in Prozent entspricht.

Aufgrund der zunehmenden analytischen Schwierigkeiten bei der Analyse und der Lösung von Gleichungen mit diskreten Zufallsvariablen (vgl. Abschnitt 2.2.1) werden diese im Zuge der stochastischen Bestellmengenrechnung häufig durch kontinuierliche Zufallsvariablen unter Annahme eines stetigen Bedarfs angenähert [Hillier und Lieberman 1997, S. 626].³ Einzige Veränderung des Modells ist die Verwendung der Dichtefunktion $f(y)dy$ und der kumulativen Verteilungsfunktion $F(y)$ (vgl. Abschnitt 3.2.2.1) für y in Gleichung (2.24):

$$\begin{aligned} C(q) &= E[C(y, q)] \\ &= \int_0^{\infty} [c_B q + c_L \max(0, q-y) + c_{V,L} \max\{0, y-q\}] f(y) dy \quad , \quad (2.36) \\ &= c_B q + \int_0^q h(q-y) f(y) dy + \int_q^{\infty} p(y-q) f(y) dy \end{aligned}$$

und einer unveränderten Bestimmung der optimalen Bestellmenge q^* für $C(q)$, für die gilt:

$$F_Y(q^*) = \frac{c_{F,V} - c_B}{c_{F,V} + c_L} \quad . \quad (2.37)$$

¹ HILLIER und LIEBERMANN [1997, S. 629 ff.] liefern eine vollständige Herleitung dieses Ergebnisses für das einperiodige Modell ohne Beschaffungskosten mit linearen Fehlmengen- und Lagerhaltungskosten.

² Der kritische Quotient ist durch $c_{F,L} > c_B$ immer kleiner Eins, da andernfalls eine Lagerung unökonomisch wäre [Hochstädter 1969, S. 36].

³ Eine Reihe von Autoren betrachtet ausschließlich Bestellmengenmodelle unter kontinuierlichem Bedarf, u. a. [Assfalg 1976; Arnold et al. 2004; Hochstädter 1969].

Die optimale Bestellmenge q^* spezifiziert ein Quantil der kumulierten Verteilungsfunktion, das in Abhängigkeit des Verhältnisses der Kostenparameter aus Fehlmengenkosten, Lagerhaltungskosten und Bestellmengenkosten bestimmt wird [Arnold et al. 2004]. Dieses Quantil kann bei vollständig spezifizierter Wahrscheinlichkeitsverteilung des Bedarfs unmittelbar aus den vergangenen Bedarfsdaten abgelesen werden. Eine grafische Lösung für eine Verteilungsfunktion des normalverteilten Bedarfs ist in Abbildung 2.12 dargestellt.

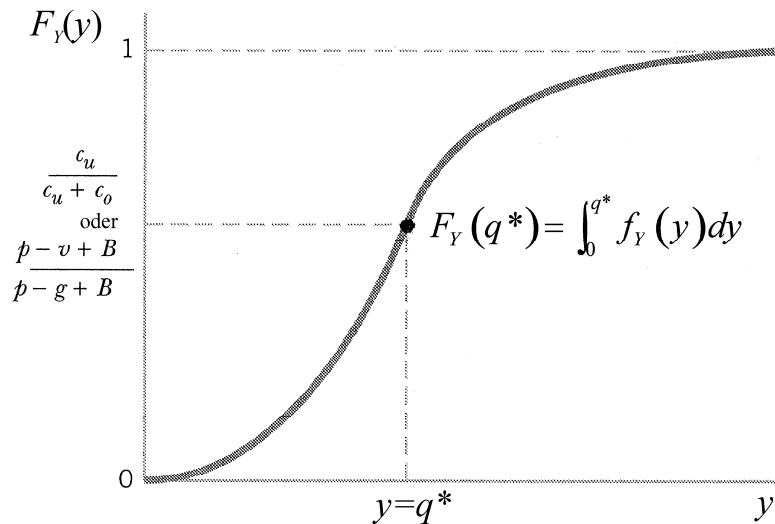


Abbildung 2.12: Bestimmung der optimalen Bestellmenge q^* im Newsboy-Modell¹

Die vorangehende Herleitung der optimalen Bestellmenge setzt die Kenntnis einer vollständig spezifizierten diskreten Wahrscheinlichkeitsfunktion bzw. stetigen Verteilungsfunktion voraus. In diesem Fall kann eine direkte Bestimmung der optimalen Bestellmenge für jede beliebige Wahrscheinlichkeitsverteilung erfolgen. Ist die tatsächliche Verteilungsfunktion jedoch unbekannt oder nur zum Teil spezifiziert, so kann $F_Y(q^*)$ in (2.35) bzw. (2.37) nur unter Annahme einer theoretischen Verteilungsfunktion und einer Schätzung ihrer ersten Momente zur Anpassung an den empirischen Bedarfsverlauf bestimmt werden (vgl. Abschnitt 3.2.2.2). In diesem Fall ist für eine spezifizierte Verteilungsfunktion der Sicherheitsfaktor k so aus den standardisierten Verteilungstabellen der jeweiligen Wahrscheinlichkeitsverteilung zu bestimmen, dass er dem durch die Kosten spezifizierten Quantil entspricht [Arnold et al. 2004, S. A3-66; Toporowski 1996, S. 128]:

$$F_Y(k^*) = 1 - \frac{c_{F,V} - c_B}{c_{F,V} + c_L} \quad (2.38)$$

Unter der Annahme eines normalverteilten Bedarfs wird eine optimale Bestellmenge q^* aus dem geschätzten Erwartungswert des Bedarfs \hat{y}_Y und einem k -fachen der empirischen Standardabweichung $\hat{\sigma}_Y$ bestimmt:³

$$q^* = \hat{y}_Y + k \hat{\sigma}_Y \quad (2.39)$$

¹ Eigene, erweiterte Darstellung in Anlehnung an SILVER et al. [1998, S. 386].

² Ein Sicherheitsfaktor k entspricht dem Wert $1-z$ der Tabellen der Standardnormalverteilung $N(0, 1)$. Ein Sicherheitsfaktor von $k=2,33$ entspricht so einem Quantil und bzw. Servicegrad von 99,0% [Makridakis et al. 1998; Silver et al. 1998].

³ Eine Normalverteilung wird vollständig durch ihre ersten beiden Momente, nämlich den Erwartungswert des empirischen Mittelwertes und die Standardabweichung der Verteilung, spezifiziert. Die Quantile anderer Verteilungsfunktionen werden jedoch durch andere Parameter spezifiziert (vgl. Abschnitt 3.2.2.2).

„Das Ausmaß $k \hat{\sigma}_Y$, in dem die Bestellmenge q über die erwartete Nachfrage \hat{y}_Y hinaus angehoben wird, ... bezeichnet den Sicherheitsbestand, der zur Absicherung gegen das Risiko zu hoher Fehlmengen eingeplant wird“ [Arnold et al. 2004, S. A3.66]. Der Zusammenhang ist in Abbildung 2.13 dargestellt.

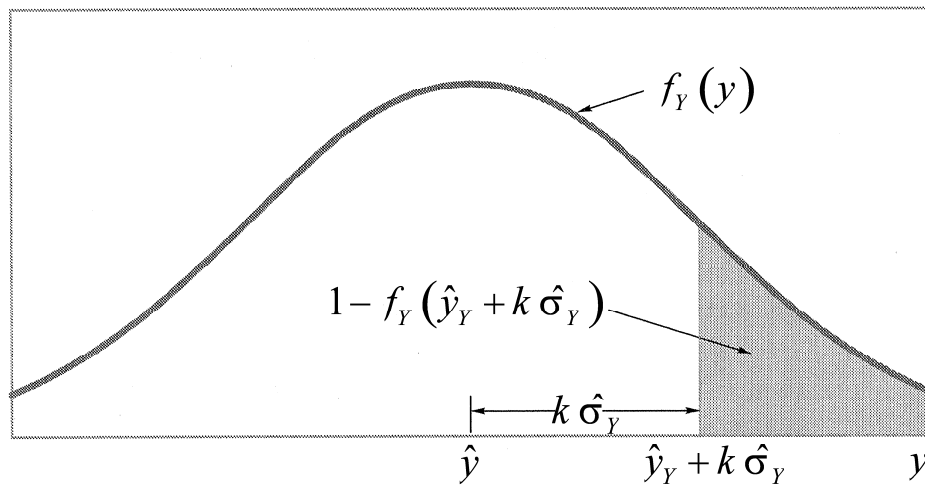


Abbildung 2.13: Bestimmung des Sicherheitsbestandes im Newsboy-Modell¹

Der Sicherheitsbestand $k \hat{\sigma}_Y$ (engl. *safety stock*) wird durch den Sicherheitsfaktor $k [-]$ (engl. *safety factor*) bestimmt, und zwar unter der Annahme, dass \hat{y}_Y und $\hat{\sigma}_Y$ erwartungstreue Schätzer ohne systematische positive oder negative Abweichungen sind. Die Fläche $1 - f_Y(\hat{y}_Y + k \hat{\sigma}_Y)$ unter der Dichtefunktion bezeichnet die Wahrscheinlichkeit, mit der der Bedarf Y die Bestellmenge übersteigt und somit Fehlmengen auftreten. Die korrekte Bestimmung des Sicherheitsfaktors und der Sicherheitsbestände erfordert eine korrekte Bestimmung der Verteilungsform. Für einen Bedarf mit Poisson-Verteilung $PV(\mu_Y)$, mit $\mu = \sigma^2$, verändert sich die Bestimmung des Sicherheitsbestandes entsprechend zu [Silver et al. 1998, S. 345]:

$$q^* = \hat{y}_Y + k \sqrt{\hat{y}_Y} \quad (2.40)$$

Für eine Diskussion der Auswahl einer geeigneten Verteilungsfunktion zur stochastischen Bedarfsprognose siehe Abschnitt 3.2.2.3 sowie [Silver et al. 1998, S. 272-274 und 284].

Der kritische Quotient der Dispositionskosten kann alternativ durch die Relation der Kosten einer Bedarfsunterdeckung (engl. *underage costs*) c_u [GE/ME] bzw. Bedarfsüberdeckung (engl. *overage costs*) c_o [GE/ME] bezeichnet werden [Lee und Nahmias 1993, S. 27], wobei folgende Relation herrscht:²

$$F_Y(q^*) = \frac{c_{F,L} - c_B}{c_{F,L} + c_L} = \frac{c_{F,L} - c_B}{(c_L + c_B) + (c_{F,L} - c_B)} = \frac{c_u}{c_o + c_u} \quad (2.41)$$

¹ Eigene, erweiterte Darstellung in Anlehnung an SILVER et al. [1998, S. 254].

² Dabei sind die Kosten einer Bedarfsunterdeckung c_u angenähert als entgangener Gewinn zu verstehen, der sich aus den Fehlmengenkosten $c_{F,V}$ aus dem Einstandspreis abzüglich unnötiger Bestellkosten c_B errechnet, mit $c_u = c_{F,V} - c_B$. Die Kosten einer Bedarfsüberdeckung entstehen durch die zusätzlichen Lagerungskosten c_L und die unnötigen Bestellkosten c_B , mit $c_o = c_L + c_B$ [Silver und Peterson 1985; Waters 2003, S. 142]. Die Notationen weichen zum Teil jedoch ab; NÄDDÖR [1971, S. 135] definiert c_u und c_o beispielsweise anders, da er in seiner Herleitung keine Bestellkosten, sondern lediglich Fehlmengen- und Lagerungskosten berücksichtigt.

Für identische Fehlmengenkosten der Unter- und Überbestände, $c_u = c_o$, gilt somit $F_Y(q^*) = 0,5$, was dem Mittelwert der Verteilung entspricht und somit $k = 0$.

Bei einer Formulierung des Newsboy-Modells unter dem Aspekt der Gewinnmaximierung [Silver et al. 1998, S. 387 ff.; Domschke et al. 1997, S. 163] resultiert ein abweichender Quotient der Dispositionskosten unter Erweiterung der Kosten bei Unterbestand durch einen Verkaufspreis p [GE/ME], mit $c_u = p - c_B + c_{F,L}$, und unter Erweiterung der Kosten bei Überbestand durch einen Schrottwert g [GE/ME], mit $c_o = c_B - g$. Somit gilt

$$F_Y(q^*) = \frac{c_u}{c_o + c_u} = \frac{p - c_B + c_{F,L}}{(c_B - g) + (p - c_B + c_{F,L})} = \frac{p - c_B + c_{F,L}}{p - g + c_{F,L}} \quad (2.42)$$

Diese abweichenden Formulierungen verdeutlichen die Möglichkeiten einer Erweiterung der Modelle unter Einbeziehung weiterer Kostengrößen bei Überbeständen durch zu hohe Prognosen und bei Unterbeständen durch zu niedrige Prognosen.¹

Obwohl einperiodigen Modellen zur Disposition nur eine eingeschränkte theoretische Bedeutung nachgesagt wird [Assfalg 1976, S. 47], finden sie in spezifischen Bedarfssituationen dennoch Anwendung in der Praxis [Lee und Nahmias 1993, S. 26; Arnold et al. 2004, S. A3-66]. Das einperiodige Modell kann Bestellmengenprobleme für Waren mit begrenzter Haltbarkeit und kurzer Lebensdauer abbilden, die (1) schnell veralten, etwa Tageszeitungen, Tannenbäume oder „Lifestyle“-Güter mit kurzen Produktlebenszyklen, die (2) schnell verderben, etwa frische Waren wie Obst, Gemüse oder Blumen, die (3) nur einmal auf Lager genommen werden können, etwa technische Ersatzteile oder saisonale Produkte mit langen Wiederbeschaffungszeiten im Textilhandel, oder die (4) über eine einzelne Periode hinaus eine unsichere Zukunft besitzen [Hillier und Lieberman 1997, S. 625; Hochstädter 1969].² Im letzteren Fall werden die Kosten eines Überbestandes c_o durch den Kapitaldienst einer Periode ersetzt, und so das Modell auf eine mehrperiodige Modellierung erweitert.

Wie für die klassische Losgrößenformel des deterministischen Bedarfs wurden auch für das Newsboy-Modell für stationär-stochastischen Bedarf diverse Erweiterungen entwickelt, die das Modell sukzessive an die stochastische Realität annähern. Eine umfassende Diskussion der Erweiterungen im historischen Zusammenhang findet sich bei HOCHSTÄDTER [1969] sowie bei LEE und NAHMIA [1993].³ Dabei sind insbesondere die Erweiterungen des Newsboy-Modells zur Berücksichtigung von Anfangsbeständen und fixen Bestellkosten relevant, da sie den Übergang zu komplexeren Modellformen aufzeigen [Hillier und Lieberman 1997, S. 627-628; Hochstädter 1969].

Das einperiodige Newsboy-Modell kann zur Berücksichtigung eines Lagerbestands a [ME], mit $a \geq 0$, am Anfang der Periode erweitert werden [siehe u.a. Assfalg 1976, S. 47 ff.; Lee und Nahmias 1993, S. 26 ff.; Hillier und Lieberman 1997, S. 627 f.]. Unter Berücksichtigung von a wird die Problemstellung von der Bestimmung der Bestellmenge q auf die Bestimmung eines optimalen Bestandsniveau S [ME] unmittelbar nach der Lieferung erweitert, mit

¹ Die Kostenkomponenten können je nach Anwendung um zusätzliche Lagerungskosten der Verschrottung, Veräußerungserlöse oder weitere Fehlmengenkosten des entstandenen Goodwill-Verlustes erweitert werden.

² Der letztgenannte Fall (4) dient auch zur Legitimation der Anwendung von Newsboy-Modellen bei Bestellmengenproblemen mit instationärem Bedarf mit kurzem Prognosehorizont unter einer rollierenden Planung, etwa bei der Warendisposition an Warenautomaten in Kapitel 6.

³ Beispielsweise untersuchen HILLIER und LIEBERMANN [1997, S. 628] die Berücksichtigung nichtlinearer Fehlmengenkosten, und SILVER et al. [1998, S. 393 ff.] erweitern das Modell zur Berücksichtigung mehrerer Produkte unter Kapazitätsrestriktion des Lagerbestandes, jedoch unter Annahme unabhängiger Bedarfe. Darüber hinaus werden auch die Annahmen einer deterministischen Liefermenge und Lieferzeit durch ihr stochastisches Äquivalent ersetzt [Hochstädter 1969; Naddor 1971]. NAHMIA [1982] untersucht Waren mit begrenzter Lebensdauer in mehreren Perioden unter direkter Erweiterung des Newsboy-Modells.

$$S = a + q \quad . \quad (2.43)$$

Das Modell mit Anfangsbeständen führt zu einer identischen Entscheidungsregel wie in (2.37), jedoch für die Bestimmung von S^* unter anschließender Berechnung von q^* aus der Differenz von Anfangsbestand und optimalem Bestandsniveau, $q = S - a$ [Hillier und Lieberman 1997, S. 627]. Für einen Anfangsbestand von Null, $a = 0$, ist das Bestandsniveau S identisch mit der Bestellmenge q , $S = q$, und führt zu einer identischen Problemformulierung wie in (2.31) und der Entscheidungsregel in (2.37) [Hillier und Lieberman 1997, S. 625]. Ferner können Restbestände an Waren W [ME] berücksichtigt werden, in Höhe von

$$W = S - \min(y, q) \quad . \quad (2.44)$$

Positive Anfangs- und Endlagerbestände gestatten so einen Übergang zu mehrperiodigen Modellen. Darüber hinaus erweitern sie die Modellform zu weiterführenden Lagerbestands- und Bestellpunktsystemen, die durch den Entscheidungsparameter des Bestandsniveaus S gekennzeichnet werden.

Die Einbeziehung fixer Bestellkosten K_B [GE] pro Bestellung erweitert die Ermittlung des optimalen Bestandsniveaus S einer Periode um den kritischen Bestellpunkt s [ME], ab dessen Erreichen eine Bestellung in Höhe von q zum Auffüllen auf das optimale Bestellniveau oder die Bestellhöhe S bestellt wird, mit $q = S - s$ [Hillier und Lieberman 1997, S. 631 ff.; Assfalg 1976, S. 47 ff.]. Die Differenz von $S - s$ bezeichnet somit die Mindestbestellmenge, ab der eine Bestellung ausgelöst wird. Die Berücksichtigung fixer Kosten erweitert die einperiodigen Modelle der Bestellmengenrechnung unter Bestimmung von q oder S zur Bestimmung von s und S , was als einperiodige (s, S) -Bestellmengenpolitik bezeichnet wird [Arrow et al. 1951].

Erweiterungen des Newsboy-Problems auf zweiperiodige Modelle [Naddor 1971, S. 143 ff.; Hochstädter 1969] bilden den Übergang zu mehrperiodigen Modellen (z.B. das Flowergirl-Modell), die im nachfolgenden Abschnitt diskutiert werden.

2.3.3.1.2 Mehrperiodige stationär-stochastische Modelle

Das einperiodige Modell für stationär-stochastischen Bedarf kann auf die Bestimmung optimaler Bestellmengen und Bestellzeitpunkte in zwei- und mehrperiodigen Modellen erweitert werden. In der Literatur existiert eine nahezu unüberschaubare Anzahl von Modellen zur Bestellmengenrechnung unter Variation der einzelnen Modell- und Kostenparameter zur Abbildung empirischer Situationen [Toporowski 1996].

Bei stationär-stochastischen Ausprägungen von Bedarf, Wiederbeschaffungszeit und Liefermengen wird der Lagerbestand durch den Erwartungswert des durchschnittlichen Warenbestandes gebildet. Bestellmenge und Bestellzeitpunkt müssen in Abhängigkeit vom tatsächlichen Nachfrageverlauf im Zeitablauf angepasst werden, was eine kontinuierliche Überwachung des Lagerbestands bzw. eine Überprüfung des Bestands in bestimmten Zeitabständen erfordert [Toporowski 1996, S. 126; Müller-Hagedorn 1998, S. 544]. Entsprechend werden Modelle hinsichtlich der Flexibilität von Bestellmenge und Bestellzeitpunkt unter Berücksichtigung von Sicherheitsbeständen unterschieden, die allgemein als Bestellmengenmodelle (engl. *inventory control systems*) bezeichnet werden [Silver und Peterson 1985, S. 256].

Die stochastischen Bestellmengenmodelle werden anhand von vier Entscheidungsvariablen charakterisiert [Götz 1999, S. 52; Toporowski 1999a, S. 198]:

- (1) Bestellpunkt s [ME] als Warenbestand, bei dem eine Bestellung ausgelöst wird
- (2) Bestellzyklus r [ZE] als konstanter Bestellzyklus $\tau_1 = \tau_2 = \dots = r$

(3) Bestellmenge q [ME] als konstante Menge im Zeitablauf

(4) Bestellniveau S [ME] als Warenbestand, bis zu dem durch eine Bestellung aufgefüllt wird

Die Entscheidungsvariablen s und r steuern den Zeitpunkt einer Bestellung, während q sowie S die Menge der Bestellung steuern. Gleichzeitig repräsentieren s und S im Zeitablauf veränderliche Entscheidungsvariablen, während r und q im Zeitablauf als konstant gelten. Durch Kombination der vier Parameter werden vier grundlegende Bestellmengenmodelle gebildet, die (s, q) -, die (r, q) -, die (s, S) - und die (r, S) -Politik, die nach den Steuerungsgrößen auch in Bestellzeitpunkt- und Bestellmengen-Modelle unterteilt werden. In Tabelle 2.3 ist die Kategorisierung der Entscheidungsvariablen und der resultierenden Bestellmengenpolitiken dargestellt, die anschließend diskutiert werden.

Tabelle 2.3: Entscheidungsvariablen der Bestellmengenmodelle¹

Bestellzeitpunkt-Entscheidung		Bestellmengen-Entscheidung		
		fixe Bestellmenge		variable Bestellmenge
		Bestellmengenmodell q		Bestellniveaumodell S
variabler Bestellzeitpunkt	Bestellpunkt- modell, s	(s, q)	(s, S)	(r, s, S)
fixer Bestellzeitpunkt	Bestellrhythmus- modell, r	(r, q)	(r, S)	

Die Bestellzeitpunktmodelle mit variablem Bestellpunkt (s, q) und (s, S) erfordern eine kontinuierliche Bestandsüberwachung, da nur so garantiert werden kann, dass genau zu dem Zeitpunkt eine Bestellung ausgelöst wird, in dem der Bestand den Bestellpunkt s erreicht [Toporowski 1998; Silver und Peterson 1985].² Entspricht der Bestand dem Bestellpunkt s , so wird durch ein Bestellmengenmodell entweder eine konstante Menge q oder durch ein Bestellniveaumodell eine variable Menge q zur Erreichung des Bestandsniveaus S bestellt, als Differenz $q = S - a$ aus dem vorgegebenen Bestellniveau S und dem Anfangsbestand a zum Kontrollzeitpunkt [Toporowski 1998]. Die bestellte Menge q trifft am Ende der Wiederbeschaffungszeit ein, die deterministischer oder stochastischer Art sein kann. Die Lagerbestandsentwicklung einer (s, q) -Politik mit Bestellpunkt und konstanter Bestellmenge ist in Abbildung 2.14 dargestellt.

¹ Eigene Darstellung in Anlehnung an [Toporowski 1999b].

² Dies erfordert zudem die Prämisse, dass der Bedarf einer Ware nur in einzelnen Mengeneinheiten auftritt.

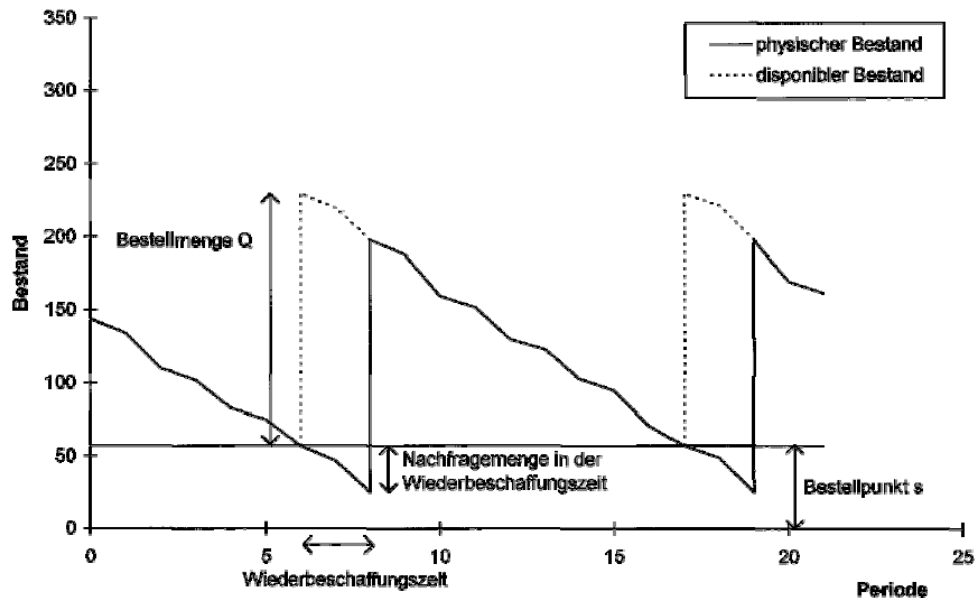


Abbildung 2.14: Lagerbestandsentwicklung bei einer stochastischen (s, q) -Politik¹

In Bestellrhythmus- bzw. Bestellzyklusmodellen (r, q) und (r, S) wird dagegen nach einem konstanten Zeitintervall r der Bestand kontrolliert und regelmäßig eine Bestellung ausgelöst, so dass keine kontinuierliche Bestandsüberwachung erforderlich ist [Müller-Hagedorn 1998].² Dabei wird durch ein Bestellmengenmodell entweder eine konstante Bestellmenge q bestellt oder durch ein Bestellniveaumodell eine variable Bestellmenge q berechnet, mit $q = S - a$, die den disponiblen Bestand als Summe von physischem und Auftragsbestand auf S aufstockt. Da der zeitliche Bestellrhythmus vorgegeben ist, sind auch die fixen Bestellkosten exogen bestimmt und somit nicht entscheidungsrelevant [Naddor 1971, S. 66]. Die Lagerbestandsentwicklung einer (r, S) -Politik mit rhythmischer Bestandskontrolle und Bestellung einer variablen Bestellmenge $q = S - a$ bis auf ein Bestellniveau S ist in Abbildung 2.15 dargestellt.

Auf die Darstellung der theoretisch denkbaren (r, q) -Modelle kann verzichtet werden, da eine regelmäßige Bestellung identischer Mengen ohne Berücksichtigung des tatsächlichen Lagerbestandes keine stochastischen Veränderungen des Bestands berücksichtigen kann [Toporowski 1998, S. 1144].

Im Rahmen einer (s, S) -Politik wird die variable Bestellmenge q so gewählt, dass der disponible Bestand auf S aufgestockt wird. Unter der Annahme einer kontinuierlichen Bestandsüberwachung beträgt der disponible Bestand zum Zeitpunkt der Bestellung exakt s , so dass die Bestellmenge q konstant ist, mit $q = S - s$. Da in diesem Fall die (s, q) - und die (s, S) -Politik identisch sind, kann auf eine separate Darstellung der (s, S) -Politik ebenfalls verzichtet werden. Unterschiede zwischen der (s, q) - und der (s, S) -Politik entstehen erst, wenn die Prämisse einer kontinuierlichen Bestandsüberprüfung aufgegeben wird oder mehrere Güter gleichzeitig entnommen werden können. In diesem Fall kann der Bestand zum Zeitpunkt der Disposition unter den Bestellpunkt s sinken [Toporowski 1999b, S. 328].

¹ Darstellung entnommen aus TOPOROWSKI [1999b, S. 199].

² Wie in Abschnitt 2.3.1.1 dargestellt beschreibt der Bestellrhythmus die konstante zeitliche Verteilung der Bestelltermine t_i über den Planungszeitraum mit dem Bestellzyklus τ_i als Zeit zwischen zwei aufeinander folgenden Bestellterminen.

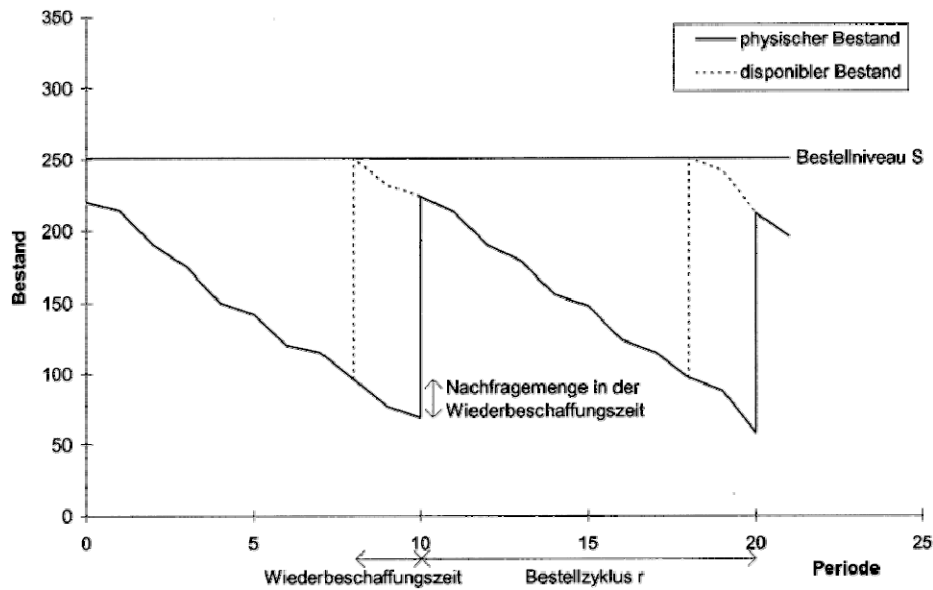


Abbildung 2.15: Lagerbestandsentwicklung bei einer stochastischen (r, S) -Politik¹

Wird auf die Annahme einer kontinuierlichen Bestandsüberwachung verzichtet und der Bestand periodisch überprüft, ohne zwingend eine Bestellung auszulösen, bezeichnet dies eine (r, s, S) -Politik mit vorgegebener Überprüfungsdauer r als Kombination der (s, S) - und (r, S) -Politiken [Toporowski 1998, S. 1145-1146]. Der Bestand wird im Abstand von r Perioden hinsichtlich seines Unterschreitens des Bestellpunktes s geprüft und in diesem Fall eine Bestellung zum Erreichen des Bestellniveaus S ausgeführt [Müller-Hagedorn 1998, S. 529]. Eine (r, s, S) -Politik mit regelmäßiger Bestandsprüfung und variabler Bestellmenge ist in Abbildung 2.16 dargestellt.

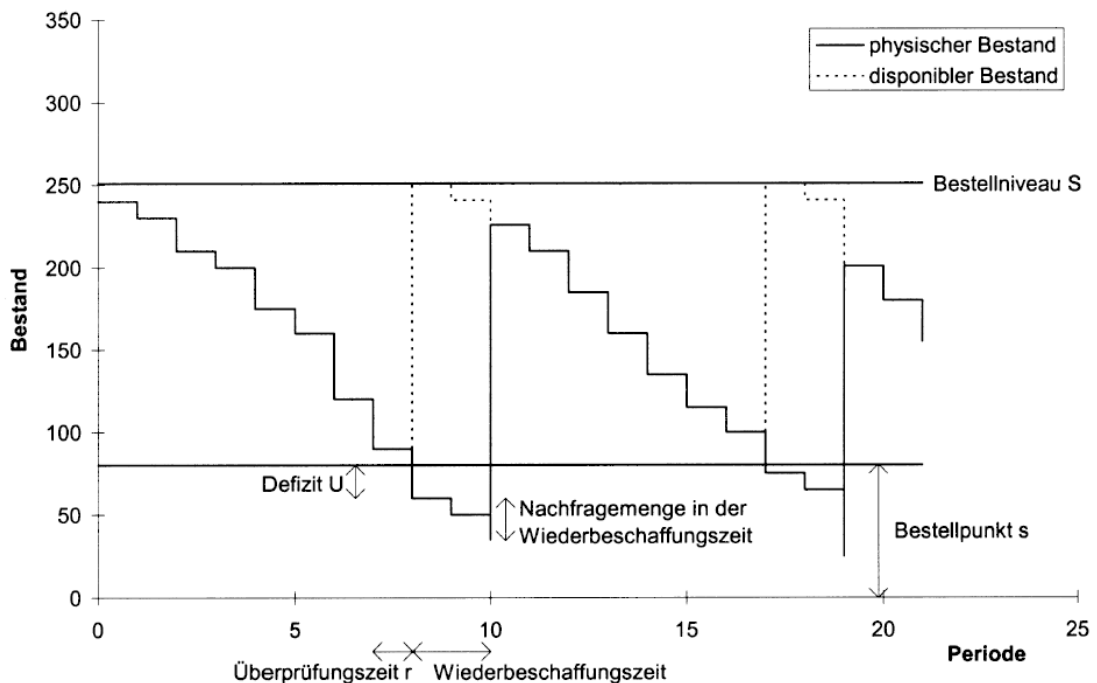


Abbildung 2.16: Lagerbestandsentwicklung bei einer stochastischen (r, s, S) -Politik²

¹ Darstellung entnommen aus TOPOROWSKI [1999b, S. 325].

² Darstellung entnommen aus TOPOROWSKI [1999b, S. 329].

Für ein Überprüfungsintervall von eins, $r = 1$, bezeichnet die (r, s, S) -Politik folglich eine $(1, s, S)$ -Politik und ist somit identisch mit einer (s, S) -Politik, die daher als Sonderfall einer (r, s, S) -Politik betrachtet werden kann. Die Bestellregeln der Lagerhaltungsmodelle in einstufigen Bestandssystemen können als Spezialfälle allgemeiner Dispositionsregeln zur Festlegung der Bestellmengen und -zeitpunkte betrachtet werden [Müller-Hagedorn 1998, S. 544]. Entsprechend lässt sich das Newsboy-Modell auch als $(1, S)$ -Politik und somit als Vereinfachung der verbreiteten (r, S) -Politiken (ohne Anfangsbestand und somit $q = s$) interpretieren, wobei dies ebenfalls die Entscheidungssituation in der letzten Periode mehrperiodiger Systeme modelliert [Assfalg 1976].¹

Die Überlegenheit einzelner Bestellmengenpolitiken wurde vielfach analytisch untersucht, jeweils für unterschiedliche Konstellationen von Entscheidungsvariablen und Parametern [siehe u.a. Naddor 1971, S. 225-289 für eine Untersuchung der Standardmodelle]. Es ist jedoch zu beachten, dass die Modellauswahl im Handel häufig extern durch die Struktur der Wertschöpfungskette, durch die Koordination der Warendisposition in der Distributionskette oder andere Marktgegebenheiten bestimmt wird. Dies kann auch durch logistische Engpassfaktoren begründet sein [Müller-Hagedorn 1998, S. 544-554; Arnolds et al. 2001, S. 95]. Obwohl in Einzelhandelsfilialen mit elektronischen Kassen eine kontinuierliche Überwachung der Warenbestände und somit eine variable Bestellpunkt-Politik möglich ist, ist im Handel eine rhythmische Artikeldisposition an einem Lieferanten-Dispositionstag üblich [Becker und Schütte 1996; Götz 1999, S. 55; Schneeweiß 1981]. Sind die Bestelltermine im Handel exogen vorgegeben, können sie nicht vom Entscheidungsträger beeinflusst werden und stellen Zustandsparameter der Bestellmengenrechnung dar [Toporowski 1999a; Toporowski 1999b]. In diesem Fall müssen Bestellrhythmus-Modelle angewendet werden. Diese erfordern weniger Bestandskontrollen und begünstigen eine koordinierte Bestellung von Waren und ihre Anlieferung in einem festen Tourenplan aus übergeordneten Lagerstufen [Arnolds et al. 2001, S. 95; Toporowski 1999b, S. 325], führen jedoch zu einer verlängerten Dispositionszeit durch die Zeit des Überprüfungsintervalls, aus der höhere Sicherheitsbestände und entsprechend höhere Durchschnittsbestände resultieren [Arnolds et al. 2001, S. 95]. Darüber hinaus können die schwankenden Bestellmengen zu warenlogistischen Problemen durch Kapazitätsrestriktionen führen [Toporowski 1999b, S. 325]. Bei der empirischen Warendisposition in Handelsfilialen dominieren Bestellrhythmusverfahren zur vereinfachten Koordination regelmäßiger Bestell- und Anlieferintervalle in standardisierten Logistiknetzwerken und zur Steuerung logistischer Engpassfaktoren, z. B. der Anzahl von Rampen zur Warenannahme durch Heuristiken.

Neben den Entscheidungsvariablen r , s , q und S sind zur vollständigen Modellbeschreibung ferner die Merkmale der Zielfunktion relevant. Dabei ist vor allem die Art der Berücksichtigung von Fehlmengen durch Kosten oder Servicegrade im Verlust- oder Vormerkfall, der Umfang und die funktionale Form der Kostenparameter, die Verteilungsannahmen der stationär-stochastischen Nachfrage sowie der Lösungsweg zur Bestimmung kostenminimaler Bestellmengen relevant [Toporowski 1998, S. 1146]. TOPOROWSKI [1998, S. 1146 ff.] untersucht unterschiedliche Modellierungsansätze zur Berücksichtigung von Fehlmengen, sowohl unter expliziter Modellierung der Fehlmengenkosten in der Zielfunktion als auch unter Abbildung eines vorgegebenen Lieferservicegrades in den Nebenbedingungen. Er zeigt, dass die Verwendung von Kosten-, α - oder β -Servicegraden zu unterschiedlichen Modellen der Bestellmengenpolitik und zudem zu unterschiedlichen Dispositionskosten führt. Das

¹ Die Verwendung der konstanten Bestellmenge q einer Periode in der Herleitung des Newsboy-Modells erfolgte aus didaktischen Gründen und in Anlehnung an die übliche Notation in der Literatur [Assfalg 1976; Domschke et al. 1997; Silver et al. 1998]. Da kein Anfangslagerbestand berücksichtigt wurde, gilt $q = S$, und die Herleitung kann identisch unter Berücksichtigung eines Bestandsniveaus S erfolgen [Hillier und Lieberman 1997]. Somit kann die $(1, q)$ -Politik in diesem Fall auch als $(1, S)$ -Politik aufgefasst werden.

Bestellmengenproblem wird als Gleichungssystem unter Einhaltung optionaler Nebenbedingungen zur Bestimmung optimaler Bestellmengen oder -zeitpunkte formuliert [Götz 1999; Toporowski 1998; Toporowski 1999a; Toporowski 1999b]. Dabei existieren alternative Lösungswege zur Festlegung der Entscheidungsvariablen in der Minimierung der Zielfunktion [Tempelmeier 2003; Silver und Peterson 1985]. In der Literatur werden sowohl Ansätze der simultanen als auch der sukzessiven Lösung spezifischer Bestellmengenmodelle bei stochastischer Nachfrage diskutiert [Götz 1999, S. 135-139; Toporowski 1996], wobei TOPOROWSKI [1999a; 1999b] Unterschiede in der Dispositionsgüte für unterschiedliche Lösungsansätze bei konstanten Parametern feststellt.

Zudem beeinflusst die korrekte Berücksichtigung der Unsicherheit in der Modellierung die Formulierung der Bestellmengenmodelle. In den dargestellten stochastischen Modellen der Bestellmengenrechnung werden Bedarf, Liefermengen und Wiederbeschaffungszeit als stationäre Zufallsvariablen mit einer spezifischen Wahrscheinlichkeitsverteilung definiert [Toporowski 1998]. In Abhängigkeit von der empirischen Verteilungsform des Bedarfs oder bei Verwendung unterschiedlicher theoretischer Verteilungsfunktionen einer Normal-, Gamma- oder Poissonverteilung ändert sich die Zielfunktion des Modells, und es werden unterschiedliche optimale Bestellmengen sowie -zeitpunkte berechnet [Toporowski 1998, S. 1154]. Die Diskussion darüber, ob die Auswahl der Verteilungsannahme einen erheblichen Einfluss auf die Höhe des Sicherheitsbestandes ausübt oder nicht, wird daher in Abschnitt 3.2.2.3 detailliert aufgegriffen [siehe auch Toporowski 1996, S. 129; Mentzer und Krishnan 1985; Eppen und Martin 1988]. Nachfolgend wird die Komplexität der Modelle durch die Berücksichtigung instationär-stochastischer Bedarfe erweitert.

2.3.3.2 Modelle für instationär-stochastischen Bedarf

Die Mehrheit empirischer Bedarfe im Handel zeichnet sich durch instationäre Zeitreihenverläufe mit Trend, Saison, Ausreißern, Niveauverschiebungen, Strukturbrüchen sowie unregelmäßigen exogenen Einflüssen durch Marketing- und Vertriebsaktivitäten aus.¹ Nach ASSFALG [1976, S. 50] bedeutet gerade die Einbeziehung dynamischer, instationärer Elemente in die stochastische Bestellmengenrechnung einen weiteren wichtigen Schritt in der Annäherung an die Lagerhaltungswirklichkeit. Somit erscheint die Berücksichtigung instationärer Bedarfe relevant und verspricht durch die gesteigerte Realitätsnähe eine zunehmende Entscheidungsqualität in der Bestellmengenrechnung.

SILVER et al. [1998, S. 341] gehen zudem davon aus, dass eine signifikante Veränderung der Verteilungsparameter – und insbesondere des durchschnittlichen Bedarfs – erhebliche Auswirkungen auf die Auswahl wirtschaftlicher Bestellpolitiken und die Bestimmung optimaler Kontrollparameter q , s und S hat. Verändert sich der mittlere Bedarf \hat{y}_y , verändern sich auch die Parameter (r, s, S) der Bestellpolitik. Somit können Bestellmengenmodelle für stationären Bedarf nicht ohne Weiteres für instationären Bedarf verwendet werden.

Im Gegensatz zu der Fülle von Veröffentlichungen zu stationär-stochastischen Modellen der Bestellmengenrechnung existieren jedoch nur wenige Untersuchungen für instationär-stochastischen Bedarf. Die Mehrheit der Monographien, unter anderem [Silver et al. 1998; Naddor 1971; Graves et al. 1993; Hochstädter 1969; Lee und Billington 1993], untersucht ausschließlich Modelle für stationären stochastischen Bedarf, nicht aber für instationären Bedarf. DONALDSON [1977] untersucht eine Politik des optimalen Bestellzeitpunktes bei stochastischem, instationärem Bedarf mit deterministischem Trend. SONG und ZIPKIN [1993] untersuchen ein Modell, in dem die Bedarfsrate durch exogene Faktoren im Zeitablauf variiert. GEUNES [2001] erweitert das Newsboy-Modell auf instationäre Bedarfe durch eine

¹ Von den 3003 empirischen Zeitreihen der M3-Competition bestand ebenfalls nur ein Bruchteil aus Zeitreihen mit stationärem Verlauf [Makridakis und Hibon 2000].

einfache Heuristik, da eine analytische Lösung bei instationären Parametern für die betrachtete Problemfamilie kaum möglich ist.¹ ASSFALG [1976] schlägt eine pragmatische Lösung für die Instationarität durch saisonale Bedarfsverläufe vor. Er postuliert eine Ausdehnung der Planungsperiode unter ausreichend langer Wahl der Beobachtungsperioden, so dass nicht „bereits saisonale oder zufällige Schwankungen der Nachfrage als Veränderung gelten“ [Assfalg 1976, S. 17]. Diese Möglichkeit ist jedoch bei exogen bestimmten Bestellzyklen oder festem Bestellrhythmus nicht gegeben, der jedoch im Handel überwiegt.²

Die geringen Forschungsanstrengungen hinsichtlich instationärer Bedarfe sind darauf zurückzuführen, dass die Form des Bedarfs auch die Komplexität der Modelle nachhaltig bestimmt [Lee und Nahmias 1993, S. 6]. Die gesteigerte Realitätsnähe geht mit einem erheblichen Anstieg der Komplexität in der analytischen Lösung stochastischer Modelle einher [Assfalg 1976, S. 50; Götz 1999, S. 75], die zu einem erheblichen Mehraufwand führt [Toporowski 1996, S. 113]. Nach GÖTZ rechtfertigt die zusätzliche Realitätsnähe in der Berechnung stochastischer Modelle nur selten den zusätzlichen Aufwand [Götz 1999; ebenso Becker und Schütte 1996].³ Somit erscheint eine exakte Modellierung instationär-stochastischer Bedarfe zu kompliziert für die regelmäßige Anwendung in der Praxis [Silver et al. 1998, S. 342].

Aus diesem Grund wird die Komplexität der dynamischen, instationären Realität durch die Verwendung von Prognoseverfahren zur instationären Bedarfsprognose und anschließender stationär-stochastischer Bestellmengenrechnung reduziert. Die Abbildung veränderlicher Bedarfsverläufe und Bestellniveaus aus instationären Zeitreihen erfordert eine adaptive Parameteranpassung, in der die Parameter der Dichtefunktion rollierend für jede Planungsperiode erneut prognostiziert werden [Arnolds et al. 2001, S. 97; Tempelmeier 1992; Silver et al. 1998, S. 342].⁴ Der Bedarf wird dazu nicht länger durch die ersten beiden Momente einer stationären Dichtefunktion, nämlich den Erwartungswert \hat{y}_Y und die Standardabweichung $\hat{\sigma}_Y$, beschrieben, sondern durch zeitabhängige Parameter eines Prognoseverfahrens. Der instationäre Mittelwert \hat{y}_{t+h} wird für einen gegebenen Prognosehorizont h oder eine Reihe von Horizonten $h = 1, 2, \dots, H$ vorhergesagt. Anstelle der Standardabweichung $\hat{\sigma}_Y$ der Bedarfswerte y wird die Streuung im instationären Fall durch die Standardabweichung $\hat{\sigma}_e$ der Residuen des Prognoseverfahrens bestimmt, die sich als Prognosefehler e_t aus der Differenz zwischen tatsächlichem Bedarfswert einer Periode y_t und prognostiziertem Bedarfswert \hat{y}_t errechnen: $e_t = y_t - \hat{y}_t$ (siehe ausführlich Abschnitt 3.3.2). Für die Zeitreihe der Prognoseresiduen gelten dieselben Anforderungen wie ursprünglich für die Dichtefunktion des Bedarfs: die Prognosefehler sind stationär (siehe Abschnitt 2.2.3.1 sowie Abschnitt 3.2.3.1) und folgen einer bekannten, vollständig spezifizierten Wahrscheinlichkeitsverteilung. Zusätzlich muss der Erwartungswert der Prognosefehler Null sein, $E(e_t) = 0$, so dass die Sicherheitsbestände ohne systematische Verzerrung (engl. *bias*) zu hoch oder zu niedrig bestimmt werden. Somit wird davon ausgegangen, dass ein ausgewähltes Prognoseverfahren die Instationarität des Bedarfs vollständig beschreibt und somit Mittelwert und Standardabweichung in den konven-

¹ Darüber hinaus existieren Erweiterungen der anderen Eingangs- und Ausgangsströme des Lagers auf instationären Verlauf, so durch GULLU et al. [1997] für Belieferungen nur zu ausgewählten, instationären Zeitpunkten, jedoch unter der vereinfachenden Annahme eines deterministischen Bedarfs, oder für mehrstufige Probleme [Joneja 1991]. SOX [1997] betrachtet instationäre Kostenparameter. SILVER et al. [1998, S. 342] geben eine Übersicht über die wenigen weiteren Ansätze in der wissenschaftlichen Literatur.

² Vielmehr sieht sich der Handel bei täglicher Disposition eher einer mehrfachen, sich überlagernden Saisonalität der Wochentage, Wochen im Monat und der Kalenderwochen im Jahr gegenüber.

³ GÖTZ [1999] erweitert seine Aussagen jedoch unzulässigerweise auf alle stochastischen Bedarfsarten – unabhängig von der Stationarität des Bedarfs – und postuliert, dass daher selbst Modelle für stationär-stochastischen Bedarf in der Handelspraxis nur selten praktischen Einsatz finden.

⁴ Nach MINNER [2000, S. 11] gerät die Stochastische Dynamische Programmierung bereits in der Lösung stationärer Bestellmengenprobleme bei einer großen Anzahl zukünftiger Zustände an die Grenzen der Berechenbarkeit, weshalb das Konzept der rollierenden Planung als Lösungsansatz vorzuziehen ist.

tionellen Formeln der Bestellmengenrechnung verwendet werden können.¹ Die so geschätzten Parameter des Bedarfs werden anschließend zur Bestimmung der Entscheidungsvariablen² q , s , und S unter Konstanz aller Parameter für die Planungsperiode und einen weiterreichenden Planungshorizont verwendet, bei rollierender Anpassung in jedem Dispositionszeitpunkt. Nach SILVER et al. [1998, S. 342] und KAUFMANN [1977] führt diese Vorgehensweise nur zu geringfügigen Ineffizienzen, obwohl sie dynamische Veränderungen des Bedarfs und somit der Bestellniveaus für die Bestimmung kostenoptimaler Bestellmengen und –zeitpunkte vernachlässigt.³

Darüber hinaus wurden einzelne Ansätze entwickelt, die zunächst den stochastischen Bedarf bestimmen und ihn anschließend in den Modellen zur Ermittlung optimaler Bestellmengen bei instationär-deterministischem Bedarf verwenden. SILVER [1978] sowie ASKIN [1981] ermitteln die instationär-stochastischen Bedarfsparameter wie zuvor, erlauben aber eine Variation der Bedarfsprognose \hat{y}_t innerhalb der Planungsperiode durch Verwendung der deterministischen SILVER-MEAL-Heuristik. VARGAS [1996] entwickelt hierzu eine stochastische Variante des WAGNER-WHITIN-Algorithmus. Die Mehrzahl der Ansätze vernachlässigt jedoch instationäre Entwicklungen der Streuung, die sich ebenfalls im Zeitablauf verändern kann [Silver et al. 1998], beispielsweise in Abhängigkeit saisonaler Muster, die zeitlich variierende Sicherheitsbestände für Phasen geringerer oder größerer Zufallsschwankungen erfordern.

Abschließend ist festzustellen, dass einperiodige Bestellmengenmodelle für instationär-stochastischen Bedarf mit den Varianten der Newsboy-Modelle für stochastisch-stationären Bedarf identisch sind, da die Bedarfsschwankungen rollierend erfasst werden können. Dies erlaubt eine Erweiterung auf mehrperiodige Problemstellungen für den Sonderfall eines Prognosehorizontes $h = 1$, wobei in jeder Periode der Prognosewert \hat{y}_{t+1} und die Standardabweichung der vergangenen Prognosefehler $\hat{\sigma}_E$ rollierend für die jeweils nachfolgende Periode bestimmt werden. Somit kann das einperiodige Modell zur Disposition an Warenautomaten im Verfahrensvergleich des Kapitels 6 trotz instationärer Bedarfe die Bestellmengen unter Verwendung des Newsboy-Modells aus Abschnitt 2.3.3.1.1 berechnen. Der instationär-stochastische Charakter der ursprünglichen Problemstellung wird somit im Experiment berücksichtigt und so eine gesteigerte Realitätsnähe in der Lösung des empirischen Vergleichs unterschiedlicher Prognoseverfahren erzielt.

Die in den vorangehenden Abschnitten dargestellten Modelle, ihre Varianten und grundlegenden Strukturdimensionen zeigen die Vielfalt der Bestellmengenpolitiken auf, insbesondere bei deterministischem Bedarf und bei stationärer, stochastischer Nachfrage. Die Bestimmung der Entscheidungsvariablen und Parameter gestaltet sich selbst im Rahmen einfacher Modelle aufwendig [Toporowski 1996] und erlaubt nicht immer eine Berücksichtigung der Komplexität des Bedarfs. Aufgrund der Vielzahl alternativer Modelle der Bedarfs- und Bestellmengenrechnung sowie zur laufenden Überprüfung der Modelle sind Gütekriterien erforderlich, die nachfolgend untersucht werden.

¹ Nach SILVER et al. [1998, S. 253] kann auch bei stationär-stochastischem Bedarf der durchschnittliche Bedarf geringfügig variieren, so dass die Entscheidungsregeln ebenfalls rollierend im Zeitablauf adaptiert werden.

² Dabei erscheint eine dynamische Anpassung des Überprüfungs- und Dispositionsintervalls r aufgrund der verbundenen ablauforganisatorischen und warenlogistischen Prozesse ungeeignet, die Veränderung von s und S dagegen möglich.

³ Die Annahme stationärer Verläufe oder eine Verwendung der klassischen Losgrößenrechnung können dagegen allenfalls als grobe Näherungen betrachtet werden [Arnolds et al. 2001, S. 97]. Eine Vereinfachung der Modellkomplexität instationär-stochastischer Bedarfsverläufe durch Prognose und anschließende Betrachtung als deterministische Bedarfsgrößen in deterministischen Bestellmengenmodellen ohne Sicherheitsbestände, u. a. bei [Toporowski 1996, S. 112; Götz 1999], oder durch einfache Bestellheuristiken [Arnolds et al. 2001, S. 97] stellt eine unzulässige Vereinfachung dar, die essentielle Modellkomponenten vernachlässigt und systematische Fehler verursacht.

2.4 Beurteilung von Dispositionsverfahren

2.4.1 Deskriptive Maße der Dispositionsgüte

In der betrieblichen Praxis der Warendisposition werden deskriptive, empirische Maße zur Beurteilung der Güte der Dispositionsentscheidungen durch die Modelle der Bedarfsprognose und der Bestellmengenrechnung verwendet.¹ Hier erfolgt zunächst eine Darstellung deskriptiver Gütemaße für Entscheidungen der Warendisposition, gefolgt von einer Analyse der Prognosefehler in Abschnitt 3.3 und der Kostengrößen in Abschnitt 5.2.

Die Dispositionsgüte wird durch die deskriptiven Gütekriterien der Häufigkeit und Menge aufgetretener Fehlmengen und Überbevorratungen bewertet.² Die Beurteilung erfolgt sowohl anhand der entscheidungsrelevanten Anzahl von Ereignissen der Planungsperiode als auch der Menge an Verkaufseinheiten [Arminger und Götz 1999, S. 12]. Die Beurteilung bezeichnet somit nicht die Güte der Prognoserechnung, sondern das entscheidungsrelevante Ergebnis der sequenziellen Bedarfsprognose und der anschließenden Bestellmengenrechnung. Im Gegensatz zur reinen Beurteilung der Prognosegenauigkeit erfolgt somit eine Beurteilung der Entscheidungsqualität. Deskriptive Fehlermaße werden in der Literatur weitgehend vernachlässigt und aufgrund ihrer Bedeutung für die Annäherung an die später zu definierenden Kostenmaße als alternative Beurteilungskriterien hier vorgestellt.

Als deskriptive Fehler im Rahmen der Warendisposition werden das Auftreten einer Fehlmengensituation (engl. *out-of-stock*, OOS) in einer Prognoseperiode $t+h$ und das Auftreten einer Überbevorratungs-Situation (engl. *overstocked event*, OVS) im Sinne des α -Servicegrades definiert. Ferner werden die mengenorientierten Prognosefehler der Anzahl unterbundener Verkäufe (engl. *out-of-stock units*, OOU) und der Anzahl überbevorrateter Verkaufseinheiten (engl. *overstocked units*, OVU) im Sinne der β -Servicegrade bestimmt und in Tabelle 2.4. dargestellt.

Tabelle 2.4: Deskriptive Fehler zur Beurteilung der Dispositionsgüte

	Auftreten eines Ereignisses		Anzahl in Verkaufseinheiten	
	Auftreten einer Fehlmenge	Auftreten eines Überbestandes	Anzahl unterdrückter Kaufhandlungen	Anzahl überbevorrateter Verkaufseinheiten
Einfache Fehler	$OOS_{t+h} = \begin{cases} 1 & \text{für } e_{t+h} > 0 \\ 0 & \text{für } e_{t+h} \leq 0 \end{cases}$	$OVS_{t+h} = \begin{cases} 0 & \text{für } e_{t+h} \geq 0 \\ 1 & \text{für } e_{t+h} < 0 \end{cases}$	$OOU_{t+h} = \begin{cases} e_{t+h} & \text{für } e_{t+h} > 0 \\ 0 & \text{für } e_{t+h} \leq 0 \end{cases}$	$OVU_{t+h} = \begin{cases} 0 & \text{für } e_{t+h} \geq 0 \\ e_{t+h} & \text{für } e_{t+h} < 0 \end{cases}$
	Out-of-stock event (2.45)	Overstocked event (2.46)	Number of out-of-stock units (2.47)	Number of overstocked units (2.48)

Die deskriptiven Fehlermaße können in einer marginalanalytischen Betrachtung in Übereinstimmung mit den statistischen Fehlermaßen und Kostengrößen grafisch abgebildet und interpretiert werden. In Abbildung 2.17 wird den Verlauf der deskriptiven Gütekriterien der Anzahl unterdrückter Kaufhandlungen OOU (a.), der Anzahl überbevorrateter Verkaufseinheiten OVU (b.), des Auftretens einer Präsenzlücke OOS (c.) und des Auftretens einer Überbevorratung OVS (d.) im Vergleich zum quadratischen Prognosefehler $SE(e)$ und der Differenz beider Gütemaße abgebildet.

¹ Dies steht somit im Gegensatz zu der intensiven wissenschaftliche Diskussion zur validen und reliablen Bestimmung der Prognosegüte durch unterschiedliche statistische Fehlermaße (siehe Abschnitt 3.3).

² In den nachfolgenden Vergleichsstudien wird die Dispositionsgüte unterschiedlicher Prognoseverfahren anhand der Anzahl aufgetretener Fehlmengen und Überbevorratungen verglichen, jeweils unter Konstanz des Modells zur Bestellmengenrechnung und möglicher Kostenparameter. Somit sind Änderungen in der Dispositionsgüte allein auf die Güte des Prognoseverfahrens zurückzuführen.

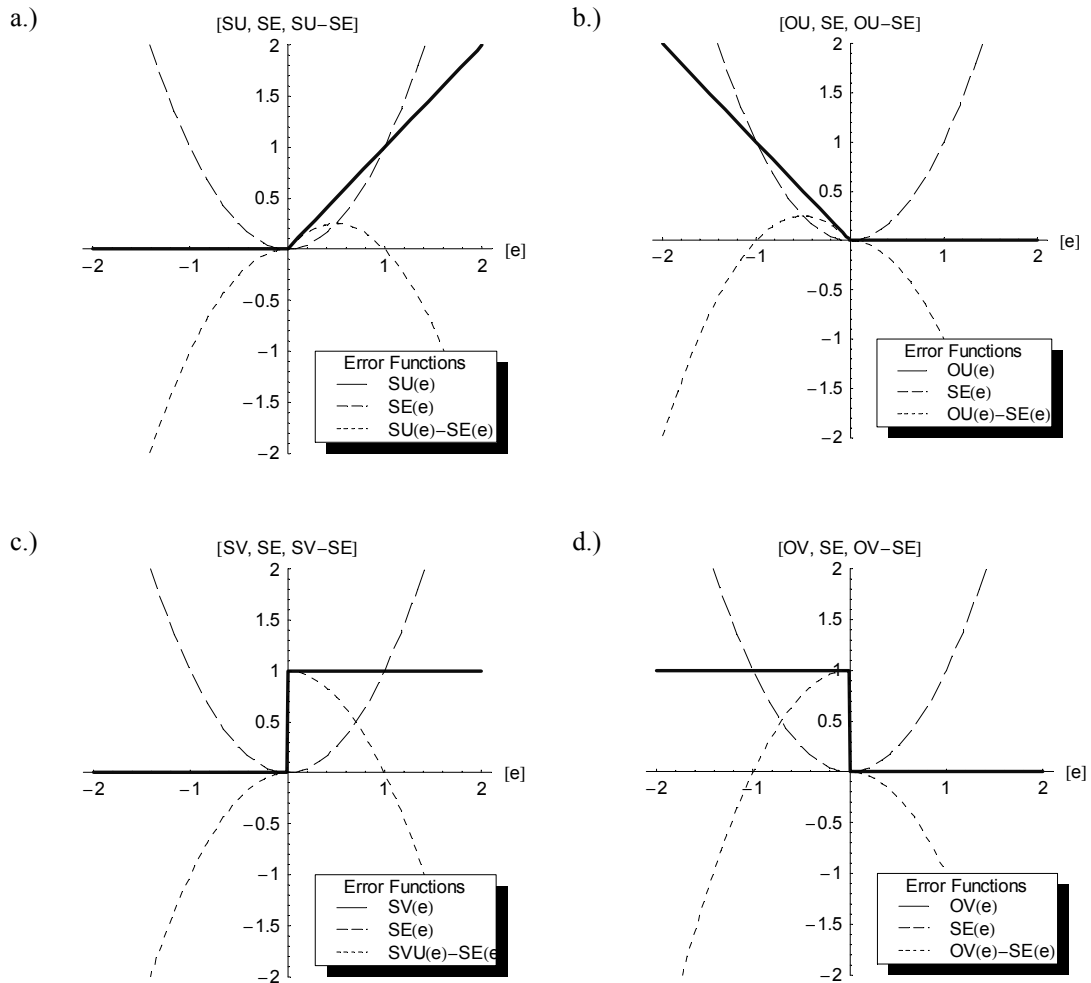


Abbildung 2.17: Fehlerverlauf deskriptiver Gütemaße des OOU (a.), OVU (b.), OOS (c.) und OVS (d.) im Vergleich zum quadratischen Fehler $SE(e)$.

Die Abbildung 2.17 verdeutlicht den asymmetrischen Fehlerverlauf aller deskriptiven Fehlermaße. Alle Fehlermaße besitzen eine Sprungstellen im Ursprung, die durch eine selektive Bewertung ausschließlich positiver oder negativer Dispositionsabweichungen entsteht [Arminger und Götz 1999, S. 12]. Der asymmetrische Verlaufs der Fehler um den Ursprung ist darauf zurückzuführen, das z. B. das Fehlermaß OOS lediglich das Auftreten einer Fehlmengensituation in einer Planungsperiode als binäre „Wahr-oder-Falsch“-Aussage bewertet, nicht aber das Auftreten eines Überbestandes (und das vice versa für OVS). Ebenso bewertet das mengenbezogene Fehlermaß OOU nur die Höhe der Fehlmenge der Periode, nicht aber die Höhe der Überbestände (und umgekehrt für OVU). Die Bewertung des Auftretens eines OOS oder OVS erfolgt dabei unabhängig von der Höhe der aufgetretenen Fehlmenge $OOS(e)$ bzw. des Überbestandes $OVS(e)$ konstant mit eins, und erlaubt eine unterschiedliche Steuerung der Warendisposition im Gegensatz zu den mengenbezogenen Kennziffern.¹ Die selektive Messung unterschiedlicher Aspekte der Dispositionsgüte durch die Fehlermaße erfordert daher ihre gemeinsame Verwendung, unter Berechnung aller vier Fehlermaße in jeder Planungsperiode.

¹ Dies impliziert trotz asymmetrischer Fehlerkosten eine identische Gewichtung von Fehlmengen und Überbeständen, im Unterschied zu den in Abschnitt 5.2.2 dargestellten asymmetrischen Kostenfunktionen.

Aus den deskriptiven Gütemaßen einzelner Dispositionshorizonte und -ursprünge können Zeitreihen von Gütemaßen gebildet und entsprechend summierte mittlere und relative Gütekriterien abgeleitet werden, die in Tabelle 2.5 abgebildet sind.¹

Tabelle 2.5: Deskriptive Fehlermaße zur Beurteilung der Dispositionsgüte

	Auftreten eines Ereignisses		Anzahl in Abverkaufseinheiten	
	Auftreten einer Präsenzlücke	Auftreten eines Überbestandes	Anzahl unterdrückter Kaufhandlungen	Anzahl überbevorrateter Verkaufseinheiten
Summierte Fehler	$SOS_l = \sum_{t=l-h}^{T-h} OOS_{t+h}$ Sum of out-of-stock events (2.49)	$SVS_l = \sum_{t=l-h}^{T-h} OVS_{t+h}$ Sum of overstocked events (2.50)	$SOU_l = \sum_{t=l-h}^{T-h} OOU_{t+h}$ Sum of out-of-stock units (2.51)	$SVU_l = \sum_{t=l-h}^{T-h} OVU_{t+h}$ Sum of overstocked units (2.52)
Mittlerer Fehler	$MOS_l = \frac{1}{T-l} \sum_{t=l-h}^{T-h} OOS_{t+h}$ Mean out-of-stock events (2.53)	$MVS_l = \frac{1}{T-l} \sum_{t=l-h}^{T-h} OVS_{t+h}$ Mean overstocked events (2.54)	$MOU_l = \frac{1}{T-l} \sum_{t=l-h}^{T-h} OOU_{t+h}$ Mean number of out-of-stock units (2.55)	$MVU_l = \frac{1}{T-l} \sum_{t=l-h}^{T-h} OVU_{t+h}$ Mean number of overstocked units (2.56)
Relativer Fehler	-	-	$MPSU_l = \frac{1}{T-l} \sum_{t=l-h}^{T-h} \frac{OOU_{t+h}}{y_{t+h}}$ Mean percent number of out-of-stock units (2.57)	$MPOU_l = \frac{1}{T-l} \sum_{t=l-h}^{T-h} \frac{OVU_{t+h}}{y_{t+h}}$ Mean percent number of overstocked units (2.58)

Die Fehlermaße des *SOU* und des *SVU* ähneln den von SCHWARZE [1980, S. 324-326] postulierten kumulierten Unter- und Überschätzungsfehlern zur Identifikation von Verfahren, die zu einer systematischen Überschätzung oder Unterschätzung der Sicherheitsbestände neigen. Das deskriptive Fehlermaß des *MOS* dient als Indikator der relativen Frequenz des Auftretens einer Fehlmengensituation im Betrachtungsintervall. Es entspricht also dem Inversen α -Servicegrad, während das Fehlermaß des *MPSU* dem Inversen β -Servicegrad entspricht (vgl. Abschnitt 2.3.1). Die deskriptiven Fehlermaße zur Ex-post-Beurteilung der Güte der Warendisposition dienen folglich der Kontrolle der bestandsorientierten Zielgrößen der Bestellmengenrechnung.

Es wird deutlich, dass die Zielkriterien der α - und β -Servicegrade die Aspekte der Überbevorratung vernachlässigen und somit kein alleiniges Zielkriterium zur Beurteilung der Dispositionsgüte sein können. Im Gegensatz zu statistischen Fehlermaßen beschränken deskriptive Fehlermaße ihre Bewertung auf einen einzelnen relevanten Tatbestand der Fehlerverursachung.² Für eine zielkonforme Steuerung und Überprüfung der Prognose- und Dispositionsgüte kann ein deskriptiver Fehler somit nicht allein angewendet werden, da er relevante Aspekte der Entscheidungssituation von Über- und Unterbeständen vernachlässigt.³ Hierfür ist eine monetäre Bewertung der aufgetretenen Fehlmengen und Überbestände in einem einheitlichen Gütekriterium erforderlich, welche die Asymmetrie von Fehl- und Überbeständen durch die Kosten der betrieblichen Entscheidungssituation valide und reliabel abbilden. Die ex-post Bewertung der wirtschaftlichen Bedeutung unterschiedlicher Prognosefehler und Warenbestände erfolgt in symmetrischen und asymmetrischen Kostenfunktionen, die in Abschnitt 5.2.2 entwickelt und analysiert werden.

¹ In Abweichung zu ARMINGER [1999] werden eine vereinfachte Notation und alternative Bezeichnungen zur Vermeidung von Verwechslungen mit statistischen Fehlermaßen gewählt.

² Dies wird auch im grafischen Verlauf aller Fehler in Abbildung 2.17 deutlich.

³ Die einseitige Zielsetzung entweder einer Minimierung von Überbeständen oder einer Verringerung der Fehlmengen entspricht jedoch vielfach der betrieblichen Praxis des Bestandsmanagements. Eine derartige Steuerung einzelner Lagerstufen reflektiert jedoch nicht die Komplexität der Warendisposition.

2.4.2 Zusammenhang von Prognosefehler und Dispositionsgüte

Für stochastischen Bedarf bestimmt die Höhe des Sicherheitsbestandes nachhaltig die Kosten der getroffenen Dispositionsentscheidung, die durch eine Reihe von Einflussfaktoren bedingt wird. "The important drivers for the amount of safety stock are the demand lead time, the variability of demand, and the service level required." [Silver et al. 1998, S. 254]. Eine Vielzahl von Firmen hat daher die Wiederbeschaffungszeiten von Lieferanten verkürzt, etwa durch schnellere Transportmittel und kürzere Wege. Darüber hinaus heben SILVER et al. [1998, S. 254] besonders die Verbesserung der Prognosegenauigkeit als Möglichkeit zur Verringerung der Variabilität der Nachfrage und somit der Sicherheitsbestände hervor.¹

Die theoretische Bedeutung der Prognosegenauigkeit wird unmittelbar aus der Gültigkeit der Formel (2.39) bzw. (2.40) für die jeweilige Verteilungsfunktion des Bedarfs deutlich. Der Sicherheitsbestand in Höhe von $k \hat{\sigma}_Y$ wird direkt durch die Standardabweichung der Prognose-residuen bestimmt. Somit steigt der Sicherheitsbestand proportional zu einer sinkenden Prognosegenauigkeit. Das Ausmaß kann – unter Annahme einer gegebenen Verteilungsfunktion – aus den Prognosefehlern bestimmt werden.

Der Standardabweichung $\hat{\sigma}_e$ der Prognosefehler e_t entspricht angenähert der Wurzel des quadratischen Fehlers (engl. *root mean squared error, RMSE*), da

$$\text{var}(e) = \sigma_e^2 = \frac{1}{n-1} \sum_{i=1}^n (e_i - \bar{e})^2 \quad , \quad (2.59)$$

und

$$\sigma_e = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (e_i - \bar{e})^2} \quad . \quad (2.60)$$

Alternativ kann die Streuung der Prognosefehler äquivalent durch das absolute Fehlermaß der mittleren absoluten Fehler (engl. *mean absolute error, MAE*) geschätzt werden, da gilt

$$\sigma_e = 1,25 MAE_e \quad . \quad (2.61)$$

SILVER et al. [1998, S. 719 f.] geben eine formale Herleitung dieser heuristischen Regel aus der Standardabweichung der Normalverteilung (siehe auch Abschnitt 3.2.2.3). Aus der Berechnung optimaler Bestellmengen (2.39), mit

$$q^* = \hat{y}_Y + k \hat{\sigma}_Y \quad , \quad (2.62)$$

und bei gegebenen Sicherheitsfaktor k führt die Veränderung des mittleren absoluten Prognosefehlers $\Delta MAE(e_t)$ zu einer proportionalen Erhöhung des Sicherheitsbestandes und der Bestellmenge Δq_e um

$$\Delta q_e = k 1,25 \Delta MAE(e_t) \quad . \quad (2.63)$$

Ferner wird die Höhe der Prognosefehler durch die unverzerrte Schätzung des Mittelwerts der Prognosefehler \hat{e}_Y beeinflusst, der ohne systematische Verzerrungen um Null schwanken muss. Erfolgt eine systematische Über- oder Unterprognose führt dies ebenfalls zu erhöhten

¹ Dagegen kann der Servicegrad nur selten durch besseres Verständnis der Kunden oder Veränderung der Kostenparameter beeinflusst werden [Silver et al. 1998, S. 254].

bzw. verringerten Warenbeständen, was die besondere Bedeutung der Überprüfung der mittleren Prognosefehler im Sinne des ersten Momentes ihrer Verteilung hervorhebt.¹

Der Einfluss der Prognosefehler auf die Bestandshöhe erscheint somit erheblich. Nach MINNER [2000, S. 13-14] erfordert die Bedeutung der Prognosefehler die Berücksichtigung exogener Marketing-Informationen in kausalen Prognoseverfahren und die korrekte Auswahl von Prognosemodellen. Sofern exogene Faktoren den Bedarf wesentlich beeinflussen müssen diese in der Bedarfsvorhersage berücksichtigt werden. Neben Aktivitäten des Wettbewerbs, der gesamtwirtschaftlichen Situation und der Phase im Lebenszyklus eines Produktes – insbesondere für neue Produkte und jene in der Reife- und Verfallsphase – hebt MINNER die Berücksichtigung von Marketingaktivitäten in Form von Preisänderungen, Werbemaßnahmen und Promotionen in der Warendisposition hervor, da diese Aktivitäten versagen wenn sie nicht hinsichtlich einer zeitgleichen Verfügbarkeit der Warenbestände koordiniert werden [Minner 2000, S. 14].

Die Betrachtung der mengenbezogenen Auswirkungen lässt jedoch keine grundsätzlichen Aussagen über die resultierenden Kosten zu. Ein Anstiegen der Prognosefehler und der Sicherheitsbestände erhöht gleichermaßen die durchschnittliche Bestandshöhe, verringert jedoch auch – bei konstanter Varianz des Bedarfs – das Risiko von Fehlmengen. Die Beurteilung des kombinierten Effektes hängt somit von dem Verhältnis der Lagerungskosten zu den Fehlmengenkosten ab, die üblicherweise nicht symmetrisch sind. Somit ist der tatsächliche Kosteneffekt in Abhängigkeit von der Asymmetrie der Kosten situativ im Einzelfall zu bestimmen.

Trotz der erheblichen Bedeutung der Prognosegenauigkeit für die Dispositions- und Lagerkosten existieren nur wenige Untersuchungen zu den empirischen Auswirkungen der Prognosegenauigkeit bei stochastischem stationärem oder instationärem Bedarf auf die Mengen und Zeitpunkte der Bestellmengenrechnung und somit auf die Kosten der Warendisposition. Nach den frühen Arbeiten durch BROWN [1956; 1959; 1963] untersuchen FILDES und BEARD {, 1991, sfc-821d23} die Auswirkungen von Prognoseverfahren und -systemen auf die Warenbestände. GARDNER [1990] analysiert den Zusammenhang von Prognosefehler und Bestandsmengen für unterschiedliche Servicegrade unter Entwicklung von Sensitivitätsanalysen. Die Studien beschränken sich jedoch auf eine Untersuchung der resultierenden Mengen, nicht aber der tatsächlichen Kosten aus den Prognoseverfahren. NADDOR [1971, S. 173] untersucht im Fall des stationären Bedarfs den Einfluss der Varianz einer Verteilungsfunktion auf die Dispositionskosten. Darüber hinaus existieren kaum Studien der Interdependenzen von Prognosegenauigkeit und Dispositionsgüte. Vielmehr beschränkt sich die Forschung im Bereich der Prognose auf die empirische Prognosegenauigkeit von Verfahren und Anwendungen losgelöst von einer nachgelagerten Entscheidung, während die umfassende Forschung im Bereich der Bestellmengenpolitiken die Problemstellungen der Prognose stochastischer, instationärer Bedarfsverläufe ignoriert und Modelle auf der Basis stationärer Verteilungsannahmen entwickelt. Die vergleichsweise geringe Betrachtung der Interaktion von Prognosefehler und Dispositionskosten in der Literatur wird von STIJBOSCH und MOORS [2005] in ihrer Untersuchung des Einflusses unbekannter Bedarfsverläufe auf die Höhe der Warenbestände zusammengefasst. Die nachfolgende Untersuchung von Neuronalen Netzen zur Prognose und Disposition wird daher auf die Analyse der Prognosegenauigkeit, der Bestandsmengen und der resultierenden Dispositionskosten erweitert.

¹ Die Höhe der Warenbestände beeinflusst über den mittleren Warenbestand und somit die Höhe der mittleren Lagerungskosten.

3 Methodische Grundlagen der Prognose

Die Erstellung von Prognosen in den verschiedensten Bereichen ökonomischer Forschung und Anwendung hat vielfältige Modelle, Verfahren und Methodologien hervorgebracht, um aus der „Analyse vergangener Ereignisse Gesetzmäßigkeiten zu entwickeln, die einen Schluss auf zukünftige Ereignisse ermöglichen“ [Hansmann 1979, S. 7]. Die Diskussion von Neuronalen Netzen zur Bedarfsprognose im Zuge der Warendisposition verlangt ihre Einordnung in die Erkenntnisse betrieblicher Prognoseverfahren und eine Erörterung im Kontext der methodischen Grundlagen der Modelle und Verfahren.

Im Folgenden werden zunächst die Determinanten objektiver Prognosemodelle analysiert, und im Kontext der betriebswirtschaftlichen Forschung und Anwendung diskutiert.¹ Die separate Untersuchung von Modell und Verfahren erscheint sinnvoll, um in späteren Kapiteln die Verfahren der Neuronalen Netze den entsprechenden Erklärungsmodellen der Prognose und konkurrierenden Verfahrensklassen zuzuordnen. Anschließend werden etablierte statistische Prognoseverfahren für die Vorhersage von stationärem und instationärem Bedarf nach der Struktur der Erklärungsmodelle und den berücksichtigten Zeitreihenkomponenten nach trendbehafteten, saisonalen und trend-saisonalen Verlauf kategorisiert, um die Komplexität der Modellselektion in Abhängigkeit der Bedarfsform herauszustellen.² Eine kurze Erörterung kausalanalytischer Verfahren von Interventionsmodellen beschließt die Untersuchung der Prognoseverfahren.

Anschließend werden statische Beurteilungsmaße der Prognosefehler sowie die Erkenntnisse zur Durchführung objektiver Verfahrensvergleiche erörtert, da diese die Durchführung der Vergleichsstudie von Neuronalen Netzen mit konkurrierenden Prognoseverfahren in Kapitel 6 nachhaltig bestimmen. Dieses Kapitel schließt mit einer prozessorientierten Analyse der einzelnen Schritte zur Durchführung von Bedarfsprognosen im Zuge der Warendisposition, die gleichermaßen auf statistische Verfahren als auch auf die Prognoseverfahren der Neuronalen Netze zutreffen.

3.1 Determinanten der Prognosemodelle

3.1.1 Merkmale der Prognose

Der Begriff „Prognose“, aus dem Griechischen wörtlich „das Vorwissen“ oder „die Vorauskenntnis“ [Kluge 2002], bezeichnet eine Aussage über ein zukünftiges, unbekanntes Ereignis, einen Zustand oder eine Entwicklung: „A forecast or prediction ... is defined as a statement concerning unknown, in particular, future, events“ [Theil 1966].

Während die Begriffe „forecast“ und „prediction“ im Englischen vielfach synonym verwendet werden [Hüttner 1986], erfolgt in dieser Untersuchung insofern eine Abgrenzung der Prognose von allgemeinen Vorhersagen, als eine Prognose stets die Formulierung von Eintrittswahrscheinlichkeiten der Ereignisse im Sinne einer Entscheidung bei Unsicherheit zulässt (siehe auch Abschnitt 2.2.1). Wenn über das zukünftige Eintreffen eines Ereignisses vollkommene Voraussicht herrscht und sein Eintreffen im Sinne einer Eintrittswahrscheinlichkeit p mit $p = 1$ mit Sicherheit behauptet werden kann, so liegt eine deterministische Vorhersage für eine Entscheidung unter Sicherheit vor – und damit keine Prognose [Brock-

¹ Die Aufstellung erhebt keinen Anspruch auf Vollständigkeit oder Überschneidungsfreiheit, insbesondere gegenüber abweichenden Taxonomien. Sie ist insbesondere auf solche Modelle und Verfahren der Prognose beschränkt, die zur betrieblichen Warendisposition angewendet werden können.

² Die Kategorisierung von Prognoseverfahren erfolgt üblicherweise zunächst nach den Verfahrensklassen, beispielsweise der Naiven Verfahren, Exponentiellen Glättung und der Autoregressiven Moving Average (ARIMA) Verfahren [Hansmann 1983; Makridakis et al. 1998; Scheer 1983].

hoff 1977, S. 17]. Wären hingegen keinerlei Informationen über das Eintreten zukünftige Ereignisse ableitbar, „so hätte dies zur Folge, dass auch die Wahlakte der Wirtschaftssubjekte nicht geordnet werden könnten“ [Brockhoff 1977, S. 17], es läge damit eine Vorhersage zur Entscheidung bei Ungewissheit vor, keine Prognose. Kann das zukünftige Eintreffen eines Ereignisses hingegen mit einer bestimmten Eintrittswahrscheinlichkeit p , mit $0 \leq p < 1$, behauptet werden, liegt nach BROCKHOFF [1977, S. 17] eine stochastische Prognose zur Entscheidung unter Risiko vor, bei der die Möglichkeit der Falsifizierung besteht. FULDA et al. erweitern diese Sichtweise: „Selbst in einer deterministischen Welt ist die Voraussicht wegen der schweren Erklärbarkeit relevanter Zusammenhänge und Daten nur begrenzt möglich“ [Fulda et al. 1989]. Somit können selbst in deterministischen Entscheidungssituationen wie der deterministischen Bedarfsrechnung stochastische Prognosen einzelner Umweltzustände erforderlich werden.

Im Gegensatz zu Vorhersagen erfolgt eine Prognose nach THEIL [1966] – neben einigen historischen Beobachtungen der zu prognostizierenden Variablen sowie verwandter Erklärungsvariablen – immer auch auf der Basis einer rationalen Theorie: „There is always some *theory*, however naive it may be. ... There are always some *observations*, however few and untrustworthy these may be“ [Theil 1966, S. 1-3]. Dabei genügen bereits die naive Annahme einer unveränderlichen Situation als Theorie und eine einzige Beobachtung als Datenbasis [Theil 1966, S. 1-3].¹ Eine Prognose ist aufgrund der Analyse der Vergangenheitsdaten, auch bei Verwendung unsystematischer Beobachtungen mit geringer Validität und Reliabilität, hinreichend empirisch fundiert [Hansmann 1979, S. 11; Brockhoff 1977, S. 17].² Ferner beruht jede realtheoretische fundierte Prognose, auch ohne explizite Theorie- oder Hypothesenbildung, auf einer sachlogischen Begründung der Beobachtungen unter Annahme von Prämissen und Randbedingungen, im Sinne einer progressiven Deduktion [Kosiol 1975, S. 46]. Der geringe Grad an analytischer Absicherung erlaubt eine Interpretation sowohl subjektiver, intuitiver Vorhersagen von Experten unter Zuhilfenahme nicht-formaler Modelle als auch objektiver Vorhersagen auf der Basis mathematisch-statistischer Ansätze unter Bildung formaler Modelle als Prognosen.³

Die Prognose auf der Basis einer rationalen Theorie erfordert deren Gültigkeit in Vergangenheit und Zukunft. Die Identität aller Grundstrukturen und Einflussfaktoren über den gesamten Beobachtungs- und Prognosezeitraum kann jedoch in der Realität nie vollständig erfüllt werden [Hansmann 1979; Stickel 1997, S. 556]. WILD formuliert hierzu eine pragmatische Zeitstabilitätshypothese: „Bei ... Rückgriff auf Vergangenheitserfahrungen wird unterstellt, daß das Ursachensystem, das die in der Vergangenheit beobachteten Sachverhalte ‚produziert‘ hat, sich zumindest bis zum Zeitpunkt, auf den sich die Prognose bezieht, nicht wesentlich ändern wird“ [Wild 1974, S. 93].⁴ Die Zeitstabilitätshypothese der Konstanz eines sich in der Vergangenheit abzeichnenden Prozesses besitzt jedoch – auch unter pragmatischen Gesichts-

¹ Dies entspricht dem einfachsten objektiven Verfahren der Naiven Prognose, in dem die Prognose der Folgeperiode dem letzten Beobachtungswert entspricht (vgl. Abschnitt 3.2.3.2.1).

² In Abgrenzung dazu erfordert eine Definition des „scientific forecasting“ gleichermaßen eine explizite Formulierung von Hypothesen und die systematische Beobachtung der Vergangenheit [Arnold et al. 2004].

³ Mit dem Erfordernis einer Theorie bzw. eines stochastischen Modells für den beobachteten Prozess wird eine Abgrenzung der Prognose von der irrationalen Vorausschau (der Prophetie bzw. der Wahrsagerei) erreicht [Schlittgen und Streitberg 1987, S. IX; Brockhoff 1977, S. 17]. Die Theorie kann jedoch schwach ausgebildet sein, sodass das Fehlen einer explizit formulierten, theoretischen Gesetzmäßigkeit – als implizite Theorie unter Annahme der Konstanz der Modellparameter – im Sinne der Prognose verstanden werden kann. Die Definition der Prognose als „the projection of the past into the future“ [Brown 1959, S. 3] setzt dagegen die Prognose unmittelbar mit einer Gruppe von Verfahren der univariaten Zeitreihenanalyse gleich, welches als Definition zu kurz greift.

⁴ Sie wird auch als „Stabilitätshypothese“ oder „Zeitinvarianzhypothese“ bezeichnet. Die Allgemeingültigkeit der Hypothese wird durch das Induktionsproblem der Wissenschaftstheorie begründet: „Eine solche Annahme scheint notwendig, da andernfalls Erfahrungen praktisch irrelevant wären und kein Lernen aus der Erfahrung möglich wäre“ [Stickel 1997, S. 557].

punkten – keine unbegrenzte Gültigkeit für einen beliebig langen Zeitraum [Rosentreter 1977, S. 2]. Je weiter eine Prognose in die Zukunft reicht, desto weniger sicher ist die Konstanz der ihr zugrunde liegenden Systemstruktur und somit die Validität des formulierten Modells. „Unterstellt man, dass die Zeitstabilitätshypothese um so weniger sicher ist, desto weiter in die Zukunft hinein sie als gültig angenommen wird, und geht man weiter davon aus, dass der Informationsgehalt der Prognose abnimmt, weil ihre Präzision mit zunehmender Ausdehnung in die Zukunft hinein ceteris paribus geringer wird, so nimmt die Prognosequalität ebenfalls ab“ [Brockhoff 1977, S. 20; ebenso Granger 1980, S. 6]. Die Wahrscheinlichkeit des Zutreffens von Prognosen wird somit umso geringer, je weiter in die Zukunft prognostiziert wird [Hüttner 1986; Brockhoff 1977, S. 41]¹ und verliert ihre Gültigkeit, wenn das postulierte Modell nicht zutrifft oder sich ändert [Schlittgen und Streitberg 1987, S. 335].² Die Gültigkeit einer überprüfbar Theorie grenzt Ansätze der Prognose ferner von technologischen Voraussagen ab, die zukünftige Veränderungen der Rahmenbedingungen der zugrunde liegenden Theorie vorhersagen und als solche nicht zu prognostizieren sind [Haderer 2000, S. 3252].³

3.1.2 Determinanten von Prognosemodellen

3.1.2.1 Gegenstand der Prognose

Ökonomische Prognosen werden nicht zum Selbstzweck durchgeführt [Makridakis et al. 1980, S. 11], sondern antizipieren wirtschaftliche Tatbestände als Mittel zur verbesserten Entscheidungsfindung [Brockhoff 1977, S. 39]. Sie sind als solche in den Zusammenhang der betrieblichen Planung und das jeweilige Entscheidungsproblem (z.B. der Warendisposition) einzuordnen.

Das Fällen beschränkt rationaler Entscheidungen in der betrieblichen Planung ist ein systematischer Vorgang der Gewinnung und Verarbeitung von Informationen [Meffert 2000, S. 147], und erfordert eine möglichst genaue Vorhersage der situativen Gegebenheiten und der zukünftigen Umweltzustände aller Entscheidungsalternativen.⁴ Prognosen liefern die für das Fällen bzw. Unterlassen von Entscheidungen nötigen Informationen [Hüttner 1986, S. 19] und sind eine notwendige Voraussetzung für jeden Entscheidungsprozess im Rahmen der Unternehmensplanung auf jeder Planungsebene und für alle Planungsobjekte [Brockhoff 1977, S. 16; Rosentreter 1977, S. 9; Stickel 1997, S. 557; Makridakis et al. 1980, S. 11]. Prognosen stellen somit die informatorische Fundierung der Planung dar [Rosentreter 1977, S. 5]. Dabei hängt die Qualität einer Entscheidung wesentlich von der Qualität der ihr zugrunde liegenden Prognose ab [Rosentreter 1977, S. 5], die durch ihre Zuverlässigkeit (Reliabilität) und ihrer Genauigkeit (Validität) bestimmt wird. Reliabilität und Validität einer Prognose sind messbar, sie werden durch unterschiedliche Gütekriterien und statistische Fehlermaße erfasst und in Abschnitt 3.3 dieser Untersuchung untersucht.

Da ökonomische Prognosen auf allen Planungsstufen und in jedem Funktionsbereich eines Unternehmens durchgeführt werden, existiert eine Vielzahl möglicher Gegenstände der Prognose⁵ mit jeweils unterschiedlichen Prognosemodellen und -verfahren. Sie sind für eine Untersuchung der Absatzprognose zur Warendisposition sinnvoll einzugrenzen. Nach

¹ Bei extrem kurzfristigen Prognosehorizonten dagegen, z. B. von nur einigen Sekunden, ist die Prognose aufgrund der Identität aller Einflussgrößen und des Umfangs der bekannten Informationen oft trivial [Granger 1980, S. 6].

² Strukturbrüche der Zeitreihe falsifizieren somit die Zeitstabilitätshypothese für eine deterministisch bestimmte Modellstruktur, nicht aber wenn die Modellstruktur stochastische Änderungen der Zeitreihe durch die Form des Erklärungsmodells implizit zulässt, z. B. im Fall der Exponentiellen Glättung.

³ Technologische Voraussagen werden auch als „Technologiefolgeschätzungen“ oder „Voraussagen im engeren Sinne“ bezeichnet [Stickel 1997, S. 556].

⁴ Für Grundlagen begrenzt rationaler Entscheidungen siehe WILD [1974, S. 93].

⁵ Sie werden auch als Prognoseobjekte oder Prognoseaussage bezeichnet [Brockhoff 1977]

PREBMAR [1995] werden die funktionalen Planungsbereiche einer Unternehmung, die verbundenen Märkte sowie die makroökonomischen Rahmenbedingungen in einem Strukturmodell abgebildet, das in Abbildung 3.1 dargestellt ist:¹

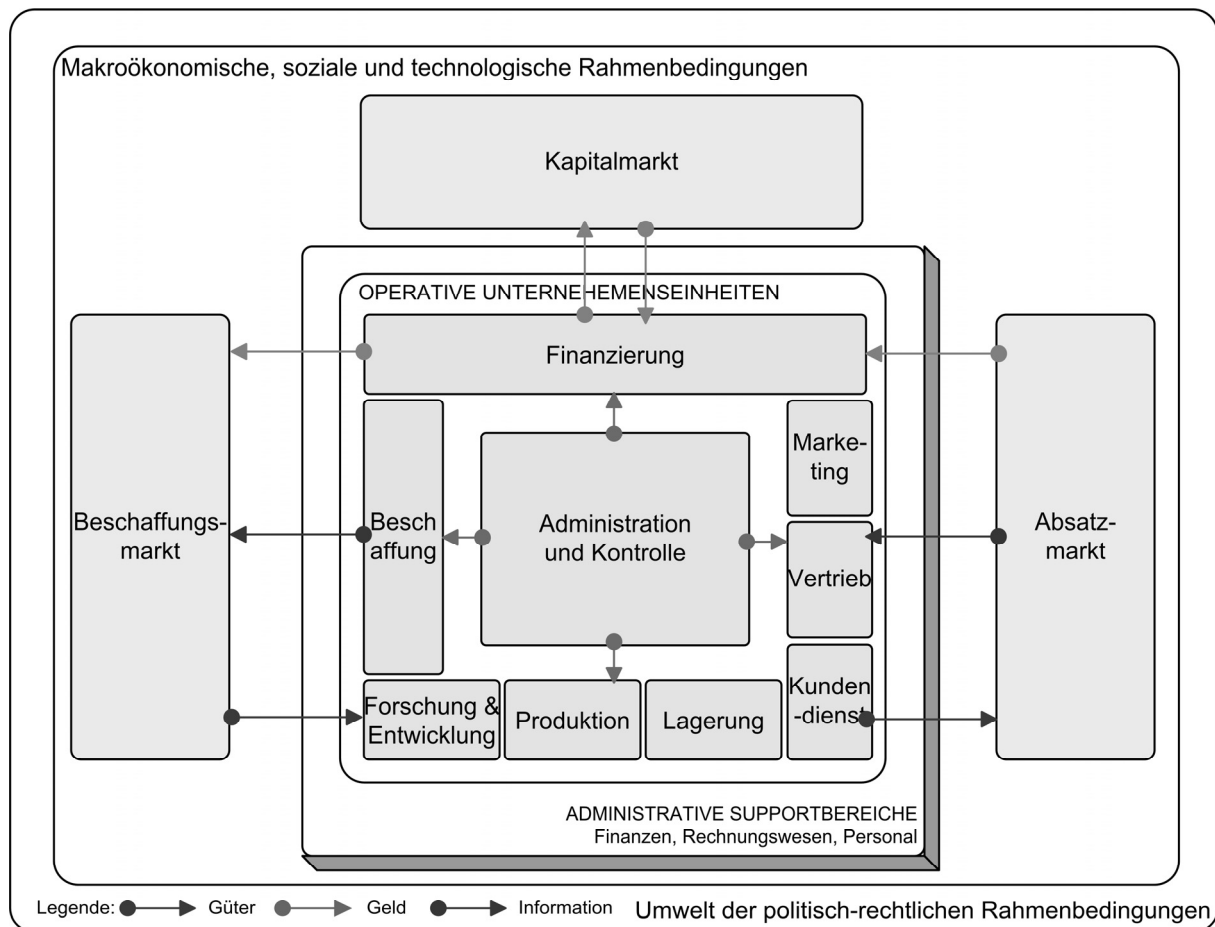


Abbildung 3.1: Strukturmodell der Einzelunternehmung²

Das Strukturmodell verdeutlicht die Interaktion eines Unternehmens mit den Beschaffungs-, Absatz- und Finanzmärkten durch Güter-, Geld- und Informationsströme, und erweitert die Betrachtung der Waren- und Informationsflüsse in Lagerstrukturen in den Abbildungen 2.3 und 2.4. Gegenstände der Prognose umfassen alle Elemente des Strukturmodells, und erlauben eine Prognose der makroökonomischen sozio-demographischen, technologischen oder politischen Rahmenbedingungen, ihrer Einflüsse auf die Märkte, der Entwicklung der Märkte und einzelner Marktteilnehmer, der Interaktion der Unternehmung mit den Märkten bis hin zur Prognose innerhalb unternehmensinterner Funktionen oder Projekte. Dabei kann der Prognosegegenstand neben den originären Güter, Geld und Informationsflüssen weitere Elemente des Strukturmodells umfassen, beispielsweise Preise, Kosten oder Servicegrade, durch Preisprognosen der Beschaffungsmärkte, Prognosen von Zinssätzen auf Kapitalmärkten oder die Prognose unternehmensinterner Kosten oder Verrechnungspreise in der Produktion. Weiter kann eine Prognose auf die Mengen und Zeitpunkte der Objekte gerichtet sein, oder auch abgeleitete Größen von Wachstumsraten oder Strukturbrüchen im Sinne von Verände-

¹ Das generische Strukturmodell kann leicht auch Handelsunternehmen erweitert werden, in dem lediglich die Funktionen von Forschung und Entwicklung sowie der Produktion fortfallen. Es dient hier jedoch einer Einordnung der Modelle und Verfahren der Bedarfsprognose in die Vielzahl möglicher Prognoseobjekte der ökonomischen Forschung unabhängig von dem Wirtschaftssektor.

² Eigene Darstellung mit formalen Veränderungen nach PREBMAR [1978; 1995]. Das Strukturmodell zeigt die funktionalen Bereiche der Aufbauorganisation in der Interaktion mit den umgebenen Märkten, in Abgrenzung zu der prozessorientierten Darstellung als Handels-H-Modell nach Becker und Schütte

rungen der typischen Ausprägung der Größen vorhersagen. Die vielfältigen Aspekte ökonomischer Prognosen erfordern somit eine Eingrenzung nach dem Prognosegegenstand.

Die Problemstellung der Bedarfsprognose zur Warendisposition erlaubt eine Spezifizierung des Prognosegegenstandes als den Bedarf an Handelswaren, wobei aufgrund der Probleme in der validen Erhebung des Bedarfs nachfolgend der Absatz synonym mit dem Begriff des Bedarfs verwendet wird (siehe Abschnitt 2.2.1.1).¹ MEFFERT [2000] charakterisiert die Absatzprognose als „eine auf die Empirie gestützte Vorhersage des zukünftigen Absatzes von Produkten einer Unternehmung an bestimmte Käuferschichten .. in einem bestimmten Zeitabschnitt und bei einer bestimmten absatzpolitischen Instrumentkombination“ [Meffert 2000, S. 165]. Die so getroffene Eingrenzung erlaubt eine Einordnung in das Strukturmodell der Unternehmung und eine Beschränkung auf die Warenflüsse des Absatzmarktes und der Absatzplanung.² Somit können die Aspekte der Vorhersage an Finanzmärkten, technologischer Innovationen in Forschung und Entwicklung sowie makroökonomischer Prognosen der gesamtwirtschaftlichen Rahmenbedingungen in der weiteren Analyse vernachlässigt werden, einschließlich der damit verbundenen Literatur und spezifischer Prognosemodelle und -verfahren.³

Gegenstände der Prognose zur Absatzplanung werden auf die Konstituenten der Absatzmärkte und den ihnen zugewandten Unternehmensfunktionen beschränkt. Da der Absatzmarkt durch das Unternehmen als Marktteilnehmer nicht direkt und nur selten indirekt beeinflusst werden kann, werden alle absatzpolitischen Entscheidungen unter Unsicherheit und unvollkommener Information getroffen und erfordern somit die Prognose einer Vielzahl von Phänomenen und Gegenständen des Absatzmarktes. „Die Vorhersage des zukünftigen Absatzes umfasst neben dem Absatzvolumen eines Produktes auch Marktpotential, Absatzpotential, Marktvolumen und Marktanteil“ [Werner 2000, S. 18]. Dabei unterscheiden sich mikroökonomische Absatzprognosen, etwa über Absatzvolumen, Marktanteil oder Absatzpotential der Unternehmung, von makroökonomischen Absatzprognosen, die sich auf Aussagen über die umgebenden Güter- und Geldmärkte auf aggregiertem Niveau konzentrieren, z. B. in Form von Prognosen über das Marktpotential, das Marktvolumen, oder das Konkurrenzverhalten von Mitbewerbern [Meffert 2000, S. 165; Gisholt 1976, S. 46].⁴

Diese Gegenstände der Absatzmärkte sind in Form von Beobachtungswerten der vergangenen Absatzwerte sowie aller den Absatz beeinflussenden Variablen in der Absatzprognose zu berücksichtigen. Die Interaktion der Objekte wird durch das wirtschaftswissenschaftliche Theoriensystem beschrieben [Hansmann 1979, S. 11] und ist als Schema nach ARMSTRONG [2001d] in Abbildung 3.2 strukturiert abgebildet:

¹ Somit werden nachfolgend auch die Begriffe der Absatzprognose und Bedarfsprognose synonym verwendet.
² In Kapitel 2 wurde der Handlungsrahmen der Warendisposition im Handel als Entscheidungsprozess mit seinen Zustandparametern (Mengen und Zeitpunkte der Kundennachfrage und der Lieferungen) sowie seinen Entscheidungsparametern (Mengen und Zeitpunkte der Bestellung) diskutiert. Demzufolge liefert die Prognose der stochastischen Nachfrage der Kunden einer Lagerstufe die Entscheidungsgrundlage für die Warendisposition. Für die Filialen des Einzelhandels stellt dies die Vorhersage des Bedarfs der Endkunden oder Konsumenten dar, der auf den Absatzmärkten entsteht.
³ Prognoseverfahren für die simultane Lösung makroökonomischer Prognosemodelle in Mehrgleichungssystemen, oder auch Modelle auf der Basis der Finanzmarkttheorie, werden folglich nicht untersucht.
⁴ Die Prognosen können sich auch auf Teilbereiche der Komponenten beziehen, beispielsweise auf einzelne Industriezweige, Regionen, oder Kundengruppen in Kombination mit Hierarchien von Produktkategorien sowie in Abgrenzung nach dem Lebenszyklus der Produkte durch neue und etablierte Waren.

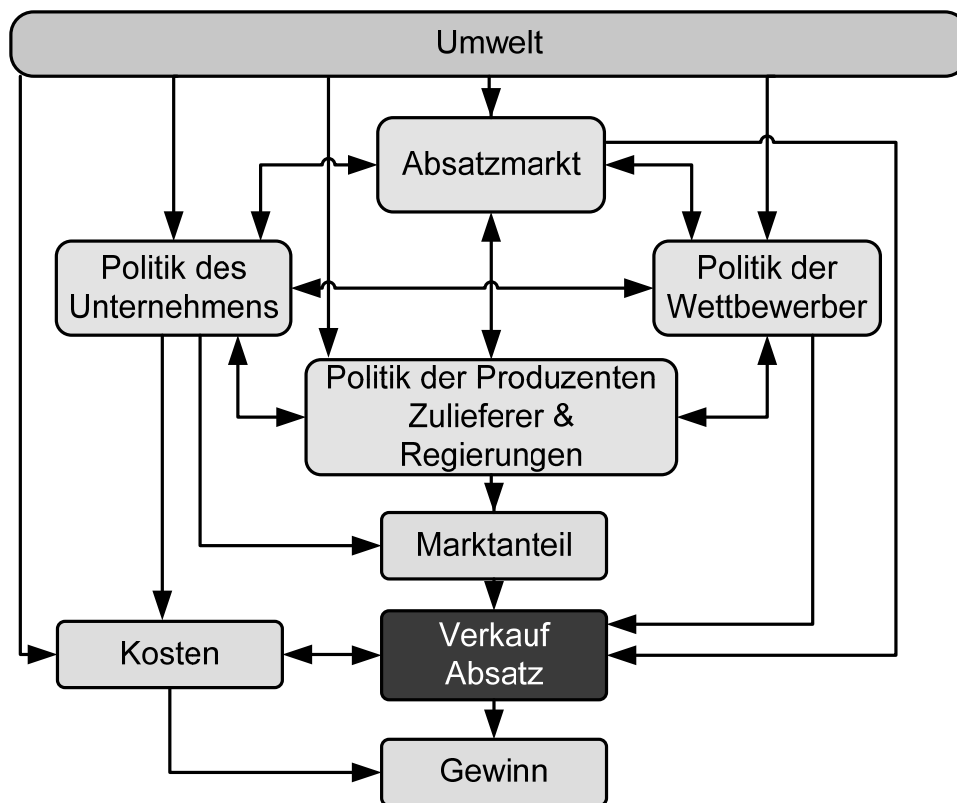


Abbildung 3.2: Prognoseobjekte der Absatzprognose und ihre Interaktion¹

Das Schema in Abbildung 3.2 identifiziert exogene, interdependente Einflüsse auf den Absatz. Im Rahmen der Marktinteraktion sind insbesondere die Preis- Produkt-, Kommunikations- und Distributionspolitiken des Unternehmens, seiner Mitbewerber im Absatzmarkt sowie seiner Zulieferer zu berücksichtigen, einschließlich der Entwicklung der betroffenen Absatzmärkte und der sie umgebenden Unternehmensumwelt. ARMSTRONG [2001d] unterscheidet ferner die verwandten Prognosegegenstände des Marktanteils, der Kosten und der Gewinne vom Absatz, die mit der Entwicklung des Absatzes interagieren und gegebenenfalls separat zu prognostizieren sind. Eine Absatzprognose zur Vorhersage der Bedarfsmengen und -zeitpunkte eines Produktes im Zuge der Warendisposition auf einem lokalen Absatzmarkt kann somit die Berücksichtigung einer Vielzahl exogener, interdependent verbundener Einflussfaktoren erfordern [Meffert 2000, S. 165]. Der Umfang und die Struktur der berücksichtigten Variablen im Erklärungsmodell der Absatzprognose bestimmt neben der Realitätsnähe und der Komplexität des Modells auch die Auswahl geeigneter Verfahren, weshalb sie nachfolgend genauer untersucht wird.

3.1.2.2 Formalisierungsgrad der Modellbildung

Als Grundlage einer Prognose wird ein implizites oder explizites Erklärungsmodell formuliert, das aus den Beobachtungen der Vergangenheit jene Einflussgrößen und diejenige Datenbasis auswählt, die den betrachteten Wirkungszusammenhang möglichst exakt erklären. Ein bestimmtes Erklärungsmodell kann anschließend durch Anwendung auf Beobachtungen der Zukunft in ein Prognosemodell überführt und schließlich angewendet werden. Je nach dem Grad der analytischen Absicherung eines Erklärungs- bzw. Prognosemodells werden subjektive und objektive Prognosemodelle unterschieden.²

¹ Erweiterte und übersetzte Darstellung nach ARMSTRONG [2001d].

² Der Einsatz objektiver Erklärungsmodelle wird durch den Einsatz statistisch-mathematischer Verfahren zur späteren Schätzung der Modelle in der Literatur vielfach als quantitative Modellierung bezeichnet, der Einsatz subjektiver Erklärungsmodelle dagegen als qualitative oder auch intuitive Modellierung [Hansmann

Erfolgt keine Aufstellung eines formalen, objektiv nachprüfbareren Erklärungsmodells oder ist die Aufstellung und Ableitung einer Prognosefunktion aufgrund der Datensituation oder der Wirkungszusammenhänge der Variablen nicht möglich, handelt es sich um ein subjektives Erklärungsmodell [Stickel 1997, S. 557]. Subjektive Modelle (engl. judgmental models) unterstellen eine nicht-formale Modellbildung durch Experten auf der Basis eines internen Modells. Diese „internen Modelle beinhalten einige in der Vergangenheit bewusst oder unbewusst erfahrene oder gespeicherte Gesetzmäßigkeiten, die sich allerdings einer intersubjektiven Nachprüfung entziehen“ [Meffert und Steffenhagen 1977, S. 34]. Die Modellbildung erfolgt hierbei intuitiv-kreativ aus Beobachtungen und Erfahrungen der Vergangenheit durch eine verbalargumentative Verknüpfung der Variablen und Wirkungsbeziehungen durch einen menschlichen Experten.

Im Fall der objektiven Modellierung erfolgt die empirische Begründung hingegen durch Aufstellung eines formalisierten Erklärungsmodells, in dem die Variablen mit Hilfe der formalen Mathematik in Gestalt eines Gleichungssystems zu einem Erklärungsmodell verknüpft werden, von dem dann eine Prognosefunktion abgeleitet wird [Hansmann 1983; Meffert und Steffenhagen 1977, S. 35]. In allgemeiner Form wird die Ausprägung eines Prognosegegenstandes y , auch als abhängige oder endogene Variable bezeichnet, durch Realisationen einer oder mehrerer unabhängiger Variablen x_i ($i = 1, \dots, I$) beschrieben, die sowohl deterministischer als auch stochastischer Art sein können.¹ „Da vollkommene Erklärungen von y in der Realität nicht gelingen, wird der ungeklärte Rest als stochastische Variable ε aufgefasst“ [Brockhoff 1977, S. 47]. In den Realisationen von ε , auch als Zufallsvariable oder unerklärter Rest bezeichnet, kommen alle nicht explizit im Modell erfassten Einflüsse zum Ausdruck. Sie können auch als Summe einer Reihe komplexer, ex ante nicht quantifizierbarer oder modellierter Phänomene interpretiert werden [Brockhoff 1977, S. 19]. So wird ein allgemeines Erklärungsmodell formuliert, mit

$$y = f(x_1, x_2, \dots, x_i) + \varepsilon \quad (3.1)$$

Die Erklärung der Variablen y durch x_i wird zur Grundlage der Prognose zukünftiger Werte. Unter Annahme der Zeitstabilitätshypothese kann das Erklärungsmodell zum Prognosemodell zukünftiger Realisationen der Variablen umgeformt werden. Die Form des Erklärungsmodells sowie die zu seiner Lösung anwendbaren Prognoseverfahren werden durch die Anzahl, das Skalenniveau, die zeitlichen Verzögerungen und die funktionale Verknüpfung der abhängigen Variable y und der unabhängigen Variablen x_i bestimmt. Aufgrund ihrer Bedeutung werden sie in den Abschnitten 3.1.2.4 ff. näher analysiert.²

1983; Scheer 1983; Makridakis et al. 1998]. HANSMANN [1979] bezeichnet subjektive bzw. qualitative Verfahren auch als heuristische Verfahren, um den Charakter zweckmäßiger, methodisch erarbeiteter Regeln und das Fehlen eines schematischen Prognosemodells herauszustellen. Da subjektive Prognosemodelle und -verfahren ebenfalls quantitative Verfahren mit mathematisch-statistischen Operationen zur Ermittlung der Prognoseergebnisse verwenden, z. B. durch Informationssammlung in Tabellenkalkulationsprogrammen, und zudem quantitative Ergebnisse in Form einer metrischen Prognoseausgabe liefern, erscheint eine Abgrenzung qualitativer versus quantitativer Modelle als wenig trennscharf. Eine Prognose, die zum selben Zeitpunkt auf der Basis identischer Daten durch andere Experten durchgeführt wird, kann in Ermangelung eines formalen Modells zu abweichenden Prognoseergebnissen führen. Zur Hervorhebung des Grades einer objektivierten Reproduzierbarkeit der Ergebnisse unabhängig von einem menschlichen Experten wird daher eine Unterscheidung in subjektive versus objektive Modelle vorgezogen.

¹ Das Erklärungsmodell ist jedoch grundsätzlich stochastischer Natur (vgl. Abschnitt 3.1.1).

² Die Unterscheidung nach dem Formalisierungsgrad der Modellierung ist für die nachfolgende Untersuchung der Neuronalen Netze als Prognoseverfahren relevant, da ein semi-parametrisches Verfahren wie die Neuronalen Netze kein formales Erklärungsmodell bilden, sondern aufgrund ihrer datengetriebenen Form die Systemzusammenhänge direkt aus den beobachteten Daten erlernen. Darüber hinaus kann das aus den Daten erlernte Erklärungs- und Prognosemodell vielfach nicht extrahiert und interpretiert werden, zumal dann, wenn es auf einer nur schwach ausgeprägten oder gar unbekanntem Theorie beruht, die in der Modellierung

Die Entscheidung für eine subjektive oder eine objektive Modellierung lässt keine Ex-ante-Rückschlüsse auf die Prognosequalität zu. Das Fehlen eines formalen Erklärungsmodells ist kein grundsätzliches Indiz für eine generelle Unterlegenheit subjektiver Prognosen in allen Entscheidungssituationen [Brockhoff 1977, S. 19]. Vielmehr wird die Anwendung objektiver, computergestützter Prognosemodelle durch die Erfordernisse zur Durchführung einer großen Anzahl von Prognosen und das Vorliegen einer ausreichenden Menge Daten bestimmt.¹ Aufgrund der Problemstellung dieser Untersuchung, die sich mit dem Vergleich von statistischen Prognoseverfahren mit NN zur Warendisposition befasst, konzentrieren sich die nachfolgenden Ausführungen auf objektive Prognosemodelle.

3.1.2.3 Gestaltung der abhängigen Modellvariablen

Das Ergebnis der Prognoserechnung ist die zur Entscheidungsfindung notwendige Information über den Prognosegegenstand y_t , und wird als Prognoseaussage bezeichnet [Brockhoff 1977, S. 47]. In objektiven Erklärungsmodellen wird sie durch eine oder mehrere abhängige Variablen des Erklärungsmodells realisiert, deren Anzahl und Skalenniveau die Modellform bestimmen.²

Die Form des Erklärungsmodells ist abhängig von der Formulierung der Prognoseaussage. BROCKHOFF [1977, S. 47] sowie WEBER [1991, S. 4] grenzen quantitative Aussagen über zukünftige Ereignisse von qualitativen Aussagen ab. Quantitative Aussagen beziehen sich auf das absolute bzw. relative Ausmaß des Prognosegegenstandes, während qualitative Aussagen nicht direkt quantifizierbar sind und „nur Art und Richtung der Entwicklung ökonomischer Variablen“ nennen [Hadelar 2000, S. 2687]. Qualitative Aussagen können als Vorhersage einer binären Klassenzugehörigkeit durch eine „Stark-oder-schwach“- bzw. eine „Steigt-oder-fällt“-Aussage erfolgen [Brockhoff 1977, S. 47]. Aufgrund der mehrdeutigen Bezeichnung qualitativer bzw. quantitativer Attribute werden strukturprüfende multivariate Analyseverfahren in Anlehnung an BACKHAUS et al. [1994, S. XVII] nach dem Skalenniveau der abhängigen Prognosevariablen y_t unterschieden.³ In Tabelle 3.1 ist eine Übersicht der zugehörigen Modelle in Abhängigkeit von der Skalierung der abhängigen und unabhängigen Variablen dargestellt.

Tabelle 3.1: Kategorisierung statistischer Modelle nach dem Skalenniveau der Variablen

	Skalierung der abhängigen Variable	Skalierung der unabhängigen Variablen		
		metrisch	ordinal	nominal
abhängige Variable bekannt	metrisch	Regressionsanalyse	↻	Varianzanalyse
	ordinal	⊙	↻	⊙
	nominal	Diskriminanzanalyse	↻	Kontingenzanalyse
abhängige Variable unbekannt	keine	Hauptkomponentenanalyse	↻	Clusteranalyse

⊙, ↻ = Richtung einer zulässigen Skalierung der Werte auf ein niedrigeres Skalenniveau zur Prognoserechnung

nicht immer evident wird. Aufgrund der Modellierung eines formalen, wenn auch schwer verständlichen Erklärungsmodells werden die Neuronale Netze somit den objektiven Prognosemodellen zugeordnet.

¹ Eine große Menge an Zeitreihen kann nicht valide und reliabel durch subjektive Prognosemodelle erfolgen, da die subjektive Prognoseerstellung durch Experten nicht automatisierbar ist und somit für große Datenmengen durch den Zeit- und Kostenaufwand weniger effizient als objektive Modelle ist.

² Eine Variation des Skalenniveaus bei mehreren abhängigen, endogenen Variablen ist insbesondere im Kontext der Modellierung von Neuronale Netzen denkbar, wird jedoch nicht durch typische Verfahren abgebildet und somit vernachlässigt.

³ Die Unterscheidung zwischen Prognose und Klassifikation erfolgt somit nicht hinsichtlich der Zukunftsbezogenheit der Aussage, sondern anhand des Skalenniveaus der abhängigen Variablen. Nachfolgend wird der Begriff der Prognose einheitlich zur Identifikation regressionsanalytischer Ansätze der Punktprognose und der Begriff der Klassifikation zur Abgrenzung nominal skalierten Vorhersagen verwendet, unter Abgrenzung nach dem Skalenniveau der abhängigen Variablen.

Ansätze der Regressionsanalyse modellieren die funktionalen Abhängigkeiten zwischen einer metrisch skalierten abhängigen Variable und einer oder mehreren unabhängigen Variablen beliebiger Skalierung, wobei im Fall der Prognose die abhängige Variable eine zukünftige Realisation y_{t+h} in einer zukünftigen Zeitperiode $t+h$ bezeichnet.¹ Die Formulierung der Prognose als Diskriminanzanalyse zur Klassifizierung oder Einteilung von Objekten anhand bestimmter Merkmale erfolgt dagegen durch Ausgabe einer nominal skalierten, anhängigen Variablen, die eine Klassenzugehörigkeit des Prognosegegenstandes zu einem Zeitpunkt z. B. in Form der Klassen „steigt“ oder „fällt“ durch $\{0; 1\}$ codiert und vorhersagt [Backhaus et al. 1994, S. XVII].² Sind die Ausprägungen der vorherzusagenden abhängigen Variablen im Sinne einer unbekanntem Zugehörigkeit zu einer Klasse hingegen unbekannt, liegen Verfahren der Hauptkomponentenanalyse sowie der Clusteranalyse vor, die nicht zur Prognose im Sinne einer Vorhersage von Realisationen in der Zukunft angewendet werden.³ Für Anwendungen der Warendisposition ist ausschließlich die regressionsanalytische Modellierung unter Annahme metrisch skalierten abhängiger Prognosevariablen relevant, da eine Berechnung von Bestellmengen nicht anhand von nominalen Prognoseaussagen erfolgen kann. Auf eine Untersuchung der anderen Verfahren wird entsprechend verzichtet.

Prognosen können in unterschiedlichen Frequenzen einmalig oder mehrmalig, in einer frei wählbaren oder feststehenden Sequenz und mit unterschiedlichen zeitlichen Horizonten erarbeitet werden [Weber 1991, S. 4]. Jede Prognose wird zu einem spezifischen Zeitpunkt der Berechnung t ($t = 1, \dots, T$) erstellt, dem sogenannten Prognoseursprung, und für eine zukünftige Periode $t+h$ ($h = 1, \dots, H$) berechnet. Der Prognosehorizont h bezeichnet den Zeitraum zwischen dem Zeitpunkt der Prognoseerstellung t und dem Zeitpunkt der Entscheidung $t+h$ [Brockhoff 1977, S. 40; Meffert und Steffenhagen 1977, S. 37].⁴ Wird für eine abhängige Variable y_t im Zeitpunkt t der zukünftige Wert \hat{y}_{t+h} der unmittelbar nachfolgenden Periode $t+1$, mit $h = 1$, prognostiziert, erfolgt eine Einschrittprognose (engl. *one-step forecast*) [Granger 1980, S. 8]. Eine Sequenz von Punktprognosen wird als Mehrschrittprognose (engl. *trace forecasts*) bezeichnet [Granger 1980, S. 37]. Sie wird zu einem Zeitpunkt t für mehrere Prognosehorizonte h_1, h_2, \dots, h_H erstellt und erlaubt eine zeitraumbezogene, dynamische Betrachtung des Prognosegegenstandes im Zeitverlauf. Mehrschrittprognosen können je nach Prognoseverfahren als iterative Einschrittprognosen unter iterativer Verwendung von Beobachtungen und Prognosewerten vorangegangener Perioden erfolgen, mit

$$\begin{aligned} \hat{y}_{t+1} &= f(y_t, y_{t-1}, \dots, y_{t-l}) \\ \hat{y}_{t+2} &= f(\hat{y}_{t+1}, y_t, \dots, y_{t-l+1}) \\ &\dots \\ \hat{y}_{t+H} &= f(\hat{y}_{t+H-1}, \hat{y}_{t+H-2}, \dots, y_{t-l+H}) \end{aligned} \quad (3.2)$$

¹ Hiervon sind Ansätze der Diagnose zu unterscheiden, in denen die abhängige Variable vergangene oder gegenwärtige Zeitpunkte bezeichnet, nicht aber zur Vorhersage zukünftiger Realisationen verwendet wird.

² Die Systematik nach BACKHAUS et al. [2003] vernachlässigt allerdings den Sonderfall einer Klassifikation durch die Vorhersage der Wahrscheinlichkeit der Klassenzugehörigkeit von neuen Beobachtungen $0 \leq p \leq 1$ durch eine oder mehrere metrisch skalierte, abhängige Variable. Obwohl die abhängige Variable nicht nominal, sondern metrisch skaliert ist wird sie dennoch als Anwendung der Klassifikation bezeichnet.

³ Diese Klassifikation korrespondiert mit der Unterscheidung in Problemstellungen des überwachten sowie unüberwachten Lernens der Verfahren der Neuronalen Netze (vgl. Abschnitt 4.3), die nachfolgend erneut aufgegriffen wird.

⁴ In der betrieblichen Planung sind vielfach für einen Prognoseursprung mehrere Prognosen für unterschiedliche, prominente Prognosezeitpunkte der operativen und taktischen Planung und Entscheidung zu erstellen. So finden sich im Handel bereits vielfache operative Planungshorizonte, z. B. in Form der täglichen Prognose von 5 bis 12 Tagen zum Bestandsmanagement, 10 bis 28 Tagen zur Personaleinsatzplanung sowie 2 bis 6 Wochen zur Planung der Distributionskapazitäten.

oder sie können als direkte Mehrschrittprognose ohne iterative Verwendung vorheriger Prognosewerte erfolgen:

$$\begin{aligned}\hat{y}_{t+1} &= f(y_t, y_{t-1}, \dots, y_{t-l}) \\ \hat{y}_{t+2} &= f(y_t, y_{t-1}, \dots, y_{t-l}) \\ &\dots \\ \hat{y}_{t+H} &= f(y_t, y_{t-1}, \dots, y_{t-l})\end{aligned}\tag{3.3}$$

Sofern nicht anders angegeben, beziehen sich die nachfolgenden Ausführungen auf den einfachen Fall einer Einschnitt-Prognose für $t+1$ mit $h = 1$.

Prognoseaussagen können sich ferner auf einzelne oder mehrere Variablen beziehen [Granger 1980, S. 10]. Bezieht sich die Aussage auf eine einzige ökonomische Variable, so spricht man von einer Einzelprognose. Dagegen bezieht sich eine multiple Prognose oder ein „Prognose-system ... auf eine Gesamtheit von Variablen, die in ihrer gegenseitigen Verknüpfung prognostiziert werden“ [Brockhoff 1977, S. 36; Meffert und Steffenhagen 1977, S. 35].¹ Im Fall der Absatzprognose $\hat{y}_{j,t+h}$ ($j = 1, \dots, J$) zur Disposition von j Artikeln zu einem Zeitpunkt handelt es sich – unter Vernachlässigung von Verbundeffekten – entsprechend nicht um eine multiple Prognose, sondern um j Einzelprognosen.² Modelle und Verfahren für interdependente Mehrgleichungssysteme werden nachfolgend nicht berücksichtigt.

Im Fall der Regression kann ferner zwischen Punktprognosen, Intervallprognosen und Verteilungsprognosen unterschieden werden. Bei Punktprognosen wird für eine ökonomische Variable y_t im Zeitpunkt t der zukünftige Wert y_{t+h} im Zeitpunkt $t+h$ gesucht.³ Ein „point forecast ... is a single guess for y_{t+h} that in some way well represents the whole distribution of possible values“ [Granger 1980, S. 10]. Dabei werden üblicherweise die zentralen Lageparameter von Mittelwert oder Median der Verteilung aller vergangenen Absatzwerte verwendet. Im Fall der Vorhersage eines spezifischen Zeitpunktes $t+h$ einer Reihe von Beobachtungswerten, z. B. eines Wendepunktes, existiert ein besonderer Fall der Punktprognose mit der Zeit als abhängigen Variablen, der nicht als Punktprognose im eigentlichen Sinne zu verstehen ist.⁴

Können bei der Angabe der Ergebnisse hingegen auch Vertrauens- bzw. Konfidenzintervalle angegeben werden, innerhalb dessen der zukünftige Wert mit einer gewissen Sicherheit oder Wahrscheinlichkeit p erwartet wird, mit

$$p(B < y_{t+h} < A) = 0.95 \quad ,\tag{3.4}$$

liegen Intervall- oder Bereichsprognosen vor [Granger 1980, S. 10]. Eine Intervallprognose erfolgt unter Bestimmung mehrerer abhängiger Variablen $y_{j,t}$ ($j = 1, \dots, J$) die j Quantile A, B, \dots der Verteilungsfunktion für einen Prognosezeitpunkt $t+h$ spezifizieren. Die Quantile

¹ Der Begriff „Prognosesystem“, bei dem das zugrunde liegende Erklärungsmodell durch ein System von Gleichungen abgebildet wird, ist von der Definition eines computergestützten Prognosesystems als Kombination unterschiedlicher Modelle und Verfahren in einem Softwaresystem zu unterscheiden.

² Entsprechend wird im Folgenden auf den Index j als Notation für unterschiedliche Artikel verzichtet.

³ Eine einheitliche Notation der Symbole, Variablen und Indizes in der Literatur zur Prognose existiert nicht. Für die vorliegende Arbeit wird eine Notation in Anlehnung an die betriebswirtschaftliche Prognoseliteratur nach BROCKHOFF [1977] und SCHEER [1983] verwendet. Inkonsistenzen werden, soweit möglich, angepasst und Originalnotationen auch in Zitaten weitestgehend harmonisiert.

⁴ GRANGER [1980] bezeichnet die Vorhersage von Zeitpunkten als „Event Timing Forecasts“ auf der Basis von Frühwarn-Indikatoren (engl. *leading indicators*), in Abgrenzung zu Zeitreihenprognosen (engl. *time series forecasts*) [ebenso Brockhoff 1977, S. 48].

werden vielfach symmetrisch um einen zentralen Lageparameter gewählt, und auch als Mehrschrittprognose zur Abbildung von Konfidenzintervallen erweitert. Wird hingegen nur ein einseitiges Quantil einer stetigen Verteilungsfunktion angegeben, z. B. als

$$p(y_{t+h} < A) = 0.975 \quad , \quad (3.5)$$

liegt eine Punktprognose vor, wobei anstelle des Mittelwertes lediglich ein alternatives Quantil der gesamten Wahrscheinlichkeitsverteilung angegeben wird. Man spricht auch von einer Quantil-Regression (engl. quantile regression) [KOENKER UND BASSETT 1978; KOENKER UND DOREY 1987], die in der betriebswirtschaftlichen Theorie und Praxis jedoch kaum Beachtung findet [Koenker 2005].

Als Erweiterung von Punkt- und Intervallprognosen nennt Granger [1980] die Verteilungsprognose, in der die vollständige bedingte Dichte- bzw. Verteilungsfunktion der abhängigen Variablen in der Vorhersage spezifiziert wird. Eine Prognose der Verteilungsfunktion ist theoretisch vorzuziehen, da sie alle denkbaren Punktprognosen des Mittelwertes, Medians oder einzelner Quantile und darüber hinaus alle symmetrischen und asymmetrischen Intervallprognosen umfasst. Aufgrund der hohen Komplexität ihrer Berechnung wird sie jedoch nur vereinzelt untersucht [Granger 1980]. Für die Problemstellung der Warendisposition ist zu berücksichtigen, dass eine Punktprognose keine Aussagen über die Sicherheit der getroffenen Prognose im Sinne einer Eintrittswahrscheinlichkeit oder Aussagen über die Verteilungsform der Prognosewerte erlaubt. Punktprognosen besitzen somit weniger Informationen als Intervall- oder Verteilungsprognosen, werden aber in Theorie und Praxis weit mehr beachtet als komplexere Prognoseformen [Brockhoff 1977, S. 47 f.].¹ Die nachfolgend diskutierten statistischen Prognoseverfahren bestimmen grundsätzlich Punktprognosen der zentralen Lageparameter. Im Gegensatz dazu berechnet das in Abschnitt 5 entwickelte und in Abschnitt 6 beurteilte Neuronale Netz unter Verwendung asymmetrischer Kostenfunktionen eine Punktprognose eines bestimmten Quantils der Verteilungsfunktion im Sinne der Quantil-Regression.

3.1.2.4 Gestaltung der unabhängigen Modellvariablen

Bei objektiver Modellierung können die formalen Erklärungsmodelle je nach Umfang der verwendeten unabhängigen Variablen in statische versus dynamische, kausale und zeitreihenanalytische sowie lineare und nichtlineare Modelle unterschieden werden. Die Variablen bestimmen durch die Form des Erklärungsmodells die zu ihrer Schätzung anzuwendenden Prognoseverfahren, die in Abschnitt 3.1.3.3 einander zugeordnet und diskutiert werden.

Beziehen sich sowohl alle abhängigen Variablen y_t und alle unabhängigen Variablen $x_{i,t}$ des Erklärungsmodells auf dieselbe Zeitperiode t , wird ein statisches Modell formuliert.² Werden hingegen zeitlich verzögerte Zusammenhänge zwischen der abhängigen Variablen y_t im Zeitpunkt t und den unabhängigen Variablen $x_{i,t-l}$ in vorherigen Perioden $t-l$ ($l = 1, 2, \dots, L$) oder auch zeitlich verzögerte Realisationen der abhängigen Variablen selbst y_{t-l} abgebildet, entsteht ein dynamisches Modell [Scheer 1983, S. 8]. Die zeitliche Verzögerung (engl. lag) einer Variablen $x_{i,t-l}$ wird durch l spezifiziert, und die Variable entsprechend als „Lag-Variable“ oder „prä-determinierte Variable“ bezeichnet [Brockhoff 1977, S. 36]. Das dynamische Erklärungsmodell für Realisationen der abhängigen Variablen y_t lautet in allgemeiner Form

¹ Die Bedeutung der Verteilungsform ist besonders im Fall der Bedarfsprognose zur Bestellmengenrechnung in der Warendisposition gegeben, da die Bestimmung eines Quantils der Verteilungsfunktion durch die Intervallprognose einer Berechnung der Warenbestände durch das Newsboy-Modell entspricht.

² Nachfolgend aus Gründen der Übersichtlichkeit vereinfachend von einer abhängigen Variablen ausgegangen.

$$y_t = f\left(\left(x_{1,t}, x_{1,t-1}, \dots, x_{1,t-l}\right), \dots, \left(x_{i,t}, x_{i,t-1}, \dots, x_{i,t-l}\right), \left(y_{t-1}, y_{t-2}, \dots, y_{t-l}\right)\right) + \varepsilon_t. \quad (3.6)$$

Die Erklärung der Variablen y_t durch zeitgleiche Realisationen von $x_{i,t}$ sowie verzögerte Realisationen $x_{i,t-l}, y_{t-l}$ erfolgt im Erklärungszeitraum $t = 1, \dots, n$. Der Zeitpunkt $t = n$ des letzten historischen Beobachtungswertes¹ wird zur Grundlage der Prognose zukünftiger Werte. Unter Annahme der Zeitstabilitätshypothese (vgl. Abschnitt 3.1.1) kann das Erklärungsmodell zum Prognosemodell zukünftiger Realisationen der Variablen \hat{y}_{t+h} für beliebige Zeitpunkte $t+h$ im Prognosezeitraum ($h = 1, \dots, H$) umgeformt werden², mit

$$\hat{y}_{t+h} = f\left(\left(x_{1,t+h}, x_{1,t+h-1}, \dots, x_{1,t+h-l}\right), \dots, \left(x_{i,t+h}, x_{i,t+h-1}, \dots, x_{i,t+h-l}\right), \dots, \left(y_t, y_{t-1}, \dots, y_{t-l}\right)\right). \quad (3.7)$$

Kausale Erklärungs- (3.6) bzw. Prognosemodelle (3.7) berücksichtigen dementsprechend zeitgleiche sowie zeitlich verzögerte Realisationen der unabhängigen Variablen x_i sowie der abhängigen Variablen $y_{i,t-l}$ zur Modellierung kausaler Ursache-Wirkungszusammenhänge. Der Zusammenhang von Erklärungs- und Prognosezeitraum ist in Abbildung 3.3 dargestellt.

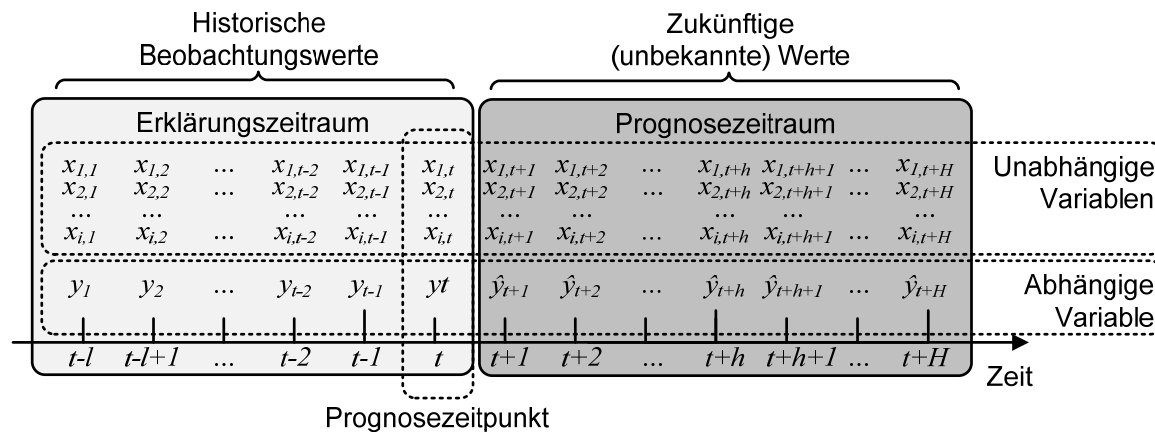


Abbildung 3.3: Modellbildung im Erklärungs- und Prognosezeitraum³

Der Beobachtungs- bzw. Erklärungszeitraum $t = 1, \dots, n$ bildet die Datengrundlage und umfasst alle historischen Beobachtungswerte.⁴ Die Auswahl der Vergangenheitswerte der abhängigen Variablen y_t und der unabhängigen Variablen $x_{i,t}$ im Erklärungszeitraum ergibt die zur Modellierung verwendete Datenbasis. Die Bestimmung des Erklärungszeitraums erlaubt auch die Vernachlässigung von Beobachtungswerten früherer Perioden, die nicht in das Erklärungsmodell eingehen sollen [Brockhoff 1977, S. 36].⁵ Die Verwendung zeitlich verzögerter Realisationen der abhängigen Variable y_t und der unabhängigen Variablen $x_{i,t}$ sowie ihre funktionale Verknüpfung bestimmt die Form des Erklärungsmodells.

¹ In Anlehnung an die Notation von GRANGER [1980] mit $n := \text{„now“}$ im Sinne des Zeitpunktes „jetzt“.

² Prognostizierte Werte der abhängigen Variablen werden nach allgemeiner Notation zur Unterscheidung von vergangenen Beobachtungswerten durch ein Dach gekennzeichnet. Die Notation von Variablen mit Dach $\hat{y}_{t+h}, \hat{x}_{i,t}, \hat{\mu}, \hat{\sigma}$ symbolisiert ihre Schätzung durch stochastische Prognosemodelle.

³ Eigene Darstellung in Anlehnung an SCHEER [1983, S. 8 f.].

⁴ Häufig werden die Daten des Beobachtungszeitraums zudem in eine weitere Untermenge zur Parameteranpassung der Prognoseverfahren und eine unabhängige Menge zum Test der Generalisierungsfähigkeit der Modelle zerlegt im Sinne einer retrospektiven Ex-ante-Prognose unter Festlegung eines beliebigen Prognosezeitpunktes t' in der Vergangenheit des Erklärungszeitraums [Küsters 2006]. Der Prognoseursprung t entspricht somit nicht immer dem Zeitpunkt der Gegenwart.

⁵ Die Interpretation des Erklärungszeitraums als Zeitfenster von v Perioden ($v = 1, 2, \dots, V$) im Beobachtungszeitraum ($t = 1, \dots, T$) mit $V < T$ würde allerdings eine abweichende Notation bei späteren Formeln bedingen, weshalb vereinfachend darauf verzichtet wird.

Ein Erklärungs- oder Prognosemodell soll alle für eine Entscheidungssituation relevanten Einflüsse der makroökonomischen, sozialen und technischen Umweltdimensionen, das Verhalten anderer aktiver Systemelemente auf den Märkten sowie mikroökonomische Tatbestände umfassen, z. B. alle Einflussfaktoren und Teilentscheidungen die die Nachfrage von Konsumenten beeinflussen [Hadelar 2000, S. 2679; Stickel 1997, S. 557]. MEFFERT [1980, S. 147] fordert die Berücksichtigung folgender Elemente in Erklärungs- und Prognosemodellen:

- Entwicklung der Rahmenbedingungen (ökonomische, technische, soziale, rechtliche),
- Verhalten der Marktteilnehmer (Konkurrenz, Käufer, Serviceanbieter), und
- Wirkung der Marketinginstrumente unter wechselnden Marktbedingungen.

Eine im Marketing verbreitete statische Ausprägung der kausalen Prognosemodelle sind die sogenannten Wirkungsprognosen. „In Wirkungsprognosen dokumentieren sich die für eine spezifische Unternehmenssituation ermittelten Gesetzmäßigkeiten der Marktreaktion“ [Meffert 1980, S. 89].¹ Die Abhängigkeit einer endogenen Prognosevariablen von jeweils veränderbaren Aktivitätsniveaus einer oder mehrerer exogener Instrumentvariablen wird dabei durch eine statische Marktreaktionsfunktion zu einem gegebenen Zeitpunkt bestimmt:

$$y_t = f(x_{1,t}, x_{2,t}, \dots, x_{n,t}) + \varepsilon_t \quad (3.8)$$

Damit soll ein kausales Erklärungsmodell zur Klärung der unbekanntem substitutionalen, komplementären oder limitationalen Wirkung der Marketinginstrumente hinsichtlich der Marktreaktion dienen [Meffert und Steffenhagen 1977, S. 36]. Aufbauend auf den Erkenntnissen des Erklärungsmodells kann anschließend ein zielführendes Niveau der kontrollierbaren Instrumentvariablen in Interaktion mit anderen Variablen des Marketing-Mix festgelegt werden.² Ihre funktionale Darstellung erfolgt – bezogen auf die jeweilige Umweltsituation – als zeitlose, statische Aussage [Meffert und Steffenhagen 1977, S. 37] und repräsentiert somit einen statischen Sonderfall kausaler Modelle.

Dynamische Erklärungsmodelle der Zeitreihenanalyse verzichten auf eine Berücksichtigung kausaler Einflüsse durch die Modellierung exogener Erklärungsvariablen x_t [Hruschka 1980], z. B. bei Unkenntnis kausaler Einflussfaktoren oder ihres Wirkungsverbundes. Diese Erklärungsmodelle ermitteln Regelmäßigkeiten allein aus den vergangenen Beobachtungswerten der abhängigen Variablen $y_t, y_{t-1}, y_{t-2}, \dots, y_{t-l}$, also der Zeitreihe der zu prognostizierenden Größe selbst. Konkurrent werden sie als Modelle der Zeitreihenprognose oder als Entwicklungsprognosen bezeichnet [Meffert und Steffenhagen 1977, S. 37; Hüttner 1986].³ Das Erklärungsmodell der Zeitreihenprognose lautet

¹ Sie werden auch als Instrumental- oder Entscheidungsprognosen sowie als Responsefunktionen bezeichnet [Meffert und Steffenhagen 1977]. Die Definition der Wirkungsprognose ist jedoch von BROCKHOFFS [1977] indirekter Wirkung der Veröffentlichung der Prognose auf die Entscheidungssituation zu unterscheiden.

² Nach MEFFERT ist „die Abschätzung und Messung der Marktreaktion als Entscheidungsergebnisse bzw. -konsequenzen ... das Kernproblem der gesamten Absatzpolitik. ... Es ist die zentrale Aufgabe der Marketingtheorie, allgemeine Gesetzesaussagen über die Wirkung der Marketingaktivitäten, d. h. generelle Reaktionshypothesen, aufzustellen und empirisch zu testen“ [Meffert und Steffenhagen 1977, S. 37]. Diese Feststellung im Kontext von Entscheidungen zur Marketingpolitik besitzt in dieser Untersuchung zur Bedeutung der Absatzprognose auf die Güte der Warendisposition keine Gültigkeit, was auf die Unterschiede in der Entscheidungssituation und dem Prognosegegenstand zurückzuführen ist.

³ Sie werden irreführend auch als autonome Prognose, nicht-konditionale Prognose, Informationsprognose oder Trendprognose bezeichnet [Meffert und Steffenhagen 1977]. Auch eine Bezeichnung als formale oder univariate Verfahren [Hansmann 1979] wird nachfolgend vermieden, da kausale Verfahren mit einer einzelnen unabhängigen Variablen ebenso univariaten Charakter besitzen [Hansmann 1979, S. 12] und die Modellbildung zeitreihenanalytischer Modelle ebenso formalen Charakter hat [Scheer 1983]. Zudem erscheint die Unterscheidung nach der Anzahl erklärender Variablen in multivariate (im Sinne von kausalen) versus univariate (im Sinne von Modellen der Zeitreihenanalyse) Verfahren wenig trennscharf, da die Abbil-

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-l}) + \varepsilon_t \quad (3.9)$$

das aus dem Erklärungsmodell abgeleitete Prognosemodell lautet

$$\hat{y}_{t+h} = f(y_{t+h-1}, y_{t+h-2}, \dots, y_{t+h-l}) \quad (3.10)$$

„In Entwicklungsprognosen ... fehlt die explizite Berücksichtigung von Wirkungen des eingesetzten und in Zukunft einsetzbaren Marketing-Instrumentariums der Unternehmung bzw. aller am Markt agierenden Unternehmungen“ [Meffert 1980]. Implizit wird somit für einen spezifischen Prognosezeitraum eine Fortschreibung aller bisherigen den Markt beeinflussenden Aktivitäten gemäß der Zeitstabilitätshypothese unterstellt [Meffert und Steffenhagen 1977, S. 36]. Die Prognose absatzpolitischer Entscheidungen durch Modelle der Zeitreihenprognose wird daher von MEFFERT [1980, S. 201] als „kausal logisch unbefriedigend“ bezeichnet. In Abbildung 3.4 erfolgt eine Einordnung der kausalanalytischen und zeitreihenanalytischer Erklärungsmodelle sowie der Wirkungsprognose.

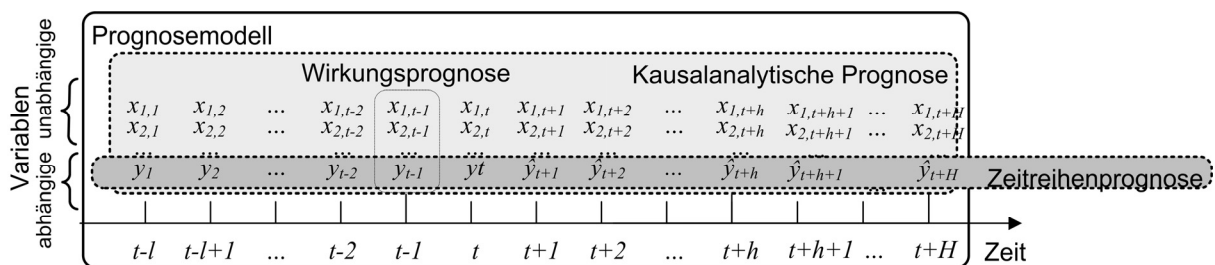


Abbildung 3.4: Prognosemodelle der Zeitreihenprognose, der Wirkungsprognose und der kausalen Prognose

Modelle der Zeitreihenprognose und der Wirkungsprognose repräsentieren somit vereinfachte Sonderfälle der kausalen Prognose. Die Betrachtung des Absatzes y_t [ME] in Abhängigkeit von einer einzelnen kausalen Variable x_t [ME], z. B. der Höhe der Werbeausgaben, und der Zeit t [ZE] erlaubt die grafische Veranschaulichung der Modelle in Abbildung 3.5.

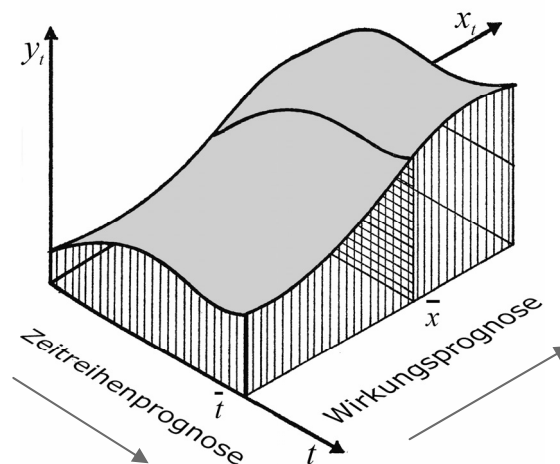


Abbildung 3.5: Abgrenzung von Zeitreihenprognose und Wirkungsprognose¹

dung mehrerer zeitlich verzögerter Variablen y_{t+h-l} gleichfalls eine multivariate Modellierung repräsentiert. Dagegen ist zu berücksichtigen, dass dynamische Regressionsmodelle der kausalen Verfahren ebenfalls auf Zeitreihen von Beobachtungswerten beruhen, die dieser Terminologie widersprechen.

¹ Erweiterte Darstellung nach MEFFERT [1980, S. 189]. Die Zeit t wird aus Gründen der Vereinfachung im als diskret angenommen, so dass jedes t in [ZE] ein Zeitintervall zwischen zwei Zeitpunkten bezeichnet.

Zeitreihenprognosen analysieren die zeitliche Entwicklung des Absatzes y_t unter Konstanz des absatzpolitischen Instrumentariums mit $y_t = f(\bar{x}_t, t)$, Wirkungsprognosen hingegen unter Variation der Absatzpolitik unter Konstanz der Zeit mit $y = f(x_t, \bar{t})$ und dynamische Modelle der kausalen Prognosen die Variation aller Variablen $y_t = f(x_t, t)$ [in Anlehnung an Meffert 1980, S. 189]. Die Veränderung der Werbeaktivitäten, z. B. von $x_t = 0$ auf $x_t = \bar{x}_t$ würde durch Verfahren der Zeitreihenanalyse als Verschiebung des Absatzniveaus im Sinne eines Strukturbruches interpretiert werden. Die Erweiterungen zeitreihenanalytischer Modelle um eine einzelne exogene Variable, z. B. zur Erklärung eines dominanten absatzpolitischen Instruments oder von Strukturbrüchen im Erklärungsmodell, werden als Interventionsmodelle bezeichnet [Makridakis et al. 1980], die einen Übergang von der zeitreihenanalytischen zur kausalanalytischen Modellierung darstellen.

Die Form der funktionalen Verknüpfung der einzelnen Variablen legt das endgültige formale Prognosemodell fest. Entsprechend den Annahmen über $f(y_i, x_i)$ kann das Erklärungs- und Prognosemodell linear oder nichtlinear sein. Ein funktionaler Zusammenhang der Form

$$f(x_i) = a + bx_i \quad (3.11)$$

heißt linear, wenn die Bedingungen der Superposition

$$f(x_1 + x_2) = f(x_1) + f(x_2) \quad (3.12)$$

und der Homogenität

$$f(ax_i) = af(x_i) \quad (3.13)$$

für alle a erfüllt sind. Danach gilt jeder funktionale Zusammenhang als nichtlinear, der nicht in jedem Bereich proportional auf das veränderte Eingangssignal a reagiert, z. B. bei Interaktion des Niveaus einer Zeitreihe mit der Schwankungsbreite einer multiplikativ verknüpften Saisonalität.¹ Man unterscheidet Modelle die linear in den Parametern sind aber nichtlinear in den Variablen des Erklärungsmodells, mit

$$f(x_i) = b_0 + b_1x_i + b_2x_i^2 + \varepsilon_i \quad (3.14)$$

sowie Modelle die linear in den Variablen sind aber nichtlinear in den Parametern:

$$f(x_i) = b_1e^{-bx_i} + \varepsilon_i \quad (3.15)$$

Obwohl die meisten in der Realität und der betriebswirtschaftlichen Praxis auftretenden Systemzusammenhänge nichtlinear sind, werden sie nahezu ausschließlich in linearen Erklärungsmodellen abgebildet, was zu erheblichem Aufwand in der Vorverarbeitung und Transformation der Daten und zu systematischen Fehlern in der Modellbildung führt.

Die Modellspezifikation wird durch die Prämissen bezüglich des Definitionsbereichs, des Wertevorrats und der stochastischer Eigenschaften der Variablen und Parameter vervollständigt [Meffert und Steffenhagen 1977, S. 36]. Diese erlauben eine verbesserte Modellbildung

¹ Auf eine Unterscheidung in statische, dynamische, einwertige und mehrwertige Nichtlinearität wird hier verzichtet. Da zu nichtlinearen Systemen keine geschlossene mathematische Theorie existiert, gibt es auch keine allgemeine Methode zur Analyse unbekannter nichtlinearer Systeme und der sie abbildenden Verfahren wie der Neuronalen Netze.

durch die Verwendung von vorliegendem Wissen, z. B. durch Identifikation der veränderlichen Variablen und der als konstant aufzufassende Modellparameter [Brockhoff 1977, S. 36]

3.1.3 Auswahl von Prognosemodell und Prognoseverfahren

3.1.3.1 Umfang des Erklärungsmodells

Die Forderung MEFFERTS [1980] nach einem kausallogisch befriedigenden Modell aller relevanten Einflussfaktoren der Absatzmärkte (vgl. Abschnitt 3.1.2.4) erscheint plausibel, jedoch kaum realisierbar. Sie erfordert die Bildung eines mikroökonomischen Totalmodells der betrieblichen Realität durch die Berücksichtigung aller realen wirtschaftlichen Einzeltransaktionen einer künftigen Periode und ihrer wechselseitigen Beziehungen auf den speziellen Produktmärkten [Gerfin 1964, S. 22]. Aufgrund der Komplexität eines derartigen Modells für unterschiedliche Produkte und lokale Märkte ist die Forderung daher als ein unmöglich zu realisierender Idealzustand zu verstehen [Gisholt 1976, S. 46]. Zudem ist die Erhebung, Modellierung und Speicherung der Variablen mit Kosten verbunden, sodass Totalmodelle trotz analytisch korrekter Abbildung der Realität gleichzeitig unwirtschaftlich sein können. Somit entsteht das Problem der wirtschaftlichen Auswahl der zu verwendenden Variablen, ihrer Lag-Strukturen und der Länge des Erklärungszeitraums in einem Partialmodell.

In der kausalanalytischen Modellierung können die unabhängigen Variablen in durch das Unternehmen kontrollierbare und nicht-kontrollierbare Variablen unterteilt werden.¹ Als nicht-kontrollierbare Variablen gelten jene, auf die die Aktivitäten des Unternehmens keinen direkten Einfluss ausüben, z. B. einzelne oder aggregierte Aktivitätsniveaus anderer Marktteilnehmer, Parameter in Form von Indikatoren der volkswirtschaftlichen, branchen- oder produktmarktbezogenen Entwicklung sowie die Störgrößen der Modellierung. Prognosemodelle mit stochastischen, nicht-kontrollierbaren Variablen erfordern, dass alle Realisationen der erklärenden Größen $x_{i,t-l}$ als Lag-Variablen noch im Erklärungszeitraum liegen [Scheer 1983, S. 8; Brockhoff 1977, S. 37]. Liegen die erklärenden Realisationen der unkontrollierbaren Variablen jedoch im Prognosezeitraum, so müssen diese vorangehend selbst prognostiziert werden, als $\hat{x}_{i,t+h}$, um sie für die Prognose verwenden zu können [Brockhoff 1977, S. 36]. „Sind die Werte der erklärenden Größen ebenso schwierig festzulegen wie die eigentlich interessierende Größe \hat{y}_{t+h} , so verlagert sich das Problem der Bestimmung von \hat{y}_{t+h} auf die Prognose der Größen $x_{i,t+h}$, und das Prognosemodell ist wenig hilfreich“ [Scheer 1983, S. 9].² Neben den stochastischen Variablen existieren jedoch auch deterministische, nicht-kontrollierbare unabhängige Variablen, deren wiederholtes Auftreten insbesondere als exogene Kalendereinflüsse in Form von Feiertagen, Urlaubszeiten oder Sonderveranstaltungen im Vorwege bekannt sind, deren Einfluss auf das Absatzniveau jedoch unbekannt ist. Diese sind in der Modellbildung auch für längere Prognosehorizonte direkt abbildbar.

Die kontrollierbaren Variablen der kausalanalytischen Modellierung umfassen hingegen jene Instrumentalvariablen, die im Rahmen des absatzpolitischen Instrumentariums durch die

¹ Die Bezeichnung „unabhängige Modellvariablen“ bezieht sich dementsprechend nicht auf die Kontrollierbarkeit durch den Entscheider. Exogene Variablen können sowohl kontrollierbar als auch nicht-kontrollierbar sein, wobei die Bezeichnung „kontrollierbar“ eine gezielte Steuerung der Instrumentalvariablen durch die Unternehmenspolitik impliziert. Die Klassifikation bezieht sich ferner auf ein spezifisches Entscheidungsmodell und die ihm zugrunde liegende Zeitspanne. Bei längerfristiger Betrachtung können z. B. unkontrollierbare Variablen durch das Unternehmen beeinflussbar werden.

² Dies ist z. B. der Fall, wenn sich die Variablen auf Aktivitäten der Konkurrenten, gesamtwirtschaftliche oder Umweltentwicklungen beziehen. So liegt z. B. in Kanada aufgrund des Rauchverbotes in Gebäuden eine hohe negative Korrelation des Absatzes von Zigaretten mit dem Wetter bzw. der Sonneneinstrahlung in der Beobachtungsperiode vor, da bei Temperaturen um -40° Celsius weniger auf der Straße geraucht wird. Im Gegensatz dazu existiert in Deutschland eine positive Korrelation, die entsprechend zu berücksichtigen wäre. Dieses bekannte Phänomen verlagert jedoch das Problem der Absatzprognose auf die Problemstellung einer weitaus komplexeren Wetterprognose, sodass die Variablen nicht berücksichtigt werden können.

Unternehmung variiert werden können. Sie erfassen explizit die direkte oder indirekte Wirkung unterschiedlicher Aktivitätsniveaus der absatzpolitischen Mittel auf den Prognosegegenstand. Zukünftige Realisationen $x_{i,t+h}$ der kontrollierbaren Variablen können durch das Unternehmen im Prognosezeitraum festgelegt werden. Sie gelten somit für die Prognose als deterministisch bekannt und erfordern keine Schätzung *ex ante*. Da der Wirkungszusammenhang zwischen Instrumenteinsatz und Ergebnis in der Realität jedoch nicht deterministisch ist, verlagert sich das Prognoseproblem auf die *Ex-ante*-Bestimmung einer Reaktionsfunktion der kontrollierbaren Variablen. Existieren Probleme der empirischen Ermittlung von Responsefunktionen einzelner Marketing-Instrumente¹ in den Modellen der Wirkungsprognose [Meffert 1980, S. 201], können die Auswirkungen der kontrollierten Variablen nicht quantifiziert werden. Die Variable kann entsprechend nicht in das Erklärungs- und Prognosemodell einfließen. Ferner erfordert die Integration kontrollierter Variablen deterministische Kenntnis über Zeitpunkt und Intensität des Einsatzes der absatzpolitischen Instrumente im Prognosezeitraum. Die Festlegung des Aktivitätsniveaus und des Zeitpunktes der Instrumente wird jedoch in der Praxis gerade auf der Basis der Erkenntnisse der Prognose variiert, um ein verbessertes betriebliches Ergebnis als das prognostizierte zu erreichen. Dies ist als Interdependenzproblem von Prognose und Planung bekannt [Brockhoff 1977] und führt häufig zu einer Falsifikation der Prognoseergebnisse sowie zu weiteren Problemen bei der Integration kontrollierbarer Variablen.

Ferner ist die Erhebung beobachtbarer Werte der abhängigen oder unabhängigen Variablen im Beobachtungszeitraum sowie der deterministisch-kontrollierbaren Variablen im Prognosezeitraum mit Aufwand verbunden. Die Schätzung der Reaktionsfunktion kontrollierbarer Variablen, die Prognose nicht-kontrollierbarer Variablen sowie die Erhebung beobachtbarer, vergangener und zukünftiger Realisationen kontrollierbarer und nicht-kontrollierbarer Variablen erfordert eine vorrausgehende Informationsbeschaffung. Daraus ergibt sich ein erweitertes Entscheidungsproblem des wirtschaftlichen Umfangs der Datennutzung hinsichtlich des Zeitfensters und einzelner Variablen. „Da die Informationsbeschaffung Kosten verursacht, wird sie selbst zum Entscheidungsproblem“ [Hüttner 1986, S. 17], das den Modellumfang entscheidend determiniert. Die Interdependenz von Prognosequalität, Kosten und Informationsversorgung kann nach GRANGER [1980] formal durch den Informationsgrad I_t im Zeitpunkt t dargestellt werden. Der Informationsgrad bzw. die verfügbare Informationsmenge (engl. *information set*) zum Zeitpunkt t repräsentiert die Grundlage der Prognose, bestehend aus den Zeitreihen der Beobachtungen der Gegenwart und der Vergangenheit [Granger 1980, S. 11]:

$$I_t : y_{t-l}, y_{t-l-1}, \dots, y_{t-L} \quad \text{mit } l, L \geq 0 \quad (3.16)$$

bzw. auf der Grundlage umfassenderer „information sets“ auf der Basis vergangener und gegenwärtiger Beobachtungen der betrachteten Zeitreihe sowie anderer Zeitreihen mit

$$I_t : y_{t-l}, x_{1,t-l}, x_{2,t-l}, \dots, x_{i,t-l} \quad , l \geq 0 \quad (3.17)$$

Für die Beschaffung unterschiedlicher Beobachtungen können unterschiedliche Kosten entstehen. Sind bestimmte Daten praktisch nicht verfügbar, bewegen sich die Kosten zur Datenbeschaffung gegen unendlich. Übersteigt der antizipierte Aufwand zur Bildung des

¹ Diese entstehen durch dynamische „Carry-over“-Effekte aus dem Einsatz der Instrumente der Absatzpolitik in vorangegangenen Perioden, durch sogenannte „Spill-over“-Effekte aus der Beeinflussung durch den Einsatz der absatzpolitischen Instrumente auf anderen Teilmärkten sowie den interdependenten Wirkungsverbund der absatzpolitischer Instrumente im Marketing-Mix

Prognosemodells den erwarteten Nutzenzugewinn in Form einer gesteigerten Prognosegenauigkeit, muss aus Gründen der Wirtschaftlichkeit auf die Berücksichtigung der Variablen im Erklärungsmodell verzichtet werden. Grundsätzlich wird eine höhere Prognosegüte bei zunehmender Informationsmenge erwartet [Granger 1980, S. 12 f.]. Dabei steigt die Prognosegenauigkeit degressiv mit zusätzlichen Kausalvariablen und bietet nur einen abnehmender Grenznutzen zunehmender Modellkomplexität. Umgekehrt sinkt die Prognosegenauigkeit durch ein stärker abstrahiertes Erklärungsmodell, wodurch die Findung einer effizienten Informationsversorgung datenspezifisch erfolgen muss [Brockhoff 1977, S. 51]. Die Komplexität der realen Entscheidungssituation muss somit nicht mit der Komplexität des verwendeten Erklärungsmodells identisch sein, sondern muss situativ in Abhängigkeit des Grenznutzens der Prognosegüte bzw. der Kosten einer falschen Entscheidung bestimmt werden. Dies erfordert die Bildung unterschiedlicher Kandidaten der „information sets“ und korrespondierender Prognosemodelle, und deren iterative Beurteilung mit passenden Prognoseverfahren (vgl. Abschnitt 3.4). Bei geringer Bedeutung der Entscheidung und somit der Prognosefehler für das Unternehmen kann eine starke Vereinfachung des Prognosemodells somit wirtschaftlich gerechtfertigt sein.

Eine spezifische Problemstellung kann somit in unterschiedlichen Erklärungs- und Prognosemodellen und mit speziellen Prognoseverfahren abgebildet werden, die eine Informationsversorgung in der Entscheidungssituation mit unterschiedlichen Genauigkeitsgraden und Kosten erlauben. Die Wahl des Erklärungsmodells zur Lösung empirischer Fragestellungen entspricht somit selbst einem Entscheidungsproblem, bei dem wirtschaftliche Überlegungen der Informationsbeschaffung und Modellanwendung zum Abwägen von Entscheidungsgüte und Entscheidungskosten dominieren und keine einfache Zielsetzung in Richtung auf eine „maximale“ Prognosequalität möglich ist.

Ein Spezialfall der Vereinfachung des Erklärungsmodells mit $i = 1$ abhängigen Variablen wird als univariates oder univariates kausales Modell bezeichnet:

$$y_t = f(x_t) + \varepsilon_t \quad . \quad (3.18)$$

Zeitreihenanalytische Modelle können daher als besonders vereinfachte Erklärungsmodelle für kurzfristige Prognosen betrachtet werden, bei denen das Modell auf die einzig erklärende Variable der Kalenderzeit t unter Konstanz aller exogenen Einflussfaktoren reduziert wird:

$$y_t = f(t) + \varepsilon_t \quad . \quad (3.19)$$

„Die Zeit wird als eine Ersatzgröße für die wirklichen Kausalfaktoren aufgefasst, deren Wirkung die Entwicklung der Zeitreihe ... unterliegt“ [Brockhoff 1977, S. 19]. In diesem Sinne abstrahieren Zeitreihenprognosen von den Instrumentalvariablen des Entscheidungsproblems [Lo 1979] und interpretieren das System aller kausalen Einflussgrößen als einen Ursachenkomplex im Sinne einer Black Box.¹ Ihre Genauigkeit hinsichtlich der Entscheidungsgüte und des verringerten Datenumfangs im Sinne der Entscheidungskosten ist somit gegen multivariate Verfahren im Vergleich festzustellen. Eine grundsätzliche Unterlegenheit zeitreihenanalytischer Modelle kann unter Berücksichtigung der Wirtschaftlichkeit der Informationsbeschaffung nicht festgestellt werden.

Unter Berücksichtigung pragmatischer Vereinfachungen werden der Umfang und die formale Struktur des Erklärungs- und des Prognosemodells als Partialmodell bestimmt. Dabei sind die

¹ Die stärkste Vereinfachung der zeitreihenanalytischen Modellierung wird durch eine Naive Prognose (NF1) erreicht, $\hat{y}_{t+1} = y_t + \varepsilon_t$, die lediglich den letzten Beobachtungswert der abhängigen Variablen extrapoliert (vgl. Abschnitt 3.2.3.2.1):

in den vorangegangenen Abschnitten entwickelten Strukturdeterminanten zu berücksichtigen, die in Tabelle 3.2 als morphologischer Kasten dargestellt sind.

Tabelle 3.2: Systematik der Determinanten von Erklärungs- und Prognosemodellen¹

Art des Prognoseergebnisses	Quantitatives Prognoseergebnis			Qualitatives Ergebnis
Formalisierung des Modells	Objektive Prognosen			Subjektive Prognosen
Skalenniveau der Prognosevariablen	metrisches Skalenniveau		nominales Skalenniveau	beliebiges Skalenniveau
Art der Modellvariablen	Zeitreihenprognose	kausale Prognose	kausale Klassifikation	kausale Prognose- und Klassifikationsmodelle
Zeitbezug der Modellvariablen	dynamische Modelle		statische Modelle	
Prognosehorizont	Einschrittprognosen ($t+1; h=1$)		Mehrschrittprognosen ($t+1, \dots, t+H; h=1, \dots, H$)	
Funktionale Form des Modells	lineares Modell		nichtlineares Modell	
Prognoseintervall	Punktprognose	Intervallprognose	Verteilungsprognose	

Die Systematik in Tabelle 3.2 erlaubt eine Konkretisierung zulässiger Erklärungsmodelle für die Problemstellung der Bedarfsprognose in der Warendisposition. Die schattierten Felder verdeutlichen Ausprägungen der Erklärungsmodelle, die aufgrund subjektiver Modellformen, qualitativer oder nominaler Prognoseergebnissen oder eine Vernachlässigung der Dynamik nicht zur Absatzprognose und Warendisposition angewendet werden können. Die Warendisposition erfordert quantitative Prognoseergebnisse in Form von Ein- oder Mehrschrittprognosen in Abhängigkeit des Prognosehorizontes. Dabei können sowohl Prognosemodelle der Punkt-, Intervall- oder Verteilungsprognose verwendet werden, um die einzelnen ökonomischen Variablen des Bedarfs bzw. des Absatzes je Ware vorherzusagen.

3.1.3.2 Zeitlicher Prognosehorizont

Als Entscheidungsmerkmal der Auswahl geeigneter Prognosemodelle dient bei vorgegebenem Prognosegegenstand zudem eine Unterscheidung nach der Fristigkeit des Prognosezeitraumes. Eine Reihe von Autoren unterscheiden so Prognosemodelle [Brockhoff 1977, S. 41] oder methodische Vorgehensweisen ihrer Bestimmung [Makridakis et al. 1980, S. 17] in kurzfristige, mittelfristige und langfristige Prognosemodelle.²

Der Umfang des Informationsgrades wird auch durch die Länge des Prognosezeitraumes bestimmt. Bei einer Ausdehnung des Prognosezeitraums müssen die Parameter eines kurzfristigen Erklärungsmodells als im Zeitablauf veränderbare Variable erfasst werden, um auf der Basis einer verfeinerten Kausalanalyse eine vergleichbare Prognosequalität für längere Prognosehorizonte zu erzielen, während dies bei kurzfristiger Prognose nicht erfolgen muss [Brockhoff 1977, S. 41]. „Wird der Prognosezeitraum ausgedehnt, so muß über den Zeiteinfluß hinaus auf ökonomische Variablen zur Erklärung des Artikelabsatzes zurückgegriffen werden, wenn die Prognose befriedigend sein soll“ [Brockhoff 1977, S. 41].

Die Kategorisierung der Zeiträume, die Abgrenzung der einzelnen Fristen in Bezug auf bestimmte Kalenderperioden und somit die Zuordnung von Verfahren zu Zeiträumen erfolgt

¹ In Tabelle 3.2 repräsentieren die schattierten Felder Ausprägungen der Erklärungsmodelle, die nicht zur Absatzprognose und Warendisposition angewendet werden können. Die dunkleren Felder identifizieren Erklärungsmodelle, die zur Warendisposition erforderlich sind, und die heller hinterlegten Felder mögliche Erweiterungen der Erklärungsmodelle, die in dieser Untersuchung vernachlässigt werden.

² Dabei ist anzumerken, dass generell keine „kurzfristigen Prognoseverfahren“ existieren, sondern lediglich kurzfristige Problemstellungen mit bestimmten Verfahren valide und reliabel gelöst werden können.

jedoch in der Literatur uneinheitlich. In Tabelle 3.3 ist eine Auswahl der Fristen von Prognosemodellen verschiedener Autoren abgebildet, welche diese Heterogenität belegt.

Tabelle 3.3: Fristen und Kalenderzeiträume von Prognoseverfahren

Fristen	Kalenderzeitraum						
	HANSMANN, MAKRIDAKIS	SCHEER	MEFFERT	HÜTTNER	GISHOLT, GRAF	SICKEL	HADELER
Kurzfristig	1 Monat bis 3 Monate	täglich bis 1,5 Jahre	< 1 Jahr	< 1 Jahr	< 1 Jahr	einige Monate	< 1 Jahr
Mittelfristig	3 Monate bis 2 Jahre	täglich bis 5 Jahre			1 Jahr bis 3 Jahre	1 Jahr bis 5 Jahre	< 5 Jahre
Langfristig	> 2 Jahre	1 Jahr bis 5 Jahre	≥ 10 Jahre	3 Jahre bis 5 Jahre	3 Jahre bis 10 Jahre	> 5 Jahre	> 10 Jahre

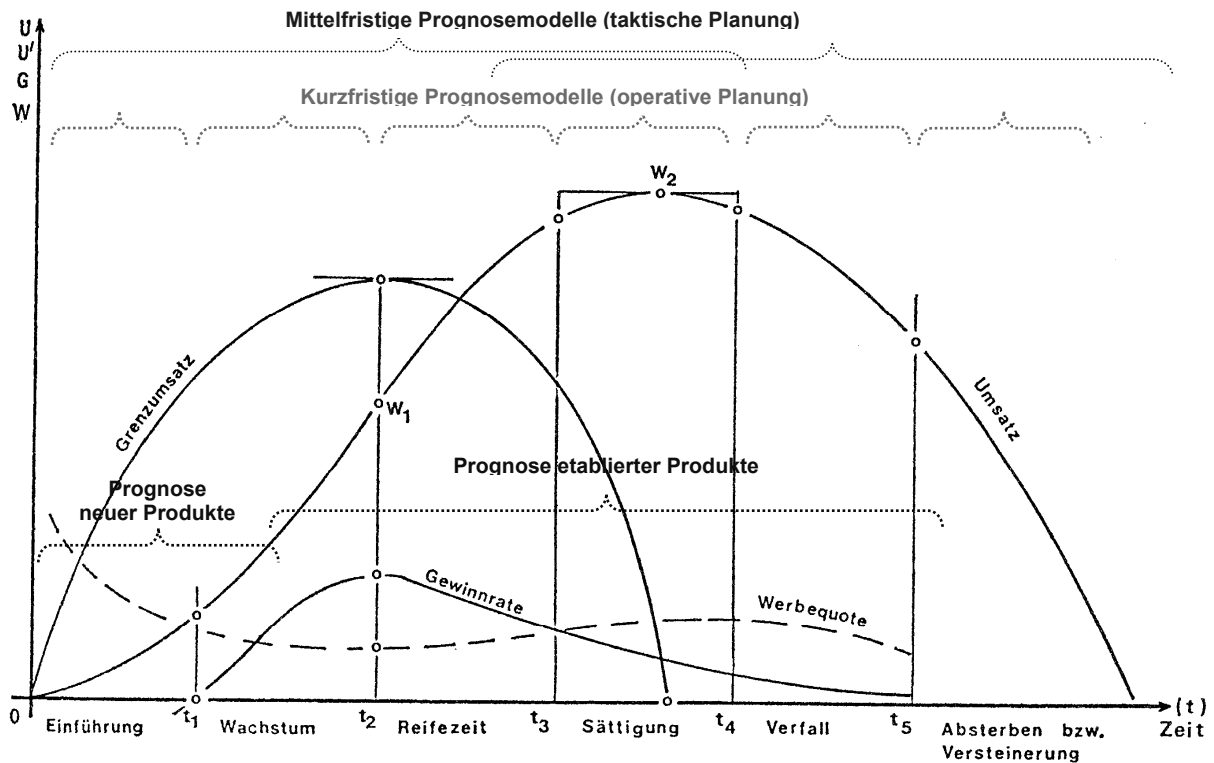
Details entnommen aus [Hansmann 1979, S. 13; Makridakis et al. 1980; Meffert 2000, S. 166; Scheer 1983, S. 51; Stickel 1997, S. 557; Hadelers 2000; Graf 1968, S. 22; Gisholt 1976, S. 47; Hüttner 1986, S. 20]. Weitere Abgrenzungen finden sich in [Brockhoff 1977, S. 40; Meffert und Steffenhagen 1977, S. 37].

Die Tabelle verdeutlicht, dass erhebliche Unterschiede in der Zuordnung der Zeiträume zu kurzfristigen, mittelfristigen und langfristigen Prognosen bestehen, so dass eine verlässliche Auswahl geeigneter Prognoseverfahren nicht anhand der Zeiträume erfolgen kann. GRANGER [1980] führt dies darauf zurück, dass die Klassifikation von Prognosen nach Zeithorizont abhängig von dem Prognosegegenstand und der Entscheidungssituation ist.¹ Zudem wird die relative Länge des Prognosezeitraumes von der zeitlichen Frequenz der Beobachtungswerte in Tagen, Quartalen, Jahren, Jahrzehnten etc. bestimmt [Hansmann 1979, S. 13] und der Prognosehorizont umso weiter gefasst, „je schneller Neuerungen sich im Gegenstandsbereich der Prognose durchsetzen und selbst wieder verdrängt werden“ [Brockhoff 1977, S. 41]. Eine Zuordnung von Kalenderzeiten oder Prognosezeiträumen zu spezifischen Prognoseverfahren erscheint somit wenig aussagekräftig, was eine valide Kategorisierung von Verfahren und Zeiträumen verhindert.

Da die Prognose als Informationsgrundlage für die betriebliche Planung dient, bestimmt der Zeithorizont der Entscheidungssituation den Prognosehorizont, und bedingt so die Auswahl eines geeigneten Prognoseverfahrens [Makridakis et al. 1980, S. 17]. Daher erscheint eine Zuordnung von kurz-, mittel- und langfristigen Prognosemodellen zu den Planungsebenen operativer, taktischer und strategischer Planung sinnvoll, erlaubt jedoch noch keine Rückschlüsse auf die Wahl spezifischer Prognosemodelle und -verfahren für einzelne Produkte. Zur Lösung erfolgt eine Zuordnung des Prognosehorizonts zu einzelnen Phasen des produkt-spezifischen Lebenszykluskonzeptes, der den idealtypischen Lebenszyklus von Produkten als zeitbezogenes Marktreaktionsmodell in disjunkten Phasen der Planung postuliert [Meffert 1980, S. 339]. Abbildung 3.6 gibt eine Übersicht über typische Verläufe der unterschiedlichen Prognoseobjekte des Umsatzes bzw. Absatzes, der Gewinnrate, der Werbequote und des Grenzumsatzes in den Phasen von der Einführung ($t_0 - t_1$) bis zur Auslistung ($t_5 - t_6$), die auch im Handel Gültigkeit besitzt [Oeldorf und Olfert 2002, S. 130].²

¹ So könne ein Prognosehorizont von zwei Wochen z. B. bei Wettervorhersagen langfristig sein, während er bei makroökonomischen Variablen, z. B. der Arbeitslosigkeit, sehr kurzfristig wäre [Granger 1980, S. 6].

² Nach MÜLLER-HAGEDORN kann das Lebenszyklus-Konzept auch auf Nachfrager und Haushalte in einzelnen Warengruppen sowie Betriebstypen des Handels zur sogenannten „store erosion“ ausgedehnt werden [Müller-Hagedorn 1998, S. 228ff, 369]. Diese Dynamik wird hier jedoch nicht berücksichtigt.

Abbildung 3.6: Absatzprognose im Produktlebenszyklus¹

Ein Produkt durchläuft einen Lebenszyklus aus charakteristischen Phasen, die unabhängig von der absoluten Lebensdauer des Produktes (von Monaten für Halbleiterchips bis zu Jahrzehnten für Industrieanlagen) durchlaufen werden [Meffert 1980, S. 340]. Innerhalb dieser Phasen wird die Marktreaktion, die Unternehmensstrategie und der Einsatz des marketingpolitischen Instrumentariums als weitgehend konstant betrachtet. Prognosen mit einem Prognosehorizont innerhalb einer dieser Phasen können somit als kurzfristig bezeichnet werden, und man kann von einer hohen Validität der Zeitstabilitätshypothese ausgehen. Dabei kann innerhalb jeder Phase eine Konstanz der Gesetzmäßigkeiten unterstellt und als Grundlage zur Zeitreihenextrapolation oder zur kausalen Modellierung genutzt werden. Im Fall einer zeitreihenanalytischen Interpretation entspricht dies einem Modell mit anhaltendem, progressivem trendbehaftetem Verlauf in der Einführungsphase ($t_0 - t_1$), einem linearen Trend in der Wachstumsphase ($t_1 - t_2$), einem degressiven Trend in der Reifephase ($t_2 - t_3$) und einem stationären Verlauf in der Sättigungsphase ($t_3 - t_4$) etc. (siehe auch Abschnitt 3.2.3.1).² Zu Beginn des Produktlebenszyklus müssen spezifische, analogiebasierte Prognosemodelle zur Vorhersage des Absatzes von neuen Produkten ohne Daten der Absatzhistorie eingesetzt werden, die nachfolgend aufgrund der Datenlage nachfolgend vernachlässigt werden. Alle weiteren Ausführungen betrachten etablierte Produkte, von denen Absatzdaten vorliegen.

Mittelfristige Prognosemodelle zur taktischen Planung, deren Prognosehorizont über eine oder mehrere Phasen hinaus reicht, erfordern hingegen bei verminderter Gültigkeit der Zeitstabilitätshypothese abweichende Prognosemodelle zur Berücksichtigung veränderbarer Umweltzustände und Marketingpolitiken. Eine taktische Planung erfordert vielfach die Prognose über den gesamten Lebenszyklus unter Verwendung spezifischer Prognoseverfahren der Marketing-Modellierung, z. B. durch das Diffusionsmodell nach BASS [1969], welches

¹ Abbildung entnommen aus [Meffert 1980, S. 341].

² Das Lebenszykluskonzept ist neben der Einführung neuer Produkte auch in der dynamischen Sortimentsplanung für die Modellierung und Prognose empirischer Bedarfsverläufe relevant, sowohl bei der Einführung, dem sogenannten „Phase-In“ in bestehende Sortimente, beim „Relaunch“ von Produkten als auch beim Abverkauf auslaufender Produkte, dem sogenannten „Phase-out“.

den Adaption- und Diffusionsprozess eines Produktes über mehrere Phasen betrachtet. Eine strategische Planung von Produkten und Marken über mehrere Lebenszyklen hinweg beinhaltet dagegen die Variation der Rahmenbedingungen der Märkte und der Unternehmenspolitik. Somit verbieten langfristige Prognosemodelle die Extrapolation von Beobachtungswerten in die Zukunft auf Basis der Zeitstabilitätshypothese, weshalb sowohl zeitreihenanalytische als auch kausale Prognoseverfahren für die strategische Planung ungeeignet erscheinen¹

Der Prognosehorizont der Warendisposition wird somit durch den Entscheidungshorizont der operativen Absatzplanung bestimmt, die entsprechend kurzfristige Prognosemodelle erfordert. Dies erlaubt einen Ausschluss der Betrachtung von Modellen und Verfahren der taktischen und strategischen Planung auf Absatzmärkten über mehrere Phasen des Lebenszyklus eines Produktes, z. B. der Verfahren der Trendextrapolation [Hansmann 1983, S. 104 ff.] oder von Diffusionsmodellen.² Für die verbleibenden Modelle der kurzfristigen Absatzprognose kommen unterschiedliche Verfahren zur Anwendung, die nachfolgend spezifiziert werden.

3.1.3.3 Verfahren der Bedarfsprognose

Der Gegenstand der Prognose kann durch die Bildung unterschiedlicher Erklärungs- und Prognosemodelle der kausalen Prognose oder Zeitreihenprognose, unter Berücksichtigung einer unterschiedlichen Auswahl kontrollierbarer oder unkontrollierbarer Erklärungsvariablen, linearer oder nichtlinearer Verknüpfung der unabhängigen Variablen und unterschiedlicher Informationsversorgungen vorhergesagt werden.

Zur Lösung der Modelle wurde eine Vielzahl konkurrierender Prognoseverfahren entwickelt, die der Berechnung spezifischer Erklärungsmodelle dienen. Ein Prognoseverfahren ist eine konkrete Anweisung zur gezielten Erstellung einer Prognose nach einer detaillierten und systematischen Handlungsvorschrift aus vorgegebenen Arbeitsschritten und Regeln. Sie bestimmt, auf welche Weise nach bestimmten Handlungsgrundsätzen (Prinzipien) planmäßig vorgegangen wird [Stickel 1997, S. 557, 731]. Ein Prognoseverfahren dient somit der systematischen Aufstellung und Berechnung von Prognosen auf der Basis der formulierten Erklärungsmodelle [Stickel 1997, S. 557].³ Dabei kann ein einfaches, zeitreihenanalytisches Erklärungsmodell eines stationären Bedarfs sowohl durch ein Naives Verfahren, einen gleitenden Mittelwert unterschiedlicher Länge und Gewichtung, Verfahren der Exponentiellen Glättung, ARIMA-Verfahren, Schätzung mit unterschiedlichen Verteilungsfunktionen oder Verfahren der künstlichen Intelligenz einschließlich der Neuronalen Netze berechnet werden (vgl. Abschnitt 4.4). Dabei bestehen Beschränkungen in der Anwendbarkeit einzelner Verfahren auf spezifische Formen, z.B. nichtlinearer Erklärungsmodelle. Zudem zeigen die Verfahren eine unterschiedliche Genauigkeit der Prognose, was zu einem Problem der validen und reliablen Verfahrensauswahl führt. Die Frage der Verfahrensauswahl kann nicht theoretisch, sondern allein durch eine empirische Bestimmung der Prognosegüte in der Entscheidungsunterstützung eines Verfahrens bestimmt werden. Im Anschluss werden zunächst jene Prognoseverfahren analysiert, welche die Struktur des quantitativen Erklärungs-

¹ Stattdessen kommen beispielsweise subjektive Verfahren der Szenariotechniken durch Experten zum Einsatz.

² Auf eine Darstellung der in der operativen Bedarfsprognose eingesetzten linearen Trend-Regression entfällt, da das Verfahren für Zeitreihen mit Niveaushiftungen nicht robust ist, die im Handel vielfach auftreten.

³ Viele Autoren nehmen keine stringente Unterscheidung zwischen Erklärungsmodell und Prognosemodell zur Spezifikation der Problemstellung getrennt von dem Prognoseverfahren der Bearbeitungsvorschrift vor [Brockhoff 1977; Hansmann 1979; Scheer 1983]. Daraus ergeben sich weniger trennscharfe Abgrenzungen insbesondere neuartiger Verfahren wie der Neuronalen Netze, die anders als statistische Verfahren Erklärungsmodelle unterschiedlicher Modellklassen abbilden können (siehe Abschnitt 4.4). Zudem entstehen so unscharfe Zuordnungen bestimmter Verfahren zu Zeiträumen, beispielsweise durch SCHEER, der langfristige Prognoseverfahren den qualitativen Prognoseverfahren sowie qualitative Verfahren den Naiven Verfahren gleichsetzt [Scheer 1983].

und Prognosemodells der Bedarfsprognose abbilden können.¹ Diese werden in Kapitel 6 einem Verfahrensvergleich auf empirischen Absatzdaten unterzogen, um ihre Prognose- und Dispositionsgüte objektiv zu bestimmen.

SCHEER [1983, beiliegende Falttafel] präsentiert einen Orientierungsbaum für die Verfahrensauswahl aufbauend auf dem Prognosehorizont, dem Vorliegen exogener Variablen und der Kenntnis der Datenmuster. ARMSTRONG [2001d] entwickelt ein abweichendes Baum-schemata zur Auswahl adäquater Prognoseverfahren, aufbauend auf der Struktur und dem Kenntnisgrad der Datenbasis durch den auswählenden Experten (siehe Abbildung 3.28). In Anlehnung an die Struktur der Determinanten der Erklärungs- und Prognosemodelle wird in Abbildung 3.7 eine Auswahl ökonomischer Prognoseverfahren dargestellt.²

Die objektiven Verfahren in Abbildung 3.7 sind nach Verfahrensklassen geordnet.³ Sie umfassen die Naiven Verfahren, Verfahren der Mittelwerte einschließlich der Exponentiellen Glättung, ARIMA-Verfahren, Neuronale Netze sowie komplexe nichtlineare Modelle zur Lösung zeitreihenanalytischer Modelle. Die Verfahren der kausalanalytischen Modellierung umfassen insbesondere die Ansätze der Regression, Erweiterungen von Verfahren der Mittelwertbildung um Interventionsvariablen sowie Neuronale Netze, die nahezu unverändert für beide Modellformen eingesetzt werden können. Aufgrund der generischen Anforderungen der Warendisposition eignet sich eine Vielzahl objektiver Verfahren der Zeitreihen- und Kausalprognose ebenso wie subjektive Verfahren zur Bedarfsprognose. Daher erfolgt eine Untersuchung der empirischen Bedeutung einzelner Verfahren, um die nachfolgende Untersuchung und den Verfahrensvergleich auf repräsentative Prognoseverfahren zu beschränken.

¹ Entsprechend der Vielzahl alternativer Prognosemodelle und Anwendungsbereiche findet sich in der Literatur eine Vielzahl unterschiedlicher Systematiken zur Klassifikation von Prognoseverfahren in unterschiedlichen Disziplinen. So gliedert HANSMANN [1983] auf oberster Ebene in subjektive versus objektive Verfahren, SCHEER [1983] hingegen nach der Fristigkeit des Prognosehorizontes und MEFFERT und STEFFENHAGEN [1977] in mikro- versus makroökonomische Verfahren. Eine Übersicht alternativer Systematiken findet sich bei EMDE [1989]

² Eine umfassende Darstellung aller konkurrierenden Verfahren ist aufgrund ihrer Vielfalt nicht möglich.

³ Die Systematik der Verfahrensklassen weicht somit von der nachfolgenden Gliederung der Diskussion nach dem Zeitreihenverlauf ab, die aufgrund der entscheidungsorientierten Betrachtung für eine Analyse konkurrierender Prognoseverfahren zur Warendisposition vorzuziehen ist.

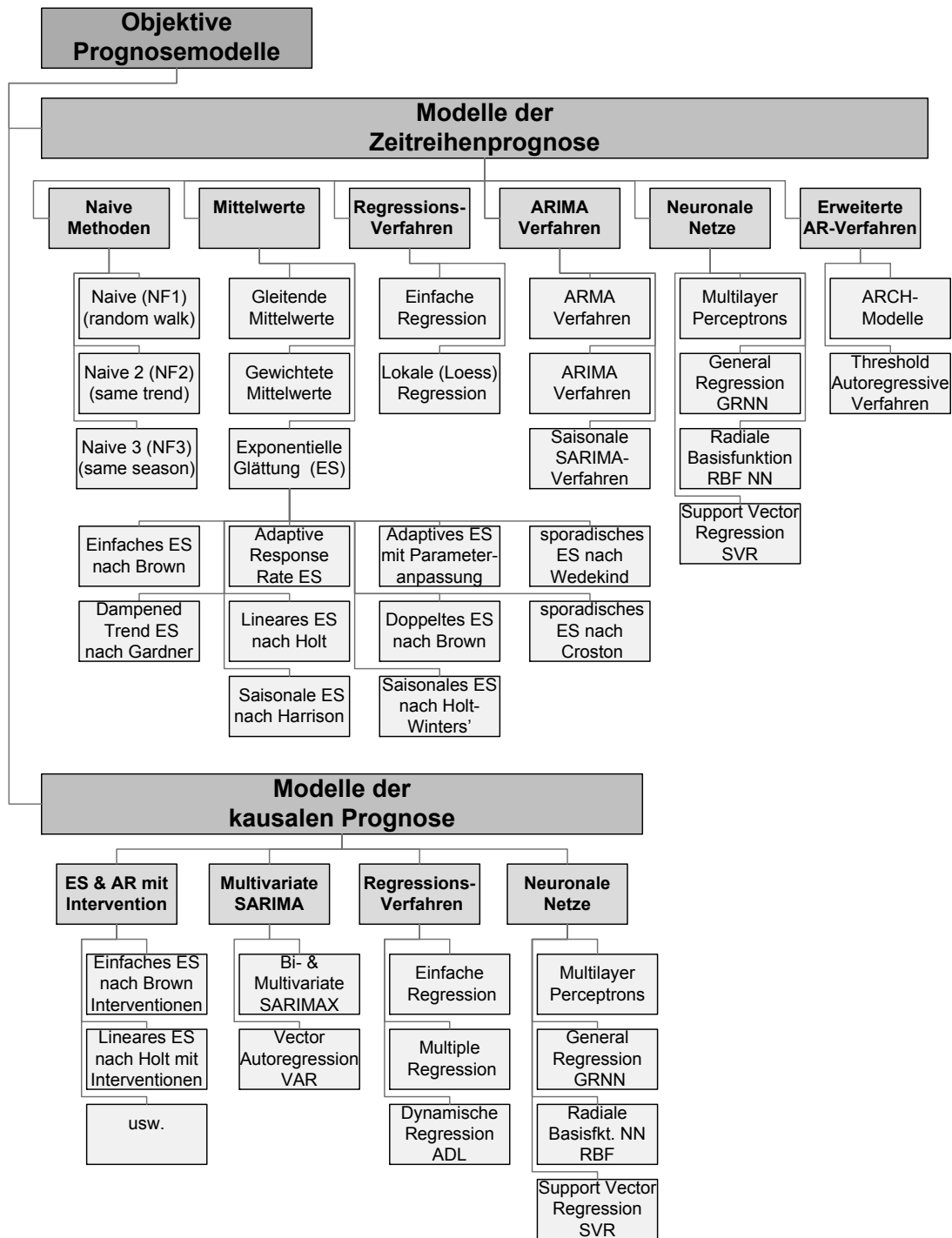


Abbildung 3.7: Systematik objektiver Prognosemodelle und Prognoseverfahren

3.1.3.4 Prognoseverfahren im Handel

Zur Feststellung der empirischen Relevanz von Modellen und Verfahren zur Absatzprognose in der Handelpraxis existieren nur vereinzelte empirische Erhebungen über den Einsatz von Prognoseverfahren. In Abbildung 3.8 sind die Ergebnisse einer Serie von Primärerhebungen an Prognostikern unterschiedlicher Industriezweige bei zehn Konferenzen des Institute of Business Forecasting (IBF) in den Jahren 2001 und 2002 abgebildet.¹

¹ Die Befragung umfasst 1072 Antworten von Absatzplanern aus 13 Industriezweigen, durch Kombination der Befragung von 379 Absatzplanern bei fünf Konferenzen der IBF im Jahr 2001 und von 693 Absatzplanern in 19 Industriezweigen bei 5 IBF-Konferenzen in 2002.

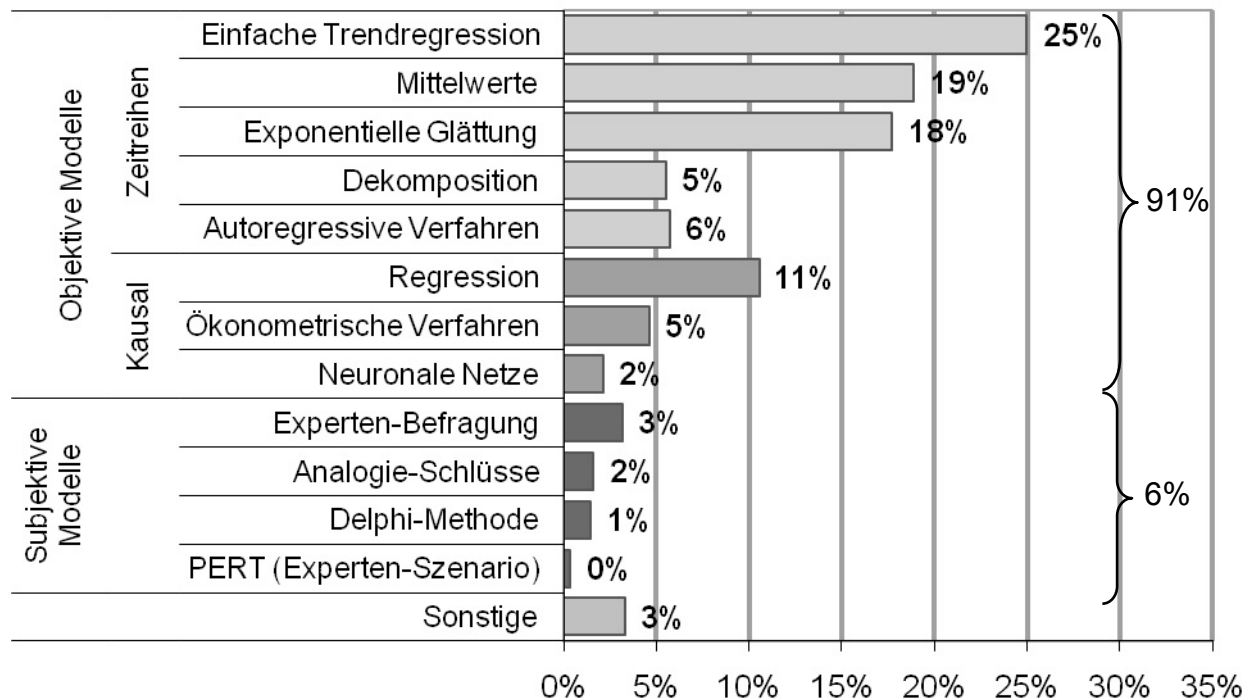


Abbildung 3.8: Anwendung von Prognoseverfahren in der betrieblichen Praxis¹

Demzufolge dominiert in der Praxis die Anwendung objektiver Verfahren in 91 % aller Betriebe, während subjektive Verfahren dagegen in nur 6 % der Betriebe angewendet werden.² Im Rahmen objektiver Verfahren setzen 73 % der Betriebe einfache Verfahren der Zeitreihenprognose ein, insbesondere die lineare Trendregression, Mittelwerte sowie die Exponentielle Glättung und zu einem geringeren Teil auch Verfahren der Dekomposition und ARIMA. Kausale Prognoseverfahren werden nur von 18 % der Betriebe eingesetzt, insbesondere univariate und multivariate Partialmodelle der Regression zur Interventionsmodellierung unter Berücksichtigung ausgewählter exogener Variablen der Kalendereffekte, der Absatzpolitik und weiterer Einflussfaktoren, deren Auftreten zum Prognosezeitpunkt bekannt ist. Auch Verfahren der künstlichen Neuronalen Netze werden als kausalanalytische Modelle vereinzelt (2%) angewendet.³ Die industriübergreifenden Erkenntnisse der IBF-Studie bestätigen frühere Befragungen [Dalrymple 1987; Moon et al. 2003], woraus jedoch noch keine Erkenntnisse für die Praxis der Absatzprognose im Handel abgeleitet werden kann.

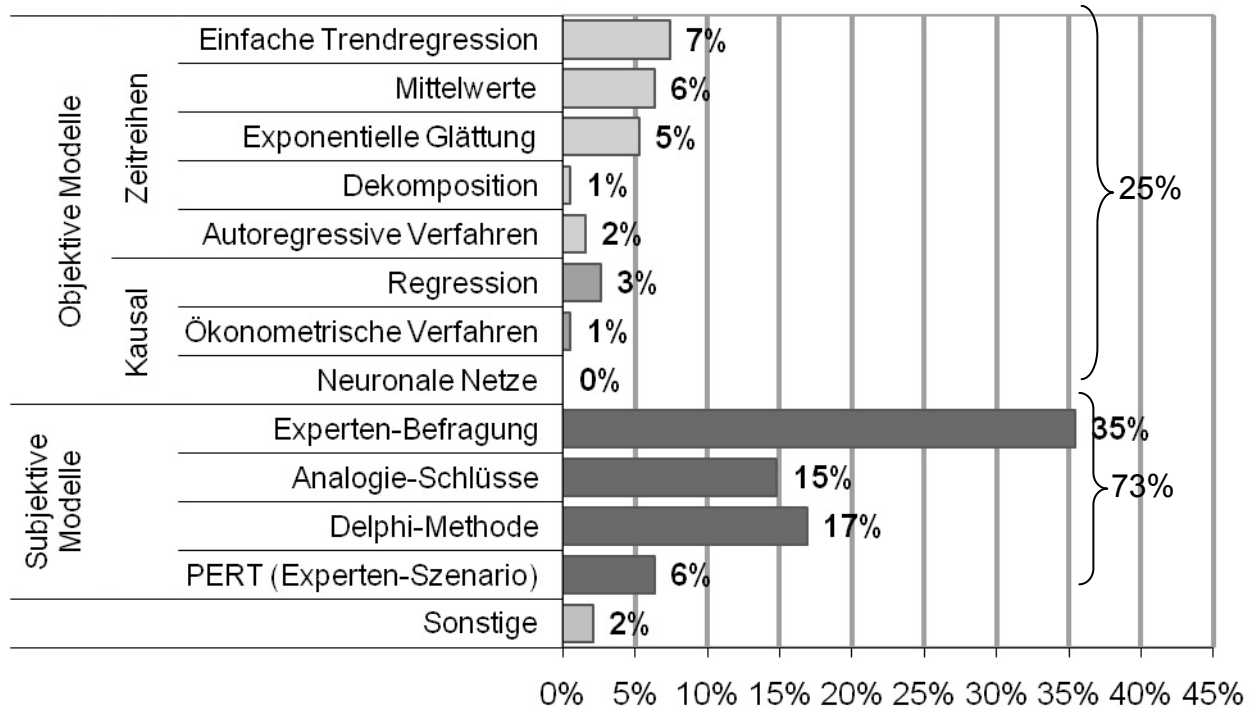
Aus der IBF-Befragung der Unternehmen des Handels in Abbildung 3.9 ergibt sich ein abweichendes Bild, in dem subjektive Verfahren mit 73 % die objektiven Verfahren der Zeitreihenprognose und der kausalen Prognose deutlich dominieren.⁴

¹ Eigene Darstellung aus den Mittelwerten der Befragungen bei IBF-Konferenzen in 2001 und 2002 über alle Industriezweige [Jain 2003, S. 39-49].

² Darüber hinaus besitzen subjektive Verfahren eine erhebliche praktische Relevanz, da sie im Prozess der Absatzplanung (engl. *demand planning* oder auch *sales & operations planning*, S&OP) vielfach mit den Werten der objektiven Prognose kombiniert werden, um subjektives, nicht quantifizierbares Expertenwissen zukünftiger Absatzperioden zu berücksichtigen. Dieses wurde in der Befragung jedoch nicht berücksichtigt.

³ In der Befragung des IBF werden Neuronale Netze fälschlicherweise der Gruppe kausaler Verfahren zugeordnet, wenngleich sie aufgrund der erheblichen Freiheitsgrade in der Modellbildung sowohl als kausal- als auch als zeitreihenanalytische Modelle konfiguriert werden können, und überwiegend als univariate Verfahren der Zeitreihenanalyse modelliert werden (vgl. Kapitel 4).

⁴ Von den 1072 Befragten stammten 10 im Jahr 2001 und 35 im Jahr 2002 aus Betrieben des Handels. Daher ist zu berücksichtigen, dass die Erhebung des Handels daher eine geringere Validität und Reliabilität besitzt.

Abbildung 3.9: Anwendung von Prognoseverfahren im Handel¹

Nur wenige der Prognoseverfahren in Abbildung 3.7 dargestellt werden im Handel praktisch eingesetzt, was auf die handelspezifischen Determinanten der Erklärungs- und Prognosemodelle zurückgeführt werden kann. Nur 25 % aller Unternehmen wenden der Befragung zufolge objektive Verfahren der Zeitreihenanalyse (21 %) oder gar der Kausalanalyse (4 %) an, im Gegensatz zur Anwendung subjektiver Verfahren bei 73 % der Befragten. Innerhalb der objektiven Verfahren dominieren erneut die Ansätze der einfachen und exponentiell gewichteten Mittelwerte sowie der Trendregression. NN werden von den befragten Unternehmen des Handels nicht eingesetzt. Die Unterschiede in der Verfahrensanwendung zwischen allen Industriezweigen und dem Handel wird in Abbildung 3.10 deutlich:

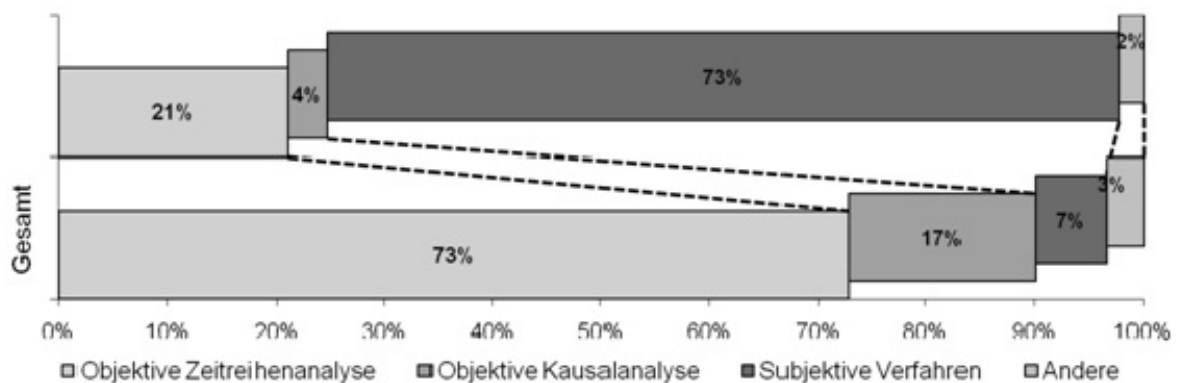


Abbildung 3.10: Anwendung von Prognoseverfahren in Industrie und Handel

Die Ergebnisse stehen in offensichtlichem Gegensatz zu den Erkenntnissen der Problemstellung der Warendisposition im Handel aus Kapitel 2, da subjektive Verfahren, z. B. durch Expertenurteile oder strukturierte Ansätze von Delphi (vgl. Abschnitt 3.2.1) aufgrund der

¹ Eigene Darstellung aus den Mittelwerten der Befragungen bei Konferenzen des IBF in 2001 und 2002 für alle Absatzplaner des Handels [Jain 2003, S. 39-49]

Einbindung von Experten zeitaufwendig, kostenintensiv und nicht zu automatisieren sind. Zudem erlauben nur 12 % der eingesetzten objektiven Verfahren¹ die Vorhersage saisonaler Zeitreihenmuster, deren Struktur im Handel vielfach vorzufinden ist.

Nach einer Erhebung durch ATZBERGER [1998, S. 9] dominierten im deutschen Handel bereits in 1997 „komplexe Verfahren“ (60 %) die Anwendung der Exponentiellen Glättung (38 %) und die der gleitender Durchschnitte (38 %), was ebenfalls nicht mit den Ergebnissen der IBF-Umfrage übereinstimmt.² Eine Studie der ABERDEEN GROUP [2004] ordnet die Prognoseverfahren zur Bedarfsvorhersage in Betrieben der Konsumgüterindustrie in unterschiedliche Reifegrade der Prognoseprozesse ein. Sie identifiziert eine kontinuierliche Weiterentwicklung der Komplexität der angewendeten Modelle zur monatlichen Prognose von den Naiven Verfahren um 1980 über zeitreihenanalytische Verfahren zur monatlichen Prognose von trend- und saisonbehafteten Zeitreihen in den 1990er Jahren bis hin zum Stand heutiger Innovationsführer mit wöchentlicher oder täglicher Prognose durch kausalanalytische Verfahren unter Berücksichtigung exogener Faktoren der Kannibalisierung, Handelsaktionen und des Absatzes verwandter Produkte und Produktfamilien [Aberdeen Group 2004, S. 3; 19].³ Für die Zukunft postuliert die Studie ein erhebliches Potenzial in der Verringerung von Prognosefehlern durch kollaborative Verfahrensansätze des Informationsaustausches, der Integration der Prognose für eine kunden- und produktindividuelle Preisdifferenzierung, und durch fortgeschrittene, kausalanalytische Prognoseverfahren der Mustererkennung durch „intelligente Verfahren“ im Sinne der Neuronalen Netze [Aberdeen Group 2004, S. 15].⁴ Die Vielzahl statistischer Prognoseverfahren aus Abbildung 3.7 findet somit in der Handelspraxis aufgrund ihrer Komplexität, der mangelnden Automatisierbarkeit oder den prohibitiven Datenanforderungen bislang keinen Einsatz bei der Warendisposition. Der Handel setzt bis auf wenige Innovationsführer überwiegend einfache, lang etablierte Prognoseverfahren der Zeitreihenanalyse ein.⁵

Neuere Verfahren der nichtlinearen Zeitreihenanalyse wie Threshold Autoregressive Modelle [Tong 1983; Tong 1990] oder Exponential Autoregressive Modelle [Ozaki 1980; Ozaki 1992] sowie von Autoregressive Conditional Heteroskedasticity-(ARCH-)Modellen [Engle 2002; Engle 1982] werden aufgrund der hohen Anforderungen in der Spezifizierung der Modelle und der geringen empirischen Relevanz für die Warendisposition im Handel vernachlässigt.⁶ Methodische Erweiterungen, die nicht auf einzelne Prognoseverfahren bezogen sind, z. B. die Bildung von Gremien mehrerer Prognoseverfahren, werden ebenfalls nicht berücksichtigt.

¹ Einzelne Varianten der der Exponentiellen Glättung, ARIMA, der Dekomposition sowie der kausalen Verfahren der Regression und der ökonometrischen Modelle sind zur Abbildung saisonaler Verläufe fähig.

² Dies kann sowohl auf Verzerrungen der Ergebnisse durch eine geringe Stichprobengröße, die Befragung spezifischer Formen der Handelsbetriebe sowie auf die Befragung von Experten außerhalb der operativen Absatzplanung des Handels zur Warendisposition zurückgeführt werden.

³ Sie spezifizieren darüber hinaus den Übergang der Prozesse zur Absatzplanung von der einfachen Kalkulation bis hin zu einem iterativen, dynamischen Prozess unter Berücksichtigung mehrfacher interner Eingangsgrößen und resultierender Szenarien [Aberdeen Group 2004].

⁴ Die in dieser deskriptiven Studie erhobenen Charakteristika können auch für eine Einschätzung des Reifegrades der Warendisposition und der Absatzplanung im Sinne einer „Demand Planning Maturity“ in Unternehmen verwendet werden.

⁵ Die Prognoseverfahren der Handelspraxis, einschließlich der multivariaten saisonalen ARIMAX-Ansätze, beruhen auf etablierten Verfahren und Techniken der 1960er Jahre. Ein Grund für die zögerliche Anwendung komplexer Verfahren von einzelnen Innovationsführern ist auf die mangelnde Automatisierung der Verfahrensanwendung zurückzuführen. So erfordert die automatische Auswahl eines Modells der Exponentiellen Glättung, die automatische Spezifikation von ARIMA-Modellen oder dynamischer Regressionsmodelle eine zeitreihenindividuelle Modellierung durch regelbasierte Methoden der Expertensysteme. Die systematische Untersuchung der Ursachen für den mangelnden Einsatz objektiver Verfahren ist jedoch nicht Aufgabe dieser Arbeit, sollte jedoch in Folgestudien analysiert werden.

⁶ Ebenso wird auf eine Darstellung der Verfahren des Kalman-Filters oder des Adaptiven Filterns [Haykin 1999] verzichtet, da diese aufgrund der historischen ADALINE-Architektur und das „Least-Mean-Square“-Gradientenabstiegsverfahren eng mit der Entwicklung von Neuronalen Netzen verwandt sind, die in Abschnitt 4 eingehend behandelt werden.

Dazu zählen auch organisatorische Ansätze innerhalb der Unternehmen, z. B. durch verbesserte Kombination von objektiven Prognoseverfahren mit subjektiven Expertenurteilen oder des Informationsaustausches zwischen Unternehmen im Sinne des Collaborative Planning, Forecasting and Replenishment (CPFR), da diese vielfach nicht analytisch oder theoretisch fundiert sind.¹ Alternative Verfahren der künstlichen Intelligenz, der Fuzzy Logik [Chen und Chen 2000; Yager 1982] oder der Support Vector Regression [Mangasarian und Musicant 2000; Tay und Cao 2001; Crone et al. 2006a; Crone et al. 2006b; Crone et al. 2006c; Crone et al. 2006d] erfordern aufgrund ihrer Neuartigkeit und der Komplexität der Modellierung eine eigenständige systematische Untersuchung durch einen empirischen Vergleich. Trotz ihres möglichen Potenzials in Anwendungen der Warendisposition müssen diese hier unterbleiben.

Im Anschluss werden die relevanten Prognoseverfahren untersucht, die im Vergleich zu Neuronalen Netzen in der Warendisposition im Handel zu beurteilen sind. Die Diskussion der methodischen Grundlagen statistischer Verfahren dient als Referenz im Vergleich mit Neuronalen Netzen und ist nicht als umfassende Darstellung aller einsetzbaren Methoden zu verstehen. Vielmehr beschränkt sich die Untersuchung auf jene Prognoseverfahren, die in der empirischen Absatzprognose zur Warendisposition etabliert sind und somit in Konkurrenz zu einer zukünftigen Anwendung von Neuronalen Netzen stehen.

3.2 Berechnung von Prognoseverfahren

3.2.1 Subjektive Prognoseverfahren

Subjektive Prognoseverfahren, „sind ein Sammelbegriff für zweckmäßige, methodisch erarbeitete Prognoseregeln ... ohne schematisches Prognosemodell“ [Hansmann 1979].² Der geringe Formalisierungsgrad subjektiver Verfahren entsteht durch eine nur schwach ausgebildete Theorie auf der Basis subjektiver, und somit nicht unmittelbar nachprüfbarer Erfahrungen von Experten, deren Expertise auf Spezialgebieten für die Prognose nutzbar gemacht wird [Hansmann 1979].³ Der subjektive Charakter der Verfahren dominiert etwaige quantitative, objektive Elemente des Verfahrens, obwohl auch mathematisch-statistische Operatoren zur Berechnung der subjektiven Prognoseaussage zur Anwendung kommen können oder im Rahmen des Stichprobendesigns für die Auswahl der Befragungssubjekte statistische Methoden eingesetzt werden können [Gisholt 1976, S. 111].

In der Modellierung objektiver Verfahren ist zu berücksichtigen, dass vielfach subjektive Entscheidungen den Prozess der Modellierung objektiver Erklärungsmodelle und die Verfahrensauswahl beeinflussen. Insbesondere die iterative Selektion einzelner Zeitreihen, Variablen und Variablen-Lags sowie die Codierung von intuitivem Expertenwissen sowie nur schwer codierbarer Information für objektive Prognosemodelle sind subjektive Aspekte der Prognose. Bezüglich des angewandten Verfahrens erlaubt der Anteil subjektiver Bearbeitungsschritte jedoch keine Zuordnung zu den subjektiven Verfahren.

Zu den subjektiven Prognoseverfahren zählen u. a. einfache Primärbefragungen von Wirtschaftssubjekten [Gisholt 1976, S. 112] oder unternehmensinterner bzw. -externer Experten auf Angebots- und Nachfrageseite, wie z. B. Händler, Endverbraucher oder Mitglieder der

¹ Selbst der prominente Ansatz des CPFR als Bestandteil der ECR-Initiative wird bisher nur geringfügig umgesetzt. Eine Studie der GMA zeigt, dass nur 7 % der Betriebe der Konsumgüterindustrie die Prognosen und Absatzpläne des Handels erhalten [GMA und IBM 2005] und dass Bedarfsprognosen noch immer auf der Basis historischer Absatzzahlen oder Kundenaufträge erfolgen [GMA und IBM 2005, S. 19].

² Sie werden von HANSMANN [1979; 1983] abweichend als heuristische Prognoseverfahren bezeichnet.

³ Dabei bezieht sich das Expertenwissen auf den Anwendungsbereich und nicht auf die Anwendung von Prognosen oder Prognoseverfahren.

eigenen Verkaufsorganisation [Meffert 1980, S. 201; Stickel 1997, S. 557]. Ferner gelten strukturierte Ansätze der manuellen Delphi-Methode [Wechsler 1978, S. 46-85; Gisholt 1976, S. 140-156] oder der computerunterstützten Delphi-Methode [Brockhoff 1979]¹, Erweiterungen zur Cross-Impact-Analyse [Gisholt 1976, S. 161], Szenario-Techniken und verwandte synoptische Methoden [Stührenberg und Töpken 1996], historische und direkte Analogiemethoden [Gisholt 1976, S. 122], die Relevanzbaumanalyse [Hansmann 1983, S. 18; Werner 2000, S. 20] sowie Expertenrunden zu den subjektiven Verfahren. Diese Ansätze zielen auf eine Objektivierung der Prognose durch die Steigerung der Validität und Reliabilität subjektiver Schätzungen unter Eliminierung systematischer Verzerrungen, z. B. durch Befragung von Gruppen zur Nivellierung der Subjektivität von Individuen [Wechsler 1978] sowie durch mehrstufige, rückgekoppelte Befragungs- und Feedbackstrukturen mit einem erhöhten Formalisierungsgrad, wie ihn etwa die Delphi-Methode aufweist.

Subjektive Prognoseverfahren erfassen und verwerten uncodiertes Wissen, Kenntnisse, Erfahrungswerte, die Fantasie und die Intuition von Experten bzw. Disponenten in Bezug auf qualitative und quantitative Sachverhalte [Stickel 1997, S. 557]. Sie eignen sich somit insbesondere dann für eine Prognose, wenn eine Formalisierung nicht möglich oder unwirtschaftlich ist und objektive Verfahren nicht eingesetzt werden können.² Dies ist insbesondere der Fall bei der Erfassung qualitativer, nicht oder nur schwer strukturierbarer oder quantifizierbarer Marktentwicklungen, in Problemstellungen ohne vorliegende Vergangenheitsdaten oder bei Strukturbrüchen im Beobachtungs- oder Prognosezeitraum, in denen die Daten keine valide Modellbildung ermöglichen. Zudem kommen sie häufig in Kombination mit objektiven Methoden zum Einsatz, indem unsystematische und nicht quantifizierbare Einflüsse der Werbung, Produktionsengpässen etc., die nicht in den Zeitreihen objektiver Verfahren enthalten sind, nachträglich mit den quantitativen Basis-Prognosen kombiniert werden. In diesen Fällen wird im Englischen von „judgemental forecasting“ gesprochen [Armstrong 2001d].

Aufgrund mangelnder Formalisierung bergen subjektive Verfahren jedoch die Gefahr, dass durch persönliche Motivation und Ziele der Experten bewusste oder unbewusste, jedenfalls unkontrollierbare Verzerrungen in die Prognose einfließen [Armstrong 2001d]. Ferner bereiten Ausbildung und Auswahl geeigneter Experten zur Angabe valider und reliabler Prognosen weitere Probleme [Meffert 1980, S. 202]. Aufgrund der Risiken und hohen Kosten der Durchführung einer subjektiven Prognose im Vergleich zur computergestützten Automatisierung durch objektive Prognoseverfahren werden subjektive Verfahren selten bei regelmäßigen und standardisierbaren Massenprognosen eingesetzt, sondern bei prominenten Entscheidungen, in der Kopplung mit objektiven Methoden und bei der Interpretation der Prognoseergebnisse [Meffert 1980, S. 203]. Somit sind subjektive Verfahren in der Absatzprognose des Handels nur bedingt effizient anwendbar.

Da subjektive Verfahren im Gegensatz zu objektiven Verfahren keine formalen Prognosemodelle besitzen, sind ihre Ergebnisse nicht in einer Ex-post-Simulation berechenbar, sondern erfordern die Dokumentation der subjektiven Prognosen zum Erstellungszeitpunkt. Daher können im Verfahrensvergleich der Neuronalen Netzen mit den statistischen Prognoseverfahren in Kapitel 6 keine subjektiven Verfahren evaluiert werden.³

¹ Sie zielen jedoch lediglich auf eine Objektivierung der Ergebnisse bei rein subjektiven Prognosen und nicht auf eine Formalisierung der Modellbildung im Sinne objektiver Prognoseverfahren.

² Zudem finden sie in der Vorhersage langfristiger Entwicklungen Anwendung, was jedoch aufgrund der abweichenden Problemstellung hier nicht relevant erscheint.

³ Subjektive Prognosen sind bei fehlenden Aufzeichnungen selbst durch denselben Experten nicht ex post replizierbar und können somit nicht in Verfahrensvergleichen berücksichtigt werden. Für die vorliegende Problemstellung lagen weder Aufzeichnungen noch ein formalisiertes Wissen zur Nachstellung subjektiver Expertenprognosen vor. Somit werden nachfolgend allein objektive Verfahren berücksichtigt.

3.2.2 Objektive Prognoseverfahren der Verteilungsschätzung

3.2.2.1 Bestimmung der empirischen Verteilungsfunktion

Im Fall eines stochastischen, stationären Bedarfs bei Risiko kann der zukünftige Bedarf durch Modelle der Verteilungsannahme geschätzt werden. Der Bedarf wird als diskrete oder stetige Zufallsvariable Y betrachtet, die beliebige Werte y_i einer bekannten, vollständig spezifizierten Wahrscheinlichkeitsverteilung annehmen kann.¹ Die Beobachtungen von y_i sind stochastisch unabhängig, und ohne serielle Autokorrelation [Domschke et al. 1997, S. 159] und repräsentieren somit eine unabhängige Folge von Zufallszahlen ohne Berücksichtigung der zeitlichen Ordnung ihres Auftretens.

Für diskrete Wahrscheinlichkeitsverteilungen beschreibt die Wahrscheinlichkeitsfunktion $P_Y(y)$ (engl. *probability mass function*, pmf) die Wahrscheinlichkeit des Auftretens von y_i anhand der Zähldichte der Verteilung von Y auf der Menge der möglichen Werte.² Somit kann jedem Wert y eine Wahrscheinlichkeit p zugeordnet werden mit der die Zufallsvariable Y diesen Wert annimmt:

$$P_Y(y) = P(Y = y) = p \quad \text{mit} \quad \sum_{i=1}^n p_i = 1. \quad (3.20)$$

Aus der Wahrscheinlichkeitsfunktion von Y wird die kumulative Verteilungsfunktion $F_Y(y)$ gebildet, die für den diskreten Fall angibt, mit welcher Wahrscheinlichkeit die Zufallsvariable des Bedarfs Y einen Wert kleiner oder gleich dem Wert y , $0 \leq Y \leq y$, annimmt:

$$F_Y(y) = P(Y \leq y) = \sum_{i=1}^y P_Y(y) \quad . \quad (3.21)$$

Die Fläche unter der kumulativen Verteilungsfunktion $F_Y(y)$, mit $F: \mathbb{R} \rightarrow [0,1]$, beträgt ebenfalls eins; sie verläuft monoton steigend und rechtsseitig stetig, mit

$$\lim_{y \rightarrow -\infty} F_Y(y) = 0 \quad \text{und} \quad \lim_{y \rightarrow +\infty} F_Y(y) = 1 \quad . \quad (3.22)$$

Im Fall einer stetigen Zufallsvariablen Y können im Gegensatz zum diskreten Fall keine Wahrscheinlichkeiten zu den einzelnen Ausprägungen der Zufallsvariablen y angegeben werden, sondern nur die Wahrscheinlichkeit $f_Y(y)dy$, dass die Werte in einem infinitesimalen Intervall dy um y liegen. Die Dichtefunktion (engl. *probability density function*, pdf) $f_Y(y)$ bezeichnet eine integrierbare Funktion, in der die Wahrscheinlichkeit, dass die stetige Zufallsvariable Y Werte zwischen a und b annimmt, durch das Integral

$$f_Y(y) = P(a \leq Y \leq b) = \int_a^b f_Y(y) dy \quad (3.23)$$

¹ Abweichend zur Notation der Materiallogistik und der Bestellmengenrechnung erfolgt die Darstellung des Bedarfs durch Y bzw. y [siehe u.a. Toporowski 1999a; 1999b; Götz 1999] anstelle von X bzw. x [siehe u.a. Naddor 1971] oder D und d [Lee und Nahmias 1993; Hillier und Lieberman 1997]. Dies erfolgt zur Anpassung der Notation an die Prognosemodelle und Verfahren, in denen x unabhängige Variablen und y die abhängige Variable des Prognosegegenstandes repräsentiert (siehe Kapitel 3).

² GÖTZ [1999] dokumentiert eine Reihe diskreter, empirischer Wahrscheinlichkeitsfunktionen $\rho(y)$ und ihre zugehörigen stetigen theoretischen Dichtefunktionen.

beschrieben wird.¹ Die Fläche unter der Dichtefunktion besitzt (wie auch im diskreten Fall) den Inhalt 1, so dass gilt:

$$\int_{-\infty}^{+\infty} f_Y(y) dy = 1 \quad . \quad (3.24)$$

Die Verteilungsfunktion $F(y)$ (engl. *cumulative distribution function*, cdf) bezeichnet analog zur diskreten Verteilungsfunktion die Wahrscheinlichkeit, dass die Zufallsvariable Y einen Wert kleiner oder gleich y annimmt [Götz 1999, S. 50]:

$$F_Y(y) = P(Y \leq y) = \int_0^y f_Y(y) dy \quad . \quad (3.25)$$

Dabei gilt

$$f_Y(y) = F'_Y(y) = \frac{dF'_Y(y)}{dy} \quad . \quad (3.26)$$

Mit Hilfe einer vollständig spezifizierten Dichtefunktion $f_Y(y)$ oder einer Verteilungsfunktion $F_Y(y)$ lässt sich der zukünftige Bedarf in der Bedarfsrechnung sowie jedes beliebige Quantil auf der Verteilungsfunktion direkt vorhersagen.²

3.2.2.2 Approximation durch theoretische Verteilungsfunktionen

Eine Bedarfsschätzung mittels Verteilungsannahme erfordert, dass die tatsächliche stationäre Wahrscheinlichkeitsfunktion $P_Y(y)$ genau spezifiziert ist, so dass $P(Y = y_i) = p_i$ für alle Werte von y_i bekannt ist. Ist dieses in empirischen Anwendungen nicht gegeben oder zu aufwendig, so wird zunächst die Verteilungsform ermittelt und anschließend ihre Momente aus der Stichprobe geschätzt, um die theoretische Verteilungsfunktion an die empirische Bedarfsverteilung anzupassen. Anstelle der empirischen Verteilungsfunktion des tatsächlichen Bedarfs wird dann vereinfacht auf eine Approximation durch eine theoretische Bedarfsverteilung zurückgegriffen.³

In Abbildung 3.11 sind die Verläufe der theoretischen Dichtefunktionen und der Verteilungsfunktionen der Normalverteilung, der Erlangverteilung und der Poissonverteilung für unterschiedliche Verteilungsparameter einander gegenübergestellt.⁴ Für kontinuierliche Variablen ist die Normalverteilung (NV) des Bedarfs $Y \sim N(\mu, \sigma^2)$ die bei Weitem populärste Approximation von empirischen Wahrscheinlichkeitsfunktionen in Theorie und Praxis [Strijbosch und Moors 2006], insbesondere in Lehrbüchern [Silver et al. 1998; Hadley und Whitin 1963].⁵ Die NV ist symmetrisch und wird durch die ersten beiden Momente der Verteilung aus Erwartungswert μ und Standardabweichung σ vollständig charakterisiert.⁶

¹ Die Grenzen a, b sind entsprechend festzulegen, so dass für $a = y$ und $b = y$ auch die Wahrscheinlichkeit ermittelt werden kann, dass die Zufallsvariable Y genau y beträgt, $y \leq Y \leq y$.

² Dabei wird in der Literatur der Lagerhaltungssysteme überwiegend die kumulative Verteilungsfunktion $F(y)$ verwendet, da daraus unmittelbar der kostenminimale Warenbestand S ermittelt werden kann, der den stochastischen Bedarf mit einer bestimmten Wahrscheinlichkeit befriedigen kann (siehe Abschnitt 2.3.3.1).

³ Neben der Anwendung theoretischer Verteilungsfunktionen können auch empirische Verteilungsfunktionen des Bedarfs aus den Beobachtungswerten der Stichprobe direkt verwendet werden, allerdings unter Aufgabe der Vorteile zur Ableitung analytischer Lösungen [Minner 2000, S. 16].

⁴ Weitere theoretische Bedarfsverteilungen finden sich bei [Silver et al. 1998, S. 272-274].

⁵ Dies ist auf die Kombination einer Reihe von wissenschaftlichen Erkenntnissen in der populären Veröffentlichung von [Silver et al. 1998] zurückzuführen, die nachfolgend kritisch erläutert werden.

⁶ Für eine Darstellung der Dichte- und Verteilungsfunktionen der Wahrscheinlichkeitsverteilungen sei auf die statistische Standardliteratur, im Kontext der Bedarfsschätzung auf GÖTZ [1999, S. 113-117] sowie MINNER [Minner 2000] verwiesen.

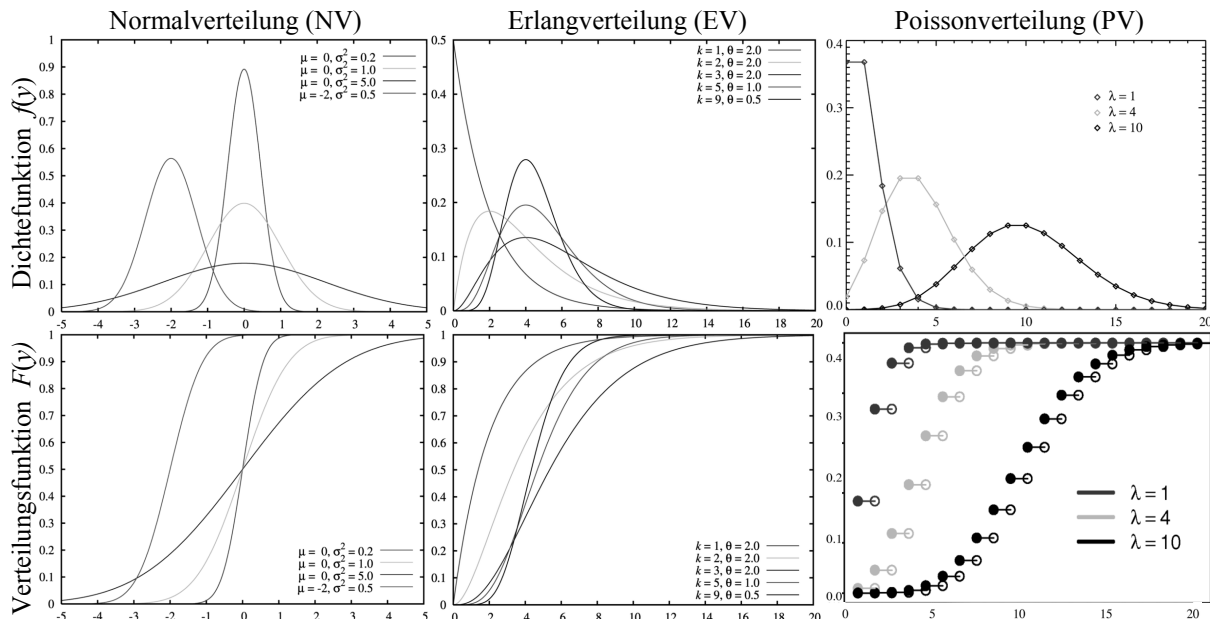


Abbildung 3.11: Dichtefunktionen und Verteilungsfunktionen der Normalverteilung, der Erlangverteilung und der Poissonverteilung

Darüber hinaus existiert eine Vielzahl theoretischer Verteilungsfunktionen, die ebenfalls sinnvoll bei der Bedarfsschätzung zum Einsatz kommen können: Die Gammaverteilung [Burgin 1972b; Burgin 1975; Burgin und Norman 1976; Tadikamalla 1978b], die Weibull-Verteilung [Tadikamalla 1978a; Jalan et al. 1996; Chakrabarty et al. 1998], die Log-Normalverteilung [Tadikamalla 1979; Silver 1980] sowie weitere stetige Verteilungen bilden sinnvolle Alternativen zur Bedarfsschätzung für Artikel mit schnellem Warenumsatz.

Von den stetigen Wahrscheinlichkeitsverteilungen zur Abbildung stationärer stochastischer Bedarfe von Waren mit hohem Umschlag werden die diskreten Wahrscheinlichkeitsfunktionen und Verteilungsfunktionen unterschieden, die zur Vorhersage von Artikeln mit langsamem stationärem Warenumsatz eingesetzt werden [Minner 2000, S. 13]. Die Poissonverteilung (PV) in Abbildung 3.11 gehört zu der Klasse der Zählmodellen, mit $Y \in \mathbb{N}$, die zur Modellierung von Häufigkeitsverteilungen seltener Ereignisse verwendet werden. Sie wird von vielen Autoren zur Bedarfsprognose mit geringen Bedarfsmengen verwendet [Axsäter 2000; Dekok et al. 1984; Forsberg 1997; Nahmias und Demmy 1982; Silver 1981]. TRUX [1972, S. 60-61] sowie GÖTZ [1999, S. 114] schlagen die Anwendung der PV bei einer Nachfrage von weniger als acht Mengeneinheiten pro Dispositionsperiode vor. Für höhere Erwartungswerte von poissonverteilten Zufallsvariablen, üblicherweise $E(Y) = \mu \geq 10$, wird als Approximation die stetige NV verwendet, und zwar unter der Annahme, dass der Bedarf $Y \sim N(\mu, \mu)$ verteilt ist [Hartung et al. 2005]. Zur Vorhersage von Zählmodellen wird auch die diskrete Negative Binomialverteilung angewendet [Agrawal und Smith 1996], sofern bei einem poissonverteiltem Bedarf die Prämisse der Identität von Erwartungswert und Varianz, $E(Y) = V(Y) = \mu$, verletzt ist [Götz 1999, S. 115-117; Sherbrooke 2004, S. 60].¹ TEMPELMEIER [2003, S. 94] rät zur Verwendung einer Poissonverteilung, sofern die Standardabweichung um weniger als 10% von der Wurzel des Mittelwertes abweicht. Für Zeitreihen mit sporadischem Bedarf und Perioden ohne Bedarf kann eine Zerlegung des Periodenbedarfs in seine Komponenten der Anzahl der Bedarfe und der Bedarfsmenge je Auftrag oder der Bedarfsmenge und der Zeitspanne bis zum nächsten

¹ Als weitere Ursache nennt SHERBROOKE [2004, S. 61] instationäre Bedarfe der Poissonverteilung mit Trend, welche ebenfalls zu einer höheren Varianz führen. Dies ist jedoch als Fehlspezifikation des Bedarfsmodells zu interpretieren und erfordert die Abbildung durch stochastische Prognosemodelle.

Bedarfsereignis erfolgen [Tempelmeier 2003, S. 90-91]. WEDEKIND [Wedekind 1968] unterstellt eine Weibull-Verteilung der Bedarfsabstände, während SHERBROOKE [2004, S. 60-62] in empirischen Verteilungen von Ersatzteilbedarfen eine Exponentialverteilung der Bedarfsabstände mit einem poissonverteilten Bedarfsmittelwert feststellt, der jedoch vielfach instationär ist.¹ Für kleine Stichproben des Bedarfs mit $n < 30$ wird darüber hinaus die Verwendung der t-Verteilung postuliert [Götz 1999, S. 117].

STRIJBOSCH und MOORS [2006, S. 805] sowie [Sherbrooke 1992; Toporowski 1998, S. 1146] stellen fest, dass neben frühen Veröffentlichungen [u.a. Burgin 1972a; Donaldson 1984] auch eine Vielzahl neuerer Veröffentlichungen aus Gründen der Vereinfachung ausschließlich einen normalverteilten Bedarf unterstellen [siehe u.a. Arto und Pylkkanen 1999; Chen und Chuang 2000; Geunes und Zeng 2001; Alstrom 2001]. Zudem werden aus Gründen der analytischen Vereinfachung in der Berechnung von Bestellmengen auch diskrete, partiell diskrete oder partiell kontinuierliche Bedarfe vielfach durch eine kontinuierliche Verteilungsform angenähert [Lee und Nahmias 1993, S. 26; Hillier und Lieberman 1997, S. 625].² SILVER und PETERSON [Silver und Peterson 1985, S. 140] begründen die Verwendung der NV neben der analytische Vereinfachung der Berechnungen sowohl theoretisch als auch empirisch.³ Nach dem zentralen Grenzwertsatz konvergiert die Summe einer großen Zahl unabhängiger, identisch verteilter Zufallsvariablen zur NV, was ihren vielfachen Einsatz theoretisch rechtfertigt. Ferner legitimieren sie die NV durch die hohe Anpassungsgüte an empirische Bedarfsverteilungen: „empirically the normal distribution usually provides a better fit to data than most other suggested distributions“ [Silver und Peterson 1985, S. 140]. NADDOR [1978] bestätigt, dass die optimalen Bestände und Kosten in Lagerhaltungssystemen überwiegend von der Standardabweichung und dem Mittelwert des Bedarfs abhängig sind und nicht von der spezifischen Form der Verteilung. Die Studie von LAU und ZAKI [1982] über mehrere Bedarfsverteilungen bestätigt dies. TYWORTH und O’NEILL [1997] untersuchen die Zulässigkeit der Approximation empirischer Bedarfe durch die theoretische NV anhand der Kosten, die durch die Fehlspezifikation im Vergleich zu einer theoretisch korrekten Vorgehensweise entstehen. Ihre Studie belegt die vergleichbare Genauigkeit der NV auf sieben typischen Bedarfsverteilungen von schnell drehenden Konsumgütern (engl. *fast moving consumer goods*). Dies legt nahe, die NV einheitlich zu verwenden.

Dagegen wurde in neueren empirischen Untersuchungen vielfach eine asymmetrische, nicht-normalverteilte Form der Bedarfsverteilung nachgewiesen [Lau und Lau 2003; Tadikamalla 1984; Wang und Subba Rao 1992; Fricker und Goodhart 2000]. Diese Erkenntnis gilt insbesondere für Güter mit geringem Warenumsatz und steht so nur bedingt im Widerspruch

¹ Da es sich beim sporadischen Bedarf streng genommen um einen instationären Bedarfsverlauf handelt, müsste dieser im folgenden Abschnitt betrachtet werden; er wird jedoch vernachlässigt. WEDEKIND [1968] unterstellt dabei instationäre Bedarfsmengen durch die Anwendung einer einfachen Exponentiellen Glättung zur Ermittlung stochastischer Bedarfsmittelwerte, während CROSTONs Methode [Croston 1972] sowohl instationäre Bedarfsmengen als auch instationäre Bedarfsabstände unterstellt. SHERBROOKE [2004] modelliert dagegen stationäre Bedarfsmengen und Bedarfsabstände durch Anwendung unterschiedlicher Verteilungsannahmen und postuliert so stationäre Verteilungen der Bedarfszeitpunkte und Bedarfsmengen ohne zeitliche Veränderung. Diese Unterscheidung geht jedoch erneut von stationären Verteilungen der Bedarfszeitpunkte und Bedarfsmengen ohne mögliche Veränderungen der Verteilungsform aus. Daher erscheint auch hier eine Prognose, zum Beispiel durch Verfahren der Exponentiellen Glättung für sporadischen Bedarf nach CROSTON, sinnvoller, die instationäre Veränderungen im Zeitablauf zulässt und für konstante Bedarfsverläufe und -zeiten ebenfalls valide Vorhersagen erzeugt. Dabei ist anzumerken, dass viele Autoren wie zum Beispiel MINNER [2000] statische und dynamische Modelle nicht unterscheiden, und diese Betrachtungsweise daher nicht berücksichtigen.

² Dies ist nach HILLIER und LIEBERMAN [1997, S. 625] dann zulässig, wenn der Bedarf eine große Bandbreite an möglichen Werten besitzt, da somit die Annäherung durch eine kontinuierliche Verteilungsfunktion nur zu kleinen Unterschieden in den numerischen Werten der optimalen Lagermengen führt.

³ Sie weisen jedoch auch auf die alternative Verwendung der Exponential-, Gamma-, Poisson- oder einer negativen Binomialverteilung hin, sofern dies für Waren mit langsamem Artikelumschlag angemessen erscheint [Silver und Peterson 1985, S. 140]

zu den früheren Studien, da diese nahezu ausschließlich schnelldrehende Konsumgüter berücksichtigt [Lau und Lau 2003]. TADIKAMALLA [1984] stellt in einer Vergleichsstudie von fünf Wahrscheinlichkeitsfunktionen für Bedarfe mit unterschiedlichen Variationskoeffizienten und Servicegraden fest, dass die Gültigkeit der NV ausschließlich für Bedarfe mit niedrigem Variationskoeffizienten gilt und sonst hohe Fehlerkosten hervorrufen kann. LAU und LAU [Lau und Lau 2003] stellen ebenfalls fest, dass die Anwendung der NV selbst für typische Bedarfsverläufe zu hohen Kosten durch die Fehlspezifikationen der Wahrscheinlichkeitsfunktion führen kann. BARTEZZAGHI et al. [1999] zeigen, dass die funktionale Form der Wahrscheinlichkeitsverteilung hinsichtlich Asymmetrie und Multimodalität einen signifikanten Einfluss auf die Bestandshöhen hat und eine Annahme der NV diesem nicht immer genügt.

Somit entstehen systematische Fehler bei der Auswahl theoretischer Verteilungsfunktionen, die durch eine statistische Bestimmung der tatsächlichen Wahrscheinlichkeitsverteilung behoben werden können.¹ Die valide Auswahl einer angemessenen Verteilungsfunktion erfordert die Berechnung etablierter statistischer Verteilungstest des Chi-Quadrat-Anpassungstest oder Kolmogorow-Smirnow-Anpassungstest auf der Stichprobe des Bedarfs.² SHERBROOK [2004, S. 92-95] beschreibt zudem einen erweiterten Chi-Quadrat-Anpassungstests für Poisson- und Binomialverteilten Bedarf.

3.2.2.3 Schätzung der Verteilungsparameter

Unter Kenntnis oder Annahme einer theoretischen Verteilungsfunktion $F_Y(y)$ des Bedarfs werden die Kenngrößen ihrer zentralen Momente von Erwartungswert, Varianz, Schiefe und Wölbung aus einer Stichprobe vergangener Bedarfswerte y_i geschätzt und zur Schätzung der Verteilungsfunktion in der Grundgesamtheit verwendet.³

Dabei entstehen weitere Fehlerquellen für die anschließende Bestimmung der Bestellmengen. Die Mehrheit der Untersuchungen beschränkt sich auf eine Betrachtung der ersten beiden Momente (Erwartungswert und Varianz), unter Vernachlässigung der dritten und vierten Momente (Schiefe und Wölbung) der Verteilung in der Stichprobe [Minner 2000, S. 15].⁴

Der Erwartungswert $E(Y)$ der Zufallsvariablen des Bedarfs Y bezeichnet jenen Wert, der unter Berücksichtigung aller Ausprägungen y der Zufallsvariablen Y und der Wahrscheinlichkeit ihres Auftretens $p_Y(y)$ erwartet wird.

¹ MINNER [2000] weist ferner darauf hin, dass Fehler in der Bestimmung der Verteilungsfunktion und ihrer Momente für geringe Stichproben zu erheblichen Verzerrungen führen können, die in der Ermittlung der Bestellmengen durch Risikozuschläge zur Berücksichtigung der Unsicherheit in der Bedarfschätzung und Modellselektion zu berücksichtigen sind. Ähnliches gelte auch bei neuen Produkten. STRIJBOSCH und MOORS [2005] zeigen zudem, dass die zusätzlichen Unsicherheiten in der Bestimmung der Momente der Wahrscheinlichkeitsfunktion anhand der Stichprobe immer zu einem Sicherheitsaufschlag zur Berücksichtigung der zusätzlichen Unsicherheit in der Modellierung führen.

² Beispiele zu einer Überprüfung auf normal- oder poissonverteiltem Bedarf anhand von Kolmogorow-Smirnow-Anpassungstests finden sich in Abschnitt 6.3. Die Forderung von SILVER und PETERSEN [1985, S. 140] nach einfachen grafischen Tests der empirischen Verteilungsfunktion auf Normalverteilung durch P-P- und Q-Q-Diagramme ist auf die geringe Verfügbarkeit von Computerressourcen zum damaligen Zeitpunkt der Veröffentlichung zurückzuführen.

³ Diese entsprechen den Kenngrößen der Parameter von Häufigkeitsverteilungen in der deskriptiven Statistik.

⁴ Die Schiefe der Verteilung wird durch die Skewness angegeben, welche eine links- oder rechtsschiefe Form der Häufigkeitsverteilung beschreibt. Die Wölbung der Verteilung, gemessen durch die Kurtosis, zeigt für Werte von Kurt=0 eine mesokurtische Wölbung im Sinne der Normalverteilung an. Werte größer null, Kurt>0, bezeichnen eine leptokurtische oder supergaußförmige Verteilung, Werte kleiner als null, Kurt<0, identifizieren dagegen eine platykurtische oder subgaußförmige Wölbung der Verteilung. LAU und ZAKI [1982] postulieren hingegen, dass die Fehler aus der Vernachlässigung von Schiefe und Wölbung der Wahrscheinlichkeitsfunktion gering sind. Dies ist jedoch im Kontext ihrer Erkenntnisse der universellen Einsetzbarkeit der NV zu verstehen, da diese keine Anpassung weiterer Parameter erfordert.

$$E(Y) = \sum_{y=0}^{\infty} y p_Y(y) \quad , \text{ bzw. im stetigen Fall mit } E(Y) = \int_{-\infty}^{\infty} y f_Y(y) dy \quad . \quad (3.27)$$

Das empirische arithmetische Mittel \bar{y} der Stichprobe gilt als unverzerrter Schätzer für den Mittelwert der Grundgesamtheit μ :

$$E(Y) = \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad , \text{ da } E(\bar{y}) = \mu \quad . \quad (3.28)$$

Der Erwartungswert wird anschließend als Prognose des Bedarfs \hat{y}_{t+h} in zukünftigen Perioden $t+h$ für beliebige Prognosehorizonte $h = 1, 2, \dots, H$ verwendet:

$$\hat{y}_{t+h} = \bar{y} \quad . \quad (3.29)$$

Ferner wird die Streuung des Bedarfs anhand der empirischen Standardabweichung s bzw. der empirischen Varianz s^2 der Stichprobe als Schätzer der Standardabweichung σ und der Varianz $\text{var}(Y) = \sigma^2$ der Grundgesamtheit ermittelt, mit

$$\text{var}(Y) = s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 \quad \text{und} \quad s = \sqrt{s^2} \quad . \quad (3.30)$$

Da die empirische Standardabweichung ebenfalls im Sinne einer Vorhersage der zukünftigen Streuung der Nachfrage zu verstehen ist und zudem im Zeitablauf veränderlich sein kann, wird sie nachfolgend mit $\hat{\sigma}_{t+h}$ bezeichnet.¹ Nach Simulationsstudien von ERHARDT [1979] wird die Dispositionsgüte durch die Verwendung der empirischen anstelle der wahren Standardabweichung σ_y nur geringfügig verschlechtert. Eine Verwendung nur der ersten beiden Momente impliziert eine symmetrische Verteilung im Sinne der NV, die bereits hinreichend durch Erwartungswert und Varianz bestimmt wird, da der mittlere Wert des Median und der häufigste Wert des Modus mit dem Mittelwert übereinstimmen. SHORE [1999, S. 477] zeigt hingegen, dass die übliche Stichprobengröße für eine valide Identifikation der dritten und vierten Momente einer theoretischen Verteilung vielfach nicht ausreicht und nur unter hohen Fehlern möglich ist. Somiz kann eine Auswahl der Verteilungsfunktion und Bedarfsschätzung unter Verwendung von Erwartungswert und Varianz und Vernachlässigung höherer Momente empirisch begründet sein.²

Die Bestimmung der Momente erlaubt eine Überprüfung der Validität der vorangegangenen Auswahl der Verteilungsform anhand heuristischer Regeln. Dazu wird vielfach der Variationskoeffizient v verwendet, der das Verhältnis der empirischen Standardabweichung s zum Mittelwert \bar{y} bezeichnet:

$$v = \frac{s}{\bar{y}} \quad . \quad (3.31)$$

Die Studien von [Naddor 1978; Lau und Zaki 1982; Tyworth und O'Neill 1997], die eine grundsätzliche Anwendung der NV unterstützen, beruhen auf Bedarfsverteilungen mit einem hohen Warenumschlag und somit einem niedrigen Variationskoeffizienten [Lau und Lau

¹ In Übereinstimmung mit SILVER et al. [1998]

² SHORE [1999] entwickelt als Lösungsansatz eine flexible Wahrscheinlichkeitsfunktion mit vier Parametern in Anlehnung an die Verteilung nach RAMBERG et al. [1979] bzw. nach SCHMEISER und DEUTSCH [1977], die jedoch allein durch Bestimmung der ersten beiden Momente spezifiziert werden kann.

2003]. SILVER et al. [1998, S. 273] schlussfolgern daraus, dass solange ein Variationskoeffizient unter $\nu < 0,45$ vorliegt, die theoretische NV zur Bedarfsschätzung auch für Zeitreihen ohne annähernd normalverteilten Bedarf verwendet werden kann.¹ Für Variationskoeffizienten mit $\nu > 0,5$ sind hingegen andere Verteilungsfunktionen zu wählen, da für einige Kombinationen von (μ, σ) unzulässige, negative Bedarfswerte mit einer positiven Wahrscheinlichkeit auftreten können. MINNER [2000, S. 16-17] untersucht die Wahrscheinlichkeit negativer Bedarfswerte in Abhängigkeit von den Variationskoeffizienten und bestätigt die Erkenntnisse anderer Autoren [u.a. Tijms und Groenevelt 1984], wonach die Anwendung der NV nur für Variationskoeffizienten $\nu \leq 0,5$ zulässig ist. Dies macht eine Anwendung der NV ungeeignet für Waren mit geringen Bedarfsmengen und langsamem Warenumsatz, die üblicherweise einen hohen Variationskoeffizienten aufweisen.² Entsprechend postulieren unterschiedliche Autoren die Verwendung alternativer, stetiger Verteilungen zur Vermeidung negativer Bedarfe, einschließlich der Gamma-Verteilung [Burgin 1972a; Moors und Strijbosch 2002]. STRIJBOSCH und MOORS [Strijbosch und Moors 2006] zeigen zwei Modifikationen der NV zur Vermeidung negativer Werte, weisen jedoch auf die Überlegenheit der Gammaverteilung hin. TIJMS [1994, S. 358-359] schlägt die Verwendung einer Mixed-Erlangverteilung vor, die durch eine Überlagerung mehrerer Erlangverteilungen (EV) in Abbildung 3.11 jede beliebige Wahrscheinlichkeitsverteilung hinreichend approximieren kann und die analytisch lösbar ist [Minner 2000]. Ihre Anwendung beschränkt TIJMS [1994, S. 360] jedoch auf Variationskoeffizienten mit $\nu \leq 1,0$ unter Anwendung der Hyper-Exponentialverteilung für $\nu > 1,0$. Aufgrund von Schätzproblemen für geringe ν beschränkt MINNER [2000, S. 19] ihre Anwendung auf das Intervall $0,5 \leq \nu < 1,0$ unter Verwendung der NV für $\nu \leq 0,5$.

Tatsächlich ergeben sich vielfältige Überschneidungen und Widersprüche bei den heuristischen Regeln, die auf eine allgemeine Unschlüssigkeit bei der Auswahl einer Verteilungsfunktion und die Grenzen einer derartigen Bestimmung einfacher heuristischer Regeln hinweisen.³ LAU und LAU stellen daher fest: „Given the hardware and software available today, instead of trying to search for a necessarily complicated rule to determine whether the normal approximation is safe in any given situation, perhaps it is time to consider ... trying to estimate more accurately the actual lead time demand distribution“ [Lau und Lau 2003, S. 165]. Demzufolge sollte die Analyse der Variationskoeffizienten eine Bestimmung der Wahrscheinlichkeitsverteilung nur ergänzen, nicht aber ersetzen.

Unter Zusammenfassung der vorangegangenen Schritte lässt sich die Bedarfsschätzung mit Verteilungsannahme als prototypische Prozesskette bestimmen, die sich als Teilprozess in die Warendisposition in Abbildung 2.5 einfügt.⁴

¹ Da nach TYWORTH und O'NEILL [1997] alle typischen Bedarfsverläufe einen solchen Variationskoeffizienten aufweisen, erschien eine einheitliche Gültigkeit und Verwendung der NV als begründet.

² Zudem kann das Integral der Verteilungsfunktion nicht analytisch gelöst werden und erfordert entweder eine Tabellierung oder eine Verwendung von Approximationen [siehe u.a. Abramowitz und Stegun 1972].

³ SHERBROOKE [2004, S. 261] weist beispielsweise darauf hin, dass bei poissonverteiltem Bedarf der Variationskoeffizient aufgrund der Identität von Mittelwert und Varianz theoretisch $\nu \leq 0,5$ betragen sollte, was im Gegensatz zu den Empfehlungen einer Nutzung der NV steht. Er berechnet anstelle des Variationskoeffizienten zudem das Verhältnis von Varianz zu Mittelwert (engl. *variance-to-mean ratio*), welches für poissonverteilten Bedarf ein Verhältnis von 1 angibt.

⁴ Sie ist an den Prozess der stochastischen Bedarfsprognose aus Abschnitt 3.4.2 angelehnt, der dort näher beschrieben ist. Der Prozess der Bedarfsschätzung wird auch für die Residuen der stochastischen Prognoseverfahren bei instationärem Bedarf durchlaufen, jedoch unter Prüfung der Prognoseresiduen anstelle der tatsächlichen Bedarfsverteilung (vgl. Abschnitt 2.3.3).

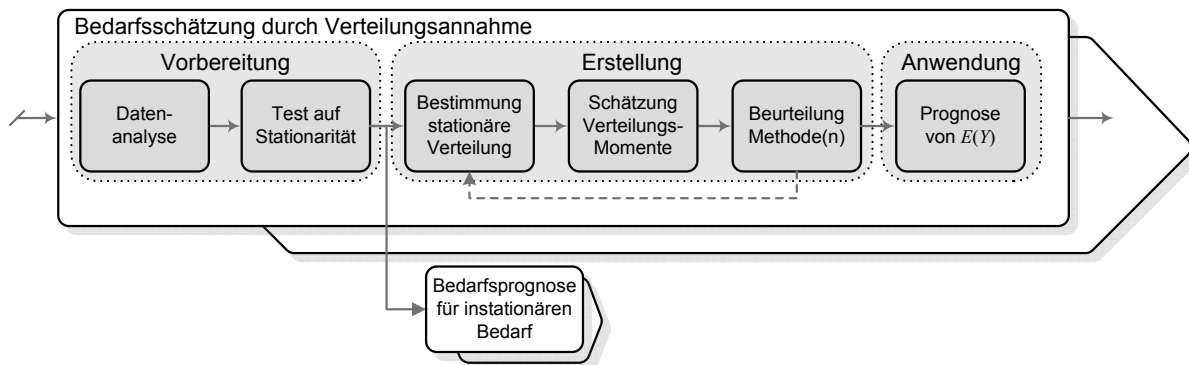


Abbildung 3.12: Prozesskette der Bedarfsrechnung mit Verteilungsannahme

Die Bedarfschätzung mittels Verteilungsannahme bedingt direkte Vorteile für die nachfolgend betrachtete Bestellmengenrechnung.¹ Die kumulative Verteilungsfunktion $F_Y(y)$ erlaubt die unmittelbare Bestimmung von optimalen Warenbeständen und Bestellmengen, welche die zu erwartende Nachfrage y mit einer vorgegebenen Sicherheitswahrscheinlichkeit $S = 1 - \alpha$ erfüllen. GÖTZ [1999, S. 113-116] zeigt u.a. die Berechnungsvorschriften für unterschiedliche Verteilungsfunktionen auf, jedoch ohne die Modelle der Bedarfsrechnung klar von jenen der Bestellmengenrechnung abzugrenzen. Zudem berücksichtigt er keine Tests auf Stationarität der Zeitreihen, sondern unterstellt für alle Zeitreihen des Absatzes in Supermärkten einen stationären stochastischen Verlauf und verwendet konsistent Ansätze der Bedarfschätzung durch Verteilungsannahme.

Die Darstellung der Modellbildung in der Prozesskette hebt hervor, dass eine Bedarfschätzung durch Verteilungsannahme nur für Zeitreihen mit stationärem Mittelwert und homoskedastischer Varianz zulässig ist, was jedoch in der Handelspraxis nur einen seltenen Sonderfall empirischer Bedarfsverteilungen repräsentiert. Ist die Prämisse der Stationarität durch Veränderungen der Momente der Verteilung nicht erfüllt, müssen stochastische Modelle der Bedarfsprognose für instationären Bedarf angewendet werden.

3.2.3 Objektive Prognoseverfahren der Zeitreihenanalyse

3.2.3.1 Zeitreihen als Grundlage der Prognose

Verfahren der Zeitreihenanalyse prognostizieren zukünftige Realisationen der abhängigen Variablen \hat{y}_{t+h} allein aus zeitlich verzögerten Beobachtungen derselben endogenen Variablen y_{t-l} mit $l \geq 0$. Die Beobachtungen liegen in Form von Zeitreihen als zeitlich geordnete Folge y_t von Beobachtungen einer Größe des jeweils gleichen Sachverhaltes vor. Für jeden Zeitpunkt t einer endlichen, diskreten Parametermenge $t = 1, 2, \dots, T$ liegt zu äquidistanten Zeitpunkten genau eine Beobachtung vor [Schlittgen und Streitberg 1987, S. 1]², wobei y_t den zum Zeitpunkt t beobachteten Wert der Zeitreihe spezifiziert. Die Beobachtungen werden als Realisationen eines stochastischen Prozesses Y_t aufgefasst, der alle Ausprägungen der Zufallsvariablen y_t umfasst und sich als statistischer Vorgang nach den Gesetzen der Wahrscheinlichkeit zeitlich vollzieht [Hansmann 1983, S. 53].

Die stochastische Bedarfsprognose muss unterschiedliche, regelmäßige und unregelmäßige Komponenten instationärer Bedarfsverläufe im Zeitablauf berücksichtigen. Da diese nachhal-

¹ Darüber hinaus hat die Bedarfschätzung auch Vorteile für die Bedarfsrechnung selbst. Die vollständige Spezifikation der Bedarfsverteilung erlaubt neben der Bestimmung des Erwartungs- und Mittelwertes im Sinne einer Punktprognose auch die problemlose Ableitung von Konfidenzintervallen der Prognose als Quantile der Verteilungsfunktion $F(y)$ [Götz 1999, S. 113], allerdings nur für stationäre Bedarfsverläufe.

² Sofern kontinuierliche Messungen vorliegen, wird für die Analyse eine Diskretisierung vorgenommen. Viele Verfahren sind zudem auch auf Beobachtungen zu unregelmäßigen Zeitpunkten bzw. zu ungleichen Zeitabständen Δt anwendbar.

tig die Modellbildung zur Bedarfsprognose bestimmen, werden sie in Abbildung 3.13 zunächst dargestellt und in Abschnitt 3.2.2.1 eingehend diskutiert.

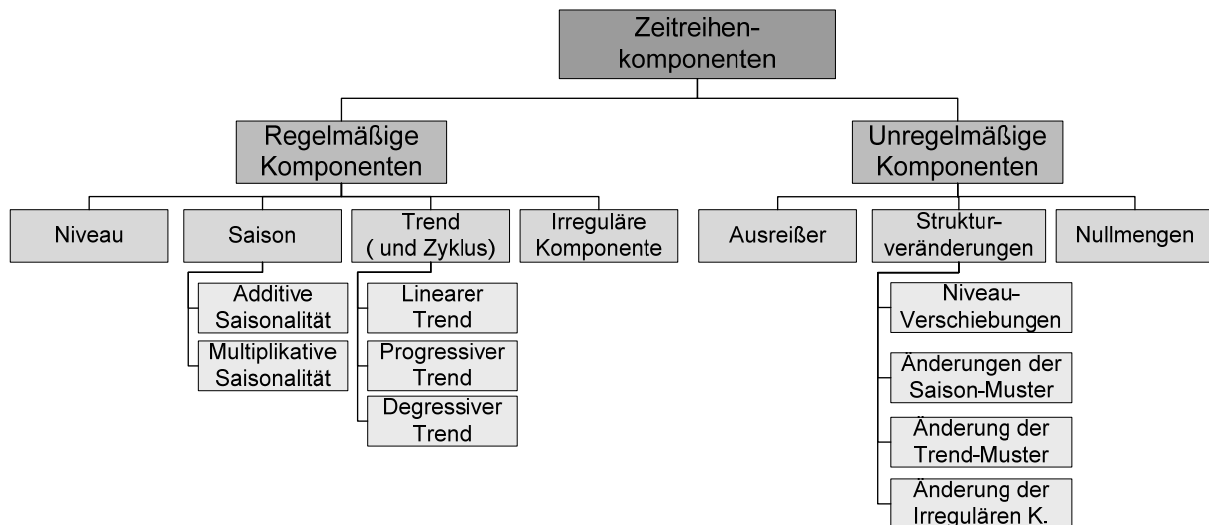


Abbildung 3.13: Regelmäßige und unregelmäßige Komponenten von Zeitreihen

Zeitreihen können als Überlagerung einer oder mehrerer regelmäßiger und unregelmäßiger Komponenten aufgefasst werden, die jedoch nicht einzeln beobachtet werden können.¹ Die regelmäßigen Komponenten umfassen das mittlere Niveau (engl. *level*) der Zeitreihe L_t , unterschiedliche Formen der Trendkomponenten T_t und der Saisonkomponenten S_t und die irreguläre Komponente E_t [Makridakis et al. 1998, S. 24-26].² Die Kombination der regelmäßigen Komponenten bildet archetypische Bedarfsmuster mit stationärem, saisonalem, trendbehaftetem sowie trend-saisonaalem Verlaufs. Verfahren der Zeitreihenprognose identifizieren diese regelmäßigen Komponenten in den Zeitreihen mit:

$$y_t = f(L_t, S_t, T_t, E_t) \quad (3.32)$$

und extrapolieren sie in die Zukunft. Die Trendkomponente repräsentiert eine kontinuierliche Veränderung des Niveaus einer Zeitreihe, gemessen durch einen Anstiegs oder die Verringerung des lokalen Mittelwerts. Die zyklische Komponente wird auf langfristige Schwankungen ökonomischer Größen um den Trend zurückgeführt.⁴ Die saisonale Komponente beinhaltet periodische Fluktuationen konstanter Länge, die direkt oder indirekt mit dem Jahreskalender zusammenhängen [Makridakis et al. 1998, S. 82].⁵ Sie können in unterschiedlichen, sich gegenseitig überlagernden Saisonalitäten halbjährlich, quartalsweise, monatlich, wöchentlich,

¹ Die Interpretation von Zeitreihen und ökonomischen Größen als Kombination unbeobachtbarer Komponenten hat eine lange Tradition in der Ökonomik. NERVOLVE et al. [1995, S. 1-20] geben einen historischen Überblick und leiten daraus die Verfahrensklasse der „unobservable components“ als Erweiterung der ARIMA-Verfahren ab [Nerlove et al. 1995]. Die Komponenten der Zeitreihen werden synonym als „systematische“ und „unsystematische“ Komponenten bezeichnet.

² Die zyklische Komponente wird vielfach mit der Trendkomponente kombiniert. Da die nachfolgende Untersuchung auf kurzfristige, operative Planungshorizonte beschränkt ist wird der Zyklus vernachlässigt.

³ Die konventionelle Notation aus [Makridakis et al. 1998] wird hinsichtlich des Niveaus der Zeitreihe erweitert, da die Intensität einzelner multiplikativer Komponenten eines Modells unmittelbar durch das Niveau der Zeitreihe bzw. Niveaushiftungen determiniert wird.

⁴ MAKRIDAKIS et al. fassen Trend- und Zyklus-Komponente zur Trend-Zyklus-Komponente zusammen, da sich die Zuordnung der nicht-saisonalen Schwankungen zur einen oder anderen Komponente lediglich durch das zeitliche Betrachtungsintervall ergibt [Makridakis et al. 1998, S. 82]. Ein Zyklus äußert sich z. B. in Form von konjunkturellen Schwankungen mit Perioden von 5 bis 7 Jahren.

⁵ LEWANDOWSKI [1974] nennt darüber hinaus die Extrasaisonalität bzw. Kalendersaisonalität, die alle durch die Kalenderstruktur bedingten Veränderlichkeiten, z. B. die unterschiedliche Anzahl von Kalender- oder Arbeitstagen in den Monaten eines Jahres oder die Anzahl und Verteilung der Kalenderwochen im Jahr, beschreibt.

täglich, stündlich oder in kürzeren Intervallen fester Dauer entstehen [Hüttner 1986].¹ Aufgrund der vergleichsweise kurzen Produktlebenszyklen und der Dominanz kurzfristiger Absatzprognosen im Handel kommt der Analyse von Trend- und Saisoneinflüssen eine besondere Bedeutung zu [Südkamp 1995, S. 12; Werner 2000, S. 18].

Die irreguläre Komponente E_t ergibt sich als die verbleibende Restgröße einer um ihre Trend-, Zyklus- und Saisonkomponenten bereinigten Zeitreihe [Hruschka 1980, S. 1, 48-49].² Sie beinhaltet alle unregelmäßigen Sondereinflüsse auf die grundlegende Struktur der Zeitreihe, z. B. durch Werbeaktionen, Strukturbrüche und andere Störungen als Summe aller deterministischen oder stochastischen Einflüsse. Dabei zählt sie dennoch zu den regulären Komponenten einer Zeitreihe: gemäß der Zeitstabilitätshypothese wird unterstellt, dass die im bisherigen Verlauf der Zeitreihe aufgedeckten Gesetzmäßigkeiten einschließlich der funktionalen Form der irregulären Komponente für die Zukunft weiter gültig sind [Hansmann 1983, S. 11]. Veränderungen von E_t , im Sinne der Heteroskedastizität (vgl. Abschnitt 2.2.3.1) gelten dagegen als unregelmäßige Komponenten. Für Modelle auf der basis der linearen Regression wird ein Erwartungswert der Zufallskomponente von null mit konstanter Varianz angenommen [Mertens und Deistler 1994, S. 30] und eine Normalverteilung der Residuen unterstellt. Die Erklärungsmodelle der Zeitreihen können einzelne Komponenten umfassen, z. B.

$$y_t = L_t + E_t \quad , \quad (3.33)$$

oder Kombinationen der Komponenten berücksichtigen. Die Erklärungsmodelle werden ferner nach der Art funktionalen Verknüpfung der Komponenten unterschieden. Grundsätzlich wird eine additive Verknüpfung mit der Annahme voneinander unabhängiger Komponenten, repräsentiert durch

$$y_t = L_t + S_t + T_t + E_t \quad , \quad (3.34)$$

von einer multiplikativen Verknüpfung – unter Interdependenz der Höhe der einzelnen Komponenten untereinander – unterschieden, im Sinne von

$$y_t = L_t \cdot S_t \cdot T_t \cdot E_t \quad . \quad (3.35)$$

Ferner sind gemischte Verknüpfungen als Kombination denkbar [Makridakis et al. 1998, S. 84-85; Lo 1979, S. 106], die individuell je Zeitreihe zu bestimmen sind. Die Überlagerung der zeitlichen Ausprägung aller regelmäßigen Komponenten L_t , S_t , T_t mit unterschiedlichen funktionalen Verknüpfungen spezifiziert den empirischen Gesamtverlauf einer Zeitreihe [Lo 1979, S. 105]. In Anlehnung an die Klassifikation nach PEGELS [1967] und die Erweiterung nach GARDNER [Gardner und MacKenzie 1989] können die Komponenten zu neun charakteristischen Zeitreihenmustern kombiniert werden, die in Abbildung 3.14 einschließlich der Zufallskomponente E_t ohne Ausreißer und Niveaueverschiebungen dargestellt sind.

Die neun archetypischen Zeitreihenmuster nach PEGELS (1969) ergeben sich durch die Kombination der saisonalen Komponenten der Zeitreihe (ohne Saisonalität E , additive Saisonalität S_A , multiplikative Saisonalität S_M) in den Spalten mit den Ausprägungen der Trend-Komponenten des linearen Trends T_L und des progressiven Trends T_P in den Zeilen. So

¹ Sie bilden somit unterschiedliche, sich überlagernder Muster eines Wochentags in der Woche, Tag im Monat, Woche im Monat, Monat im Quartal, Monat im Jahr, Quartal im Jahr oder ähnlicher Regelmäßigkeiten ab. Saisonale Komponenten erlauben die Abbildung ihrer Periodizität durch die Bildung periodischer Indizes, welche auch eine Desaisonalisierung zulassen.

² Die irreguläre Komponente wird auch als „Residualgröße“ oder „Reststreuung“ bezeichnet. In den Prognosemodellen in den Ingenieurwissenschaften wird sie auch als Rauschen (engl. *noise*) bezeichnet.

ergeben sich vier charakteristische Zeitreihenmuster: stationärer Verlauf (L.E), additive und multiplikative Trendverläufe (L.T_L.E, L.T_P.E), additive und multiplikative saisonale Verläufe (L.S_A.E und L.S_M.E)¹ und eine Kombination trend-saisonaler Verläufe mit additiver (L.S_A.T_L.E, L.S_A.T_P.E) bzw. multiplikativer Verknüpfung (L.S_M.T_L.E, L.S_M.T_P.E). Die unterschiedliche Interaktion der Komponenten gemäß ihrer Kombination im Erklärungsmodell wird in der grafischen Abbildung evident: Bei additiver Verknüpfung von Saison und Trend bleibt die Höhe der saisonalen Amplitude unabhängig vom Niveau der Zeitreihe, während bei multiplikativer Verknüpfung die saisonale Amplitude mit zunehmendem Niveau der Zeitreihe wächst, z. B. durch einen positiven Trend oder Niveauverschiebungen [Götz 1999, S. 112]. GARDNER et al. [Gardner und Dannenbring 1980; Gardner und MacKenzie 1989] erweitern diese Klassifikation um drei weitere Klassen degressiver Trends ohne Saisonalität (L.T_D.E) sowie mit additiver (L.S_A.T_D.E) und multiplikativer Saisonalität (L.S_M.T_D.E), welche er als gedämpften Trend (engl. *dampened trend*) bezeichnet.² Diese zwölf Zeitreihenmuster beschreiben die typischen Verläufe aller regelmäßigen Zeitreihen.³

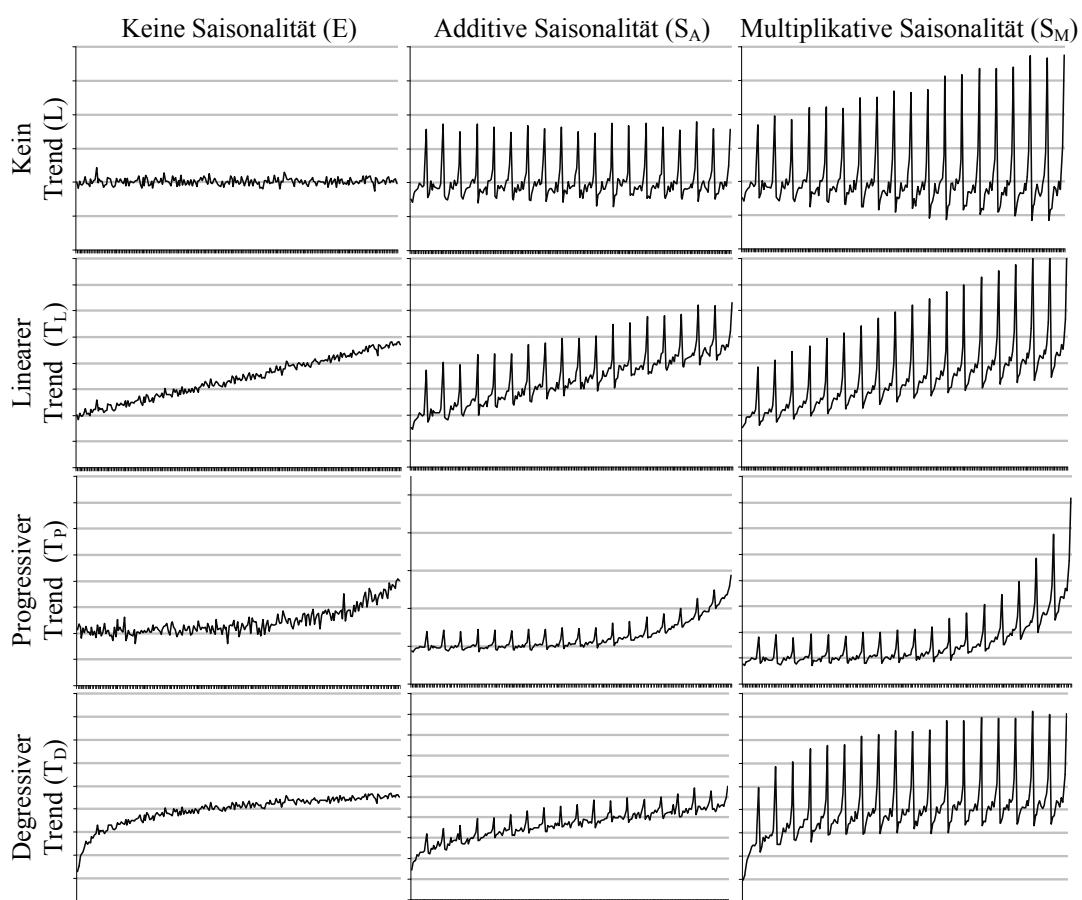


Abbildung 3.14: Zeitreihenmuster nach der PEGELS-GARDNER-Klassifikation⁴

¹ Die Abbildung der Zeitreihe ohne Trend mit multiplikativer Saisonalität nach Pegels kann theoretisch keine wachsende Saisonalität mit zunehmender Zeit aufweisen. Eine zunehmende Amplitude der Saisonalität ist hingegen bei Niveauverschiebungen und lokalen Trends möglich. Sie wird hier der Vollständigkeit halber und in Anlehnung an die ursprüngliche Klassifikation wiedergegeben.

² Die Abbildung aller 12 Zeitreihenmuster erlaubt die Abbildung aller Phasen des Produktlebenszyklus einschließlich Relaunch von Produkten (vgl. Abschnitt 3.1.3.1).

³ Die Zeitreihenmuster bei regelmäßigem Zeitreihenverlauf nach PEGELS und GARDNER können durch die Formen der unregelmäßigen, sporadischen Zeitreihenverläufe ergänzt werden, die sich durch einen Absatz von $y_t=0$ Einheiten in der Mehrzahl der Perioden und ohne systematische Muster des Trends oder der Saisonalität auszeichnen. Da sie die Modellierung und Prognose durch spezifische Verfahren, beispielsweise nach CROSTON [1972] erfordern, wird auf eine Abbildung hier jedoch verzichtet.

⁴ Eigene Darstellung in Anlehnung an [Pegels 1967, S. 312-314; Gardner und MacKenzie 1989; Makridakis et al. 1998, S. 138]

Das Konzept der Stationarität wurde bereits in Abschnitt 2.2.3.1 formal hergeleitet. Dabei entspricht allein der stationäre Verlauf L.E den Anforderungen der Stationarität von Mittelwert und Varianz. In Abbildung 3.15 sind zwei Zeitreihen mit saisonalem Verlauf um ein konstantes Niveau (a.) und mit linearem Trend (b.) dargestellt. Beide sind sowohl als Zeitreihendiagramm y_t unter Berücksichtigung der zeitlichen Ordnung der Bedarfs-Zeitpunkte t als auch als Wahrscheinlichkeitsfunktion $P_Y(y)$ ohne Zeitbezug von y_t abgebildet.¹

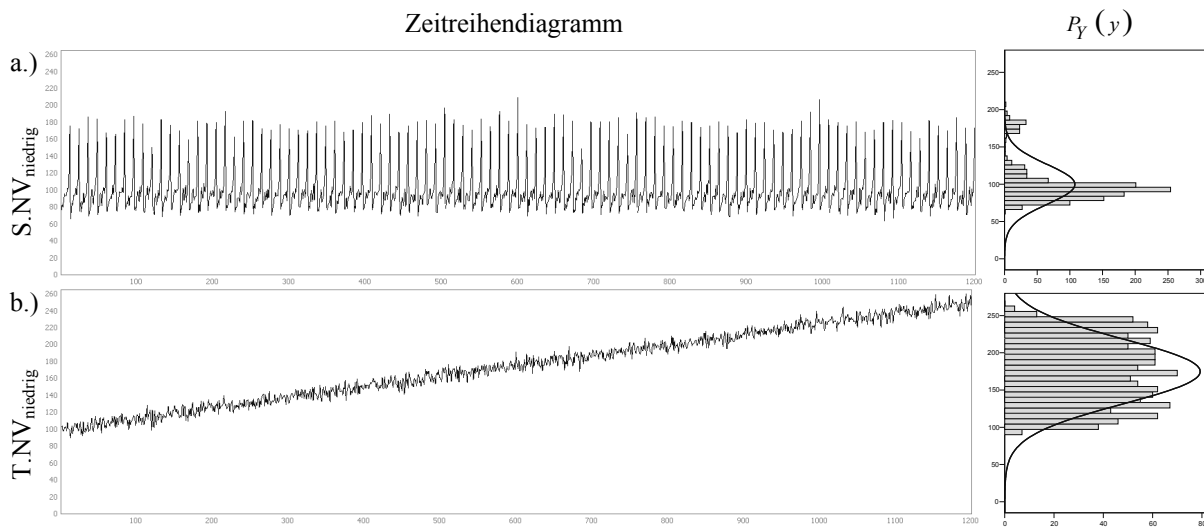


Abbildung 3.15: Instationäre Zeitreihen mit saisonalem Verlauf (a.) und trendbehaftetem Verlauf (b.)²

Zeitreihenmuster mit Trend-, Saison- oder Trend-Saison-Komponenten sind dementsprechend nicht mittelwertstationär.³ Abbildung 3.15 verdeutlicht den Effekt von Zeitreihen mit instationärem Mittelwert auf die Wahrscheinlichkeitsfunktion bei konstanter Streuung: Im Fall saisonaler Zeitreihen wird eine multimodale Verteilung mit erhöhter Varianz suggeriert, ohne die deterministische zeitliche Struktur zu berücksichtigen, im Fall trendbehafteter Zeitreihen wird eine abweichende, angenähert uniforme Verteilungsform mit erhöhter Varianz gebildet. Die Betrachtung als statische Wahrscheinlichkeitsfunktion erlaubt somit keine valide Repräsentation der dynamischen Veränderung des Bedarfs.

Da die regelmäßigen Komponenten der Bedarfszeitreihen im Zeitablauf im Sinne eines konstanten Saisonmusters bzw. eines konstanten Trendanstiegs unverändert bleiben, werden diese auch als deterministische Komponenten betrachtet. Sie unterscheiden sich von stochastischen Ausprägungen der Komponenten Niveau, Saison und Trend, die sich im Zeitablauf abrupt im Sinne eines einmaligen Strukturbruchs oder kontinuierlich strukturell verändern können. Diese bezeichnen die irregulären Komponenten der Zeitreihe, welche die regelmäßigen Komponenten überlagern. Strukturveränderungen treten in Form von Veränderungen des Niveaus, auch als Niveauverschiebungen⁴ (engl. *level shifts*) bezeichnet, Veränderungen der Saison-Muster, Veränderungen der Trend-Muster oder auch Veränderungen der irregulären Komponente als Heteroskedastizität auf. Strukturbrüche der Zeitreihe bezeichnen daher Veränderungen der regelmäßigen Komponenten durch Trendumkehr,

¹ Die Varianz der Zeitreihen ist normalverteilt und entspricht der synthetischen Zeitreihe NV_{niedrig} der Experimente in Abschnitt 5.4.

² Die Zeitreihe $S.NV_{\text{niedrig}}$ wird nachfolgend im Experiment in Kapitel 5 verwendet. Die Zeitreihe $LS.NV_{\text{niedrig}}$ zeigt eine Niveauverschiebung von $N(100, 5)$ auf $N(180, 5)$ mit konstanter Streuung nach 900 Beobachtungswerten. Die Zeitreihe $T.NV_{\text{niedrig}}$ zeigt einen Trend von 0,125 pro Periode von $N(100, 5)$.

³ Im Fall einer multiplikativen Verknüpfung der Komponenten mit der zufälligen Restkomponente sind trendbehaftete sowie saisonale Zeitreihen ebenfalls nicht varianzstationär und verstoßen so gegen die Annahme schwach stationärer Prozesse (siehe Abschnitt 3.2.2.1).

⁴ Diese treten als Veränderungen des lokalen Mittelwertes über mehrere Zeitperioden auf, z. B. verursacht durch Ein- und Auslistungen von Produkten oder den Eintritt eines Marktteilnehmers in einen bzw. Austritt aus einem verteilten Markt.

Niveau- oder Saisonverschiebungen sowie Veränderungen der irregulären Komponente und falsifizieren die Zeitstabilitätshypothese. Sie erfordern damit entweder eine erneute Schätzung des Modells, u. U. mit einer verkürzten Zeitreihe unter Ausschluss früherer Daten, oder die Berücksichtigung stochastischer Niveaus, Trends oder Saisonalitäten durch entsprechende Prognoseverfahren für stochastische Zeitreihenmuster wie z. B. der Exponentiellen Glättung.¹

Ausgewählte unregelmäßige Komponenten sind in Abbildung 3.16 als Zeitreihendiagramm und als Wahrscheinlichkeitsfunktion abgebildet.²

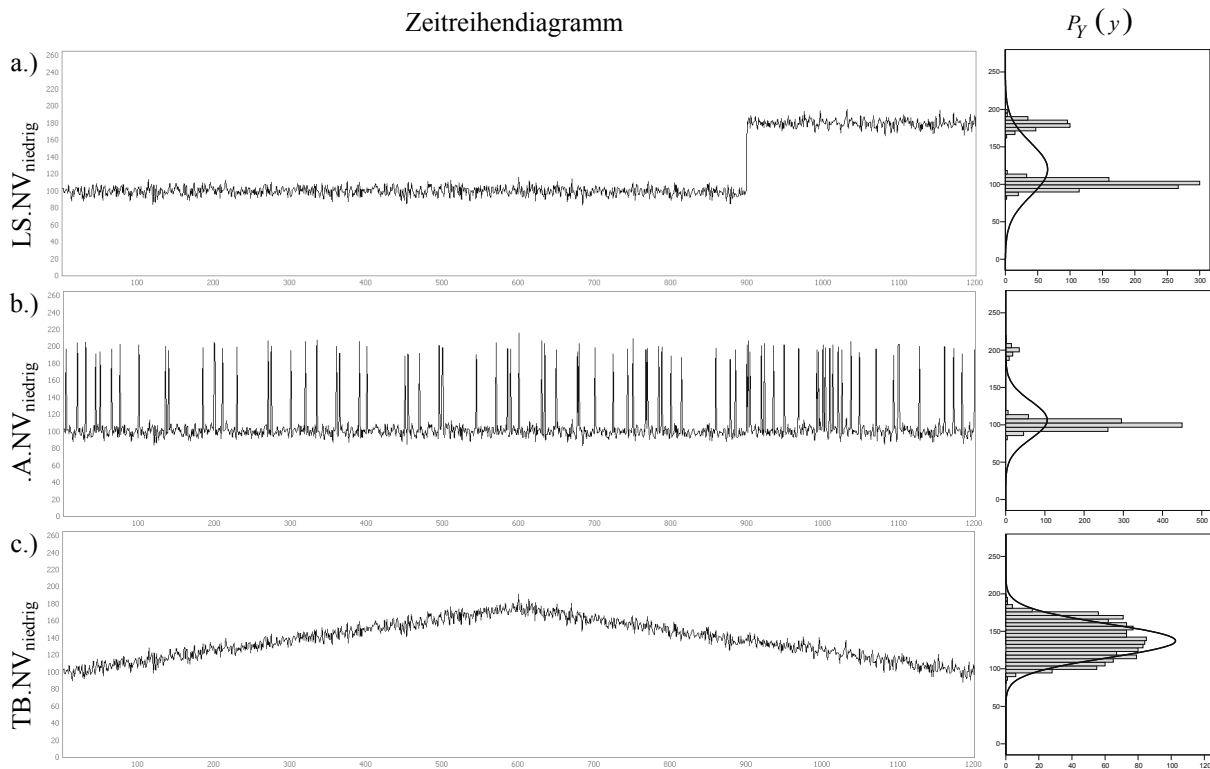


Abbildung 3.16: Instationäre Zeitreihen mit Niveauverschiebung (a.), Ausreißern (b.) und einem Strukturbruch im Trendverlauf (c.)

In Abbildung 3.16 (a.) ist eine instationäre Zeitreihe mit Niveauverschiebung abgebildet, in Abbildung 3.16 (c.) eine Zeitreihe mit Trendumkehr.³ In Abbildung 3.16 (b.) wird eine mittelwertstationäre Bedarfszeitreihe überlagert durch unregelmäßige Ausreißer abgebildet. Die instationären, multimodalen Wahrscheinlichkeitsverteilungen der Zeitreihen in Abbildung 3.16 a.) und b.) besitzen weder Saison noch Trend, sondern entstehen ausschließlich durch die Überlagerung eines mittelwertstationären Niveaus mit den unregelmäßigen Komponenten der Niveauverschiebung (a.) und Ausreißer (b.), die in empirischen Anwendungen häufig auftreten.⁴ Ausreißer (engl. *outlier*) bezeichnen einmalige Niveauveränderungen mit unterschiedlichen Formen und zeitlichen Verzögerungen.⁵ Ausreißer können auch durch

¹ Vereinzelt wird die irreguläre Restgröße E als Bestandteil der systematischen Komponente betrachtet, da ihre zentralen Parameter im Zeitablauf unverändert bleiben. Ausreißer oder Niveauverschiebungen werden dagegen als unsystematische Komponenten aufgefasst. Das Verhältnis der systematischen Komponenten L , T und S zu den irregulären, ungeklärten Komponenten bezeichnet den Signal-Rauschabstand (engl. *signal-to-noise-ratio*) durch den Variationskoeffizienten und erlaubt eine Abschätzung der Prognosequalität anhand der Regelmäßigkeiten in den Beobachtungswerten.

² Für eine analytische Definition und Abgrenzung deterministischer sowie stochastischer Saisonalität und Trends siehe insbesondere die Arbeiten von OSBORN et al. [Osborn et al. 1999; Ghysels und Osborn 2001].

³ Alle hier dargestellten Zeitreihen besitzen dieselbe niedrige, normalverteilte Streuung, einschließlich der Zeitreihen in Abbildung 3.15 und Abbildung 3.16.

⁴ Ferner sind mehrfache, aufeinander folgende Ausreißer über mehrere Perioden (engl. *outlier patches*) von Niveauverschiebungen zu unterscheiden.

⁵ Der Begriff „Ausreißer“, auch als *pulse* oder *spike* bezeichnet, beschreibt unsystematische Beobachtungswerte, die sich der Struktur der Zeitreihe außerhalb der systematischen Schwankungsbreite entziehen. Ihr

unregelmäßige, exogene Faktoren wie zum Beispiel Marketing-Aktivitäten verursacht werden, die keinem regelmäßigen, saisonalen Muster folgen und nicht im Erklärungsmodell abgebildet sind. Das Auftreten von Nullmengen, durch Zeitperioden ohne Absatz, führt zu spezifischen Verzerrungen der regulären Komponenten.¹ Ausreißer und Niveauverschiebungen sind anhand der Wahrscheinlichkeitsfunktion nicht von einem saisonalen Muster in Abbildung 3.15 (a) zu unterscheiden.

Der stationäre Verlauf L.E ist trotz seines vergleichsweise seltenen Auftretens in der Praxis hervorzuheben, da objektive Prognoseverfahren der Verteilungsschätzung die Stationarität voraussetzen. Ferner erfordern alle zeitreihen- und kausalanalytischen Prognoseverfahren der Regressionsanalyse die Stationarität der Residuen der prognostizierten Zeitreihe. Die Vernachlässigung regelmäßiger Komponenten wie Trend, Saison und Trend-Saison sowie unregelmäßiger Komponenten wie Ausreißer, Niveauverschiebung oder Nullmengen führt zu einem Ansteigen der Varianz der Bedarfswerte im Vergleich zur tatsächlichen Streuung und somit zu systematisch überhöhten Sicherheitsbeständen und Bestellmengen.² „Otherwise, if only a mean demand and a standard deviation is computed, postulating a base demand level, the forecasting error and the safety stock requirements are larger than necessary“ [Minner 2000, S. 15]. Die Betrachtung der Wahrscheinlichkeitsverteilungen der instationären Zeitreihen in Abbildung 3.15 und Abbildung 3.16 verdeutlicht somit, dass sowohl regelmäßige als auch unregelmäßige Komponenten von Zeitreihen eine valide Bedarfsschätzung anhand von empirischer bzw. theoretischer Verteilungsannahmen verbieten. Vielmehr ist eine Berücksichtigung der jeweiligen Zeitreihenkomponenten durch die Anwendung statistischer Prognoseverfahren erforderlich, die zur Vorhersage instationärer Prozesse geeignet sind.³

Prognoseverfahren der Zeitreihenanalyse modellieren einzelne oder mehrere regelmäßige Komponenten der Zeitreihe aus Beobachtungen der Vergangenheit durch Trennung von den unregelmäßigen Einflüssen zur Prognose in die Zukunft. Für jeden der 12 Verläufe wurden spezifische Verfahren der Zeitreihenprognose entwickelt, die jeweils nur die spezifische Modellform der Komponenten „valide“ und „reliabel“ prognostizieren können. So kann die Einfache Exponentielle Glättung nach BROWN lediglich zur Prognose stationärer Zeitreihen (L.E) eingesetzt werden, nicht aber zur Prognose saisonaler oder trendbehafteter Zeitreihen. Die lineare Exponentielle Glättung nach HOLT kann hingegen nur für Zeitreihen ohne Saisonalität mit linearem Trend (L.T.L.E) verwendet werden, da bei einer Anwendung auf stationäre Zeitreihen die unkorrelierten Residuen als Trendwerte fehlinterpretiert werden, im Fall anderer Trendformen die Prognosen dem tatsächlichen Trend hinterherhinken. Die Verfahren werden nachfolgend nach den archetypischen Zeitreihenmustern geordnet darge-

Gegenstück, sozusagen die „Einreißer“ (engl. „inlier“) [Reilly 2000], können hingegen nur schwer identifiziert werden, da ihre Realisationen nicht außerhalb der Wahrscheinlichkeitsverteilung des stochastischen Prozesses auftreten und somit nicht visuell hervortreten. Sie verdeutlichen jedoch die Notwendigkeit der Formulierung eines adäquaten Modells vor der Identifikation und Bereinigung von Ausreißern, im Gegensatz zur einfachen Bereinigung aller Beobachtungswerte in den Enden einer Wahrscheinlichkeitsverteilung die z. B. größer als drei Standardabweichungen ausfallen [Reilly 2000].

¹ Bei seltenem Auftreten von Nullmengen können diese durch kausalanalytische Modellierung in Interventionsmodellen berücksichtigt werden. Die Prognose von Zeitreihen mit einem regelmäßigem Auftreten von Nullmengen, was als sporadischer Bedarf bezeichnet wird und z. B. für den Bedarf an Ersatzteilen auftritt, erfordert die Berücksichtigung in spezifischen Prognoseverfahren. Diese werden jedoch nachfolgend aufgrund ihrer verringerten Bedeutung vernachlässigt.

² Da Niveauverschiebungen und Ausreißer des Bedarfs die Annahmen der Stationarität verletzen und zu Verzerrungen der Wahrscheinlichkeitsverteilung führen, sind diese im Fall der Bedarfsschätzung durch Verteilungsannahme für alle vergangenen Beobachtungswerte zu korrigieren, um eine valide Bedarfsschätzung durchführen zu können.

³ Alternativ kann die Bedarfszeitreihe durch geeignete Transformationen in einen stationären Prozess überführt werden, im Sinne einer Dekomposition und Extrapolation der stationären Zeitreihe und anschließender Rekombination zur Vorhersage zukünftiger Realisationen. Dies wird jedoch nachfolgend vernachlässigt.

stellt, um die Notwendigkeit der Übereinstimmung von Modellform und Verfahren hervorzuheben.

Unter der Annahme sich mehrfach überlagernder regelmäßiger Komponenten kann eine Zeitreihe schrittweise in ihre einzelnen Komponenten zerlegt und so genauer analysiert werden, oder die einzelnen Komponenten können getrennt prognostiziert werden. In Abbildung 3.17 wird die Zerlegung der Originalzeitreihe in eine Trend-Zyklus-Komponente, eine saisonale Komponente und den stochastischen Rest demonstriert, der von stationärem Charakter sein sollte.

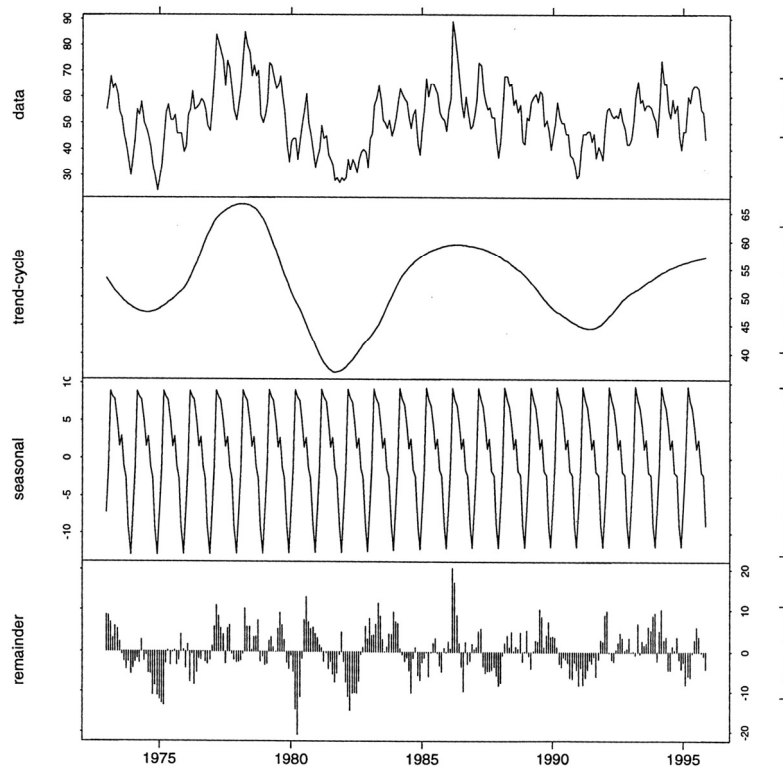


Abbildung 3.17: Dekomposition einer Zeitreihe in ihre Komponenten¹

Zur Zerlegung von Zeitreihen mit monatlicher Frequenz in ihre Einzelkomponenten können unterschiedliche Verfahren der Dekomposition verwendet werden. Die klassischen Verfahren der Dekomposition umfassen die Trendbereinigung durch Bildung erster Differenzen sowie die Saisonbereinigung durch die Bildung von Saisonindizes [Makridakis et al. 1998, S. 106-109; Hüttner 1986, S. 117] oder die Abbildung der Saisonkomponenten durch Sinus- bzw. Kosinus-Funktionen durch das Berliner Verfahren [Hansmann 1983] oder das Verfahren nach HARRISON [1967]. Zudem werden verstärkt X-11- oder X-12-Verfahren der Census-II-Dekomposition unter Verwendung von ARIMA-Verfahren eingesetzt [Shishkin und Young 1967; Dagum 1988; Findley et al. 1998], STL-Methoden [Makridakis et al. 1998, S. 121-125] sowie vereinzelt durch nichtlineare Ansätze [Ozaki und Thomson 2002] oder die Frequenz- und Spektralanalyse [Lo 1979, S. 106].

Bei der Zerlegung einer Zeitreihe können die Einzelkomponenten jedoch nur selten auf der Basis von a priori Wissen bestimmt werden. Ansätze der Dekomposition dienen daher eher der Erklärung denn der Prognose zukünftiger Werte [Makridakis et al. 1998]. Entsprechend verfolgen Verfahren der Zeitreihenprognose nicht das Ziel der Dekomposition einer Zeitreihe

¹ Darstellung des monatlichen Verkaufs von Einfamilienhäusern in den USA zwischen 1973 und 1998, entnommen aus [Makridakis et al. 1998, S. 86].

in einzelne Komponenten, sondern vielmehr die simultane Extrapolation der interdependenten Wirkung aller Komponenten in die Zukunft.

3.2.3.2 Verfahren bei konstantem Bedarfsniveau

3.2.3.2.1 Naive Verfahren und Mittelwerte

Verfahren zur Prognose von Zeitreihen mit konstantem Bedarfsniveau prognostizieren stochastische, mittelwertstationäre Zeitreihen.¹ Sie implizieren ein Erklärungsmodell mit konstantem Niveau der Zeitreihe, das lediglich durch eine irreguläre Komponente überlagert wird, also

$$y_t = f(L_t, E_t) \quad , \quad (3.36)$$

wobei eine additive oder multiplikative Verknüpfung der irregulären Komponenten mit dem Basisniveau möglich ist. Unter der Voraussetzung eines stationären Verlaufs ohne Trend, ohne Saison und mit konstanter Varianz der irregulären Komponente sowie Gültigkeit der Zeitstabilitätshypothese können Naive Verfahren sowie Verfahren der Mittelwertbildung zur Prognose angewendet werden.

Das einfache Naive Verfahren als NF1 unterstellt einen instationären Zufallsprozess des random walk mit einer stochastischen Zufallsvariablen ε_t der Form [Makridakis et al. 1980]

$$y_t - y_{t-1} = \varepsilon_t \quad . \quad (3.37)$$

Die Zufallsvariable ε_t ist stationär und folgt einem diskreten, stochastischen Prozess unkorrelierter Zufallsvariablen mit Erwartungswert null, konstanter Varianz und konstanter Wahrscheinlichkeitsverteilung, der auch als weißes Rauschen (engl. *white noise*) bezeichnet wird.

Eine Betrachtung des Prozesses über die Zeit für $t = 1, \dots, T$ erklärt die geringe Struktur des Erklärungsmodells durch die Fehlerterme [Schlittgen und Streitberg 1987, S. 72]:

$$y_T = \varepsilon_T + \varepsilon_{T-1} + \varepsilon_{T-2} + \dots + \varepsilon_1 \quad . \quad (3.38)$$

Da keine weiteren Komponenten zu verwerten sind, leitet sich durch einfache Umformung von (3.37) ein einfaches Prognosemodell ab:

$$y_t = y_{t-1} + \varepsilon_t \quad (3.39)$$

„The ... Naïve Forecast 1 or NF1 ... method uses the most recent observation available as a forecast“ [Makridakis et al. 1998, S. 46-47], mit

$$\hat{y}_{t+1} = y_t \quad . \quad (3.40)$$

Das Verfahren des NF1 führt jedoch selbst bei einem stationärem Zeitreihenverlauf der Form L.E aus Abbildung 3.14 nicht zu einer validen Trennung von Niveau und Fehlerterm, sondern extrapoliert sowohl das Niveau als auch das irreguläre Rauschen in die Zukunft. Es kann somit gezeigt werden, dass Naive Verfahren die wahre Struktur eines stationären Prozesses in Form des Mittelwertes der Zeitreihe nicht valide von den zufälligen Schwankungen trennen

¹ Sie werden häufig irreführend als konstante oder horizontale Zeitreihenverläufe bezeichnet.,

können. Sie dienen vielfach als Benchmark einer minimalen Prognosequalität und sind nicht als adäquates Prognoseverfahren - selbst für ein stationäres Zeitreihenmuster - zu verstehen.¹

3.2.3.2.2 Mittelwerte

Um die systematische Struktur des Niveaus einer stationären Zeitreihe hervortreten zu lassen, werden unterschiedliche Verfahren der Mittelwertbildung über mehrere Beobachtungsperioden eingesetzt. Die strukturelle Information über das Niveau der Zeitreihe wird von der irregulären Komponente durch Glätten der Zufallskomponente getrennt [Hansmann 1983, S. 27].² Dabei können unterschiedliche Formen der Mittelwertbildung aus arithmetischen, gleitenden und gewichteten Mittelwerten zum Einsatz kommen.³

Das arithmetische Mittel (AM) einer Zeitreihe wird zur Ausschaltung zufallsbedingter Unregelmäßigkeiten im Verlauf der Zeitreihe über alle t Beobachtungswerte der Zeitreihe $i = 1, 2, \dots, t$ gebildet und als Prognosewert zukünftiger Realisationen verwendet:

$$\hat{y}_{t+1} = \bar{y}_t = \frac{1}{t} \sum_{i=1}^t y_i \quad . \quad (3.41)$$

Unter Annahme der Stationarität können auf diese Weise irreguläre Schwankungen E und Ausreißer geglättet und das Niveau L der Zeitreihe zur Prognose gefunden werden. Treten keine Niveauveränderungen oder Ausreißer im Beobachtungszeitraum auf, erlaubt das arithmetische Mittel eine maximale Glättung und bestmögliche Schätzung. Da alle Beobachtungswerte seit Beginn der Zeitreihe berücksichtigt werden, verzerren jedoch Ausreißer dauerhaft die Schätzung von L mit $1/t$, und es kann keine Anpassung an Instationaritäten durch Niveauverschiebungen erfolgen. Daher werden gleitende Mittelwerte (engl. *moving average*, MA) mit einer definierten Fenstergröße von l Perioden zur Berücksichtigung irregulärer, stochastischer Zeitreihenkomponenten verwendet:

$$\hat{y}_{t+1} = \bar{y}_{t,l} = \frac{y_t + y_{t-1} + \dots + y_{t-(l-1)}}{l} = \frac{1}{l} \sum_{i=0}^{l-1} y_{t-i} \quad . \quad (3.42)$$

Die Intensität der Glättung wird durch die Länge des Zeitfensters l determiniert, welche die Balance zwischen Robustheit gegen Ausreißer und Adaptivität an Niveauveränderungen bestimmt. Durch Bildung eines gleitenden Mittels der Länge $l = 1$ entspricht das Verfahren dem Naiven Verfahren NF1.

In AM und MA erhält jeder Beobachtungswert unabhängig von seiner Aktualität dasselbe Gewicht $1/t$. Aufgrund der nur eingeschränkten empirischen Gültigkeit der Zeitstabilitätshypothese bezüglich möglicher Änderungen einer stationären Zeitreihenstruktur erscheint es vielfach sinnvoll, dass „jüngere Zeitreihenwerte ein stärkeres Gewicht erhalten als die älteren, weil sie die gegenwärtige Zeitreihenstruktur besser widerspiegeln“ [Hansmann 1983, S. 28]. Entsprechend werden vielfach gewichtete, gleitende Mittelwerte verwendet, mit

$$\hat{y}_{t+1} = \sum_{i=0}^{l-1} a_i y_{t-i} \quad , \quad \text{mit} \quad \sum_{i=0}^{l-1} a_i = 1 \quad . \quad (3.43)$$

¹ Trotz der geringen Verbreitung stationärer Zeitreihen im Handel dienen Naive Verfahren aufgrund ihrer Einfachheit und ihrer effizienten Berechnung in der Praxis vielfach als Grundlage subjektiver Vorhersagen von Experten. Auch die Anwendung der Naiven Verfahren des Trends NF2 und der Saison NF3 sind weit verbreitet, trotz signifikanter Unterlegenheit gegenüber selbst einfachen Verfahren der Mittelwertbildung.

² In der ingenieurwissenschaftlichen Terminologie wird dies auch als Filtern des Rauschens bezeichnet.

³ Sie bilden auch die Grundlage für die einfache Dekomposition von Zeitreihen [Makridakis et al. 1998, S. 89].

Dabei wird jede Beobachtung der Vergangenheit im Zeitfenster l durch a_l gewichtet, unter asymmetrischer Verteilung der Gewichte: $a_1 = a_2 = \dots = a_l$.¹ Erweitert man den Grundgedanken der stärkeren Gewichtung jüngerer Zeitreihenwerte auf eine exponentiell abnehmende Gewichtung der Beobachtungswerte erfolgt der Übergang zu den Verfahren der Exponentiellen Glättung.

3.2.3.2.3 Einfache Exponentielle Glättung

Verfahren der Exponentiellen Glättung (ES) repräsentieren eine Klasse gleitender Mittelwerte mit exponentieller Gewichtung der Zeitreihenwerte, die besondere Verbreitung und Bedeutung in der betriebswirtschaftlichen Praxis erlangt haben. Die auf BROWN, HOLT und MAGEE zurückgehende Verfahrensklasse der Exponentiellen Glättung (engl. *exponential smoothing*, ES) wurden als Verfahren zur wirtschaftlichen elektronischen Datenverarbeitung im Rahmen des Bestandsmanagements entwickelt [Brown 1959, S. 13; Magee 1958; Brown 1956; Brown 1963]. Alle Verfahren „apply an unequal set of weights to past data, and ... the weights typically decay in an exponential manner from the most recent to the most distant data point“ [Makridakis et al. 1998, S. 136]. Die Verfahren der ES repräsentieren eine Verfahrensklasse zur Prognose aller Modelle der 12 Zeitreihenmuster, jedoch ohne Methodologie zur adäquaten Modellauswahl. Die korrekte Verfahrensform ist somit in Anlehnung an die postulierte Modellform manuell zu bestimmen.

Für stationäre Zeitreihen modelliert das Verfahren der einfachen Exponentiellen Glättung (engl. *single exponential smoothing*, S.ES) nach BROWN [1959] einen gleitenden, exponentiell gewichteten Mittelwert.² Das Verfahren erlaubt zwei Interpretationen seiner Funktionsweise. Wird in Periode t fehlerhaft prognostiziert, so entsteht ein Prognosefehler e_t in der Periode t als Differenz des Zeitreihenwertes y_t und des in der Vorperiode $t-1$ prognostizierten Wertes \hat{y}_t ,

$$e_t = y_t - \hat{y}_t \quad . \quad (3.44)$$

Zur Ermittlung des aktuellen Niveaus L_t wird der in $t+1$ berechnete Prognosewert \hat{y}_t um einen Bruchteil α des Prognosefehlers e_t angepasst, sodass der nächste Prognosewert näher an dem wahren Mittelwert liegt. So gilt für die Berechnung des neuen Niveaus in Periode t

$$\begin{aligned} L_t &= y_t + \alpha \cdot e_t & (3.45) \\ &= y_t + \alpha \cdot (y_t - \hat{y}_t) \end{aligned}$$

welches als Prognose für die Periode $t+1$ verwendet wird:

$$\hat{y}_{t+1} = L_t \quad . \quad (3.46)$$

Die Modellform verdeutlicht die Möglichkeit der S.ES zur Berücksichtigung stochastischer Änderungen des Mittelwertes im Zeitablauf durch schrittweise Anpassung an das neue Niveau, was einen Vorteil gegenüber der Bedarfsprognose durch Verteilungsannahmen oder das arithmetische Mittel darstellt. Der Parameter α , mit $0 \leq \alpha \leq 1$, wird entsprechend als Reaktionsparameter [Hüttner 1986, S. 56] oder Glättungsparameter bezeichnet [Hansmann 1983, S. 19].

¹ Eine Diskussion weiterer Ansätze gewichteter, zentrierter, doppelter und mehrfacher gleitender Mittelwerte sowie Ansätze zur Glättung mit lokaler Regression findet sich in [Makridakis et al. 1998, S. 89-106].

² S.ES wird daher auch als „exponentially weighted moving average“ bezeichnet. In Anlehnung an MAKRIDAKIS et al. [1998] wird eine gegenüber [Hansmann 1983; Scheer 1983] veränderte Notation verwendet.

Die zweite Interpretation betrachtet das Verfahren als „weighted sum of all past demand, with the heaviest weight being placed on the most recent information“ [Brown 1959, S. 13]. Das Gewicht vergangener Zeitreihenwerte im Mittelwert nimmt mit zunehmenden Alter der Daten exponentiell ab [Hansmann 1983, S. 28]. Dies wird deutlicher bei Umformung der Gleichung (3.43) zu

$$L_t = \alpha \cdot y_t + (1 - \alpha) \cdot L_{t-1} \quad . \quad (3.47)$$

Die Prognose der Folgeperiode \hat{y}_{t+1} berücksichtigt den aktuellen Beobachtungswert \hat{y}_t mit einem Bruchteil von α und alle vorherigen Beobachtungswerte mit $(1 - \alpha)$.¹ Da auch gilt

$$\begin{aligned} L_{t-1} &= \alpha \cdot y_{t-1} + (1 - \alpha) \cdot L_{t-2} \\ L_{t-2} &= \alpha \cdot y_{t-2} + (1 - \alpha) \cdot L_{t-3} \quad \dots \end{aligned} \quad (3.48)$$

zeigt sich durch rekursives Einsetzen in die Grundgleichung der Prognose (3.47) die exponentiell abnehmende Gewichtung $\alpha(1 - \alpha)^l$ aller Zeitreihenwerte y_{t-l} mit zunehmenden Alter l [Hansmann 1983, S. 29; Scheer 1983, S. 94; Lewandowski 1974, S. 33] gemäß

$$\hat{y}_{t+1} = L_t = \alpha y_t + \alpha(1 - \alpha)y_{t-1} + \alpha(1 - \alpha)^2 y_{t-2} + (1 - \alpha)^3 \hat{y}_{t-2} + \dots \quad , \quad (3.49)$$

bis hin zu

$$\hat{y}_{t+1} = \alpha \sum_{l=0}^{t-1} (1 - \alpha)^l y_{t-l} + (1 - \alpha)^t \hat{y}_1 \quad . \quad (3.50)$$

In dieser exponentiell abnehmenden Gewichtung der Zeitreihenwerte und der Eigenschaft der Glättung der Zufallsschwankungen ist die Bezeichnung „Exponentielle Glättung“ nach BROWN begründet [Brown 1959, S. 13; Stier 2001]. Eine analytische Untersuchung der Herleitung findet sich in [Makridakis et al. 1998, S. 147-154], ein Nachweis der validen Minimierung der exponentiell gewogenen quadratischen Fehler in Anlehnung an statistische Schätzverfahren durch den Prognosewert bei [Hansmann 1983, S. 30]. Die S.ES wird durch einen einzigen Parameter α gesteuert, weshalb sie auch als exponentielle Glättung erster Ordnung bezeichnet wird. Analog zu den gleitenden Mittelwerten determiniert der Glättungsparameter α über die Stärke der exponentiellen Abnahme der Gewichtung die Länge des Zeitfensters und somit die Intensität der Glättung. Geringe Werte von α gewichten die Vergangenheit höher als den letzten Zeitreihenwert; für $\alpha = 0$ wird der erste Beobachtungswert für alle Perioden extrapoliert, für $\alpha = 1$ wird jeweils der vorige Zeitreihenwert im Sinne des NF1 extrapoliert und somit kein gleitender Mittelwert gebildet.

Abbildung 3.18 verdeutlicht den exponentiellen Verlauf der Gewichte für unterschiedliche Ausprägungen von α , mit $\alpha = \{0,05; 0,2; 0,7; 0,9\}$ sowie eine gleichmäßige Gewichtung mit 20% durch einen gleitenden, fünfperiodigen Mittelwert MA(5).

¹ Die Formel verdeutlicht ferner die geringen methodischen Anforderungen von nur drei Parametern α , \hat{y}_t und y_t zur Informationsverarbeitung, die aufgrund der hohen Speicherplatzkosten der 50er und 60er Jahre zu Durchsetzung der Verfahrensklasse der ES in der Praxis beigetragen haben.

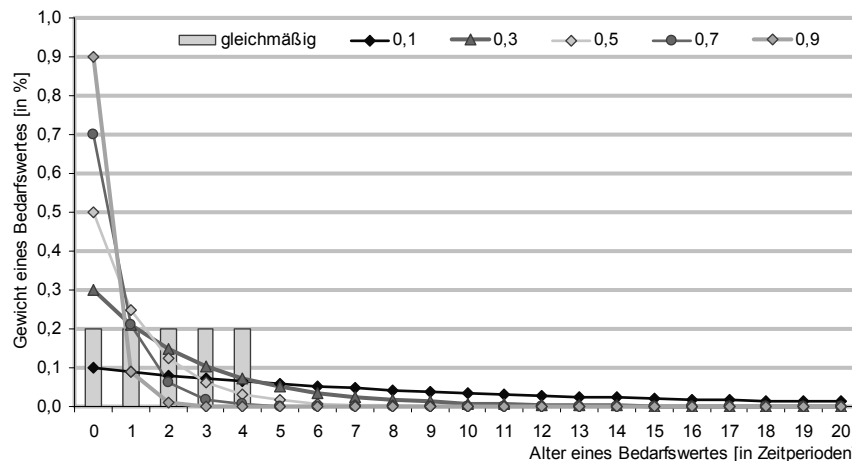


Abbildung 3.18: Gewichtung der Beobachtungswerte in der Exponentiellen Glättung

Der Parameter α balanciert somit die Intensität der Glättung von Zufallsschwankungen und Ausreißern einer stationären Zeitreihe gegen die Reagibilität auf instationäre Komponenten der Niveauverschiebungen.¹ Optimale Parameterkonstellationen sind zeitreihenspezifisch in Abhängigkeit der Schwankungsintensität und der Häufigkeit von Niveauverschiebungen und Ausreißern festzulegen und i. d. R. nicht ex ante zu bestimmen. Empfehlungen für sinnvolle Parameterwerte liegen im Bereich $0,01 \leq \alpha \leq 0,5$ [Hansmann 2001, S. 333; Hüttner 1986, S. 56; Mertens und Deistler 1994].²

Der Prozess der Bestimmung optimaler Parameter sowie das Problem der Definition der Startparameter zur Initialisierung für ein ausgewähltes Modell werden hier vernachlässigt [Götz 1999, S. 106-107]. Ansätze zur adaptiven Variation des α -Parameters im Sinne des Adaptive Response Rate Exponential Smoothing nach CHOW [Chow 1965], ebenso in SMITH [Hansmann 1983, S. 39-41; Smith 1974] oder nach TRIGG und LEACH [Scheer 1983, S. 134; Trigg und Leach 1967] erlauben theoretisch die unbeaufsichtigte Anpassung des α -Parameters an unterschiedliche Streuungsgrade, auftretende Niveauverschiebungen und Ausreißer. Dies soll eine Anwendung der S.ES auf stationäre sowie instationäre Zeitreihen ohne Benutzereingriff durch automatische Adaption des α an veränderte Niveaus und Phasen erhöhter Varianz erlauben [Makridakis et al. 1998, S. 434]. Insbesondere berücksichtigen sie keine Strukturbrüche durch Auftreten neuer Zeitreihenkomponenten und Modellfehlspezifikationen. Die genannten Ansätze werden aufgrund ihrer suboptimalen Ergebnisse in empirischen Vergleichsstudien [Ord et al. 2000a; Gardner und Dannenbring 1980] und ihrer geringen empirischen Verbreitung jedoch nachpfögend vernachlässigt.

3.2.3.2.4 ARMA-Modelle

ARMA-Modelle kombinieren die Verwendung von autoregressiven Funktionstermen (AR) und von Moving-Average-Termen (MA) in einem zeitreihenanalytischen Erklärungsmodell. Sie erklären jede Realisation eines stationären stochastischen Prozesses als lineare Kombination vorheriger Realisationen und der Fehler in der Schätzung dieser Realisationen. Seit der populären Veröffentlichung einer Methodologie und Notation zur Modellierung von ARMA-Modellen durch BOX und JENKINS [Box und Jenkins 1970; Box et al. 1994] werden sie synonym als BOX-JENKINS-Modelle bezeichnet. Sie sind in Theorie und Praxis fest etabliert

¹ Eine Abbildung der Reaktion auf Ausreißer bzw. Pulses als einmalige Störungen und Niveauverschiebungen bzw. Level-Shifts findet sich u. a. in [Scheer 1983].

² Geringere Fehler bei der Ex-post-Ermittlung von höheren Parametern als 0,5 im Stützbereich der Zeitreihe deuten vielfach auf ein fehlerhaft selektiertes Modell hin [Mertens und Deistler 1994, S. 25]. Dies setzt jedoch das Vorhandensein einer Mindestanzahl von Beobachtungswerten proportional zum Glättungsparameter voraus, da andernfalls Beobachtungen zum Beginn der Zeitreihe ein stärkeres Gewicht erhalten.

und besitzen aufgrund der unterstellten Form des Erklärungsmodells eine besondere Relevanz für die Modellierung von Neuronalen Netzen zur Prognose. Zudem enthalten sie als Oberklasse u. a. die Verfahren der Exponentiellen Glättung und der Naiven Prognosen als Sonderfälle der Spezifikation [Hansmann 1983, S. 70]. Die einzelnen Funktionsterme der autoregressiven Prozesse, der Moving-Average-Prozesse und der Integration werden nachfolgend erläutert.

Die Anwendung autoregressiver Komponenten in der Prognose beruht auf der Interpretation einer Zeitreihe als Realisation eines stationären, stochastischen Prozesses. Ein autoregressives (AR) Modell der Ordnung p entspricht einer Regression unter Verwendung p zeitverzögerter Realisationen der abhängigen Variable $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ im Sinne der Zeitreihenanalyse.¹ Die Regressionsgleichung des AR bezeichnet den konstanten Term c , den autoregressiven Parameter ϕ_l der um l Perioden zeitverzögerten Realisation der abhängigen Variable y_{t-l} im Zeitpunkt $t-l$ und e_t den zufälligen Fehlerterm im Zeitpunkt t und lautet

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + e_t \quad (3.51)$$

$$\text{mit } -1 < \phi_2 < 1; \phi_2 + \phi_1 < 1; \phi_2 - \phi_1 < 1 \quad \forall p = 2$$

Sind die ersten p Gewichte $\phi_1, \phi_2, \dots, \phi_p$, $p = l$, ungleich null und alle übrigen gleich null, sprechen BOX und JENKINS von einem autoregressiven Prozess der Ordnung p [Box und Jenkins 1970, S. 51]. Man spezifiziert die Form des Erklärungsmodells als AR(p)-Prozess, wobei p die höchste im Modell verwendete Lag-Struktur l bezeichnet. Im Zuge einer einheitlichen Notation aller ARMA- und nachfolgend ARIMA-Modellkomponenten wird auf die alternative Notation des Lag- oder Backshift-Operators B $y_t = y_{t-1}$ zurückgegriffen [Makridakis et al. 1998, S. 334; Harvey 1993]. Somit kann Gleichung (3.51) umgeformt werden zu

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) y_t = c + e_t \quad (3.52)$$

Analog zur Autoregression der abhängigen Variable y_t auf vorherige Realisationen der abhängigen Variablen y_{t-p} im AR-Modell können zukünftige Realisationen durch systematische Prognosefehler e_t des Modells in vorangegangenen Perioden erklärt werden. Dies geschieht durch eine iterativen Autoregression der abhängigen Variablen auf q vorangegangene Fehler e_t , mit $e_t = y_t - \hat{y}_t$ (3.44), im Sinne einer Korrektur der Vorhersage durch systematische Fehler in der Vergangenheit, was als Moving-Average- bzw. MA-Prozess bezeichnet wird.² Der Parameter c bezeichnet den konstanten Term und θ_l den Moving-Average-Parameter des Fehlerterms e_{t-l} in Periode $t-l$, mit

$$\begin{aligned} y_t &= c + y_t - \hat{y}_t - \theta_1 (y_{t-1} - \hat{y}_{t-1}) + \theta_2 (y_{t-2} - \hat{y}_{t-2}) + \dots + \theta_q (y_{t-q} - \hat{y}_{t-q}) \\ &= c + e_t - \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q} \end{aligned} \quad (3.53)$$

$$\text{mit } -1 < \theta_2 < 1; \theta_2 + \theta_1 < 1; \theta_2 - \theta_1 < 1 \quad \forall q = 2$$

¹ Für eine Diskussion der Annahmen und Voraussetzungen der Autoregression im Gegensatz zur Regressionsanalyse, insbesondere die Unabhängigkeit der Beobachtungen und Residuen, sei auf SCHLITGEN und STREITBERG [1987] verwiesen.

² Der Begriff des „Moving Average“-Terms ist trotz der englischen Bezeichnung nicht mit einem gleitenden Mittelwertes zu verwechseln.

Sind die ersten q Gewichte $\theta_1, \theta_2, \dots, \theta_p$ für $q = l$ ungleich null und alle übrigen gleich null, sprechen BOX und JENKINS von einem Moving-Average-Prozess der Ordnung q .¹ Analog zur Spezifikation des AR(p)-Prozesses bezeichnet q die höchste im Modell verwendete Lag-Struktur der Moving-Average-Terme zum MA(q)-Prozess. Auch der MA-Term kann in der Notation des Backshift-Operators formuliert werden:

$$y_t = c + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) e_t \quad (3.54)$$

Durch Kombination der AR- und MA-Modelle entsteht eine generelle Klasse zeitreihenanalytischer Modelle, genannt autoregressive Moving-Average- (ARMA) Modelle, die komplexe Zeitreihenstrukturen abbilden können [Makridakis et al. 1998, S. 336,345], mit

$$\begin{aligned} y_t &= c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} - \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q} \\ &= c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} - \theta_1 (y_{t-1} - \hat{y}_{t-1}) + \theta_2 (y_{t-2} - \hat{y}_{t-2}) + \dots + \theta_q (y_{t-q} - \hat{y}_{t-q}), \end{aligned} \quad (3.55)$$

oder

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) y_t = c + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) e_t \quad (3.56)$$

Die Ordnung des ARMA-Modells wird durch die Parameter p und q in der Notation ARMA(p, q) spezifiziert. Gemäß der Struktur des empirischen Prozesses können auch ausschließlich autoregressive Terme ARMA($p, 0$), verkürzt auch als AR(p)-Modell bezeichnet, oder ausschließlich Moving-Average-Terme ARMA($0, q$) bzw. MA(q)-Modelle zur Anwendung kommen.

BOX und JENKINS [Box und Jenkins 1970] definieren eine spezifische Methodologie zur validen und reliablen Datenanalyse und Modellselektion auf der Basis der Autokorrelationskoeffizienten, die drei iterative Phasen durchläuft: (1) Modell Identifikation, (2) Modell Schätzung und Test sowie (3) Modell Anwendung. Die Bildung von ARMA-Modellen durch Selektion der Parameter p und q beruht auf der Analyse und Identifikation statistisch signifikanter, absoluter und partieller Autokorrelationskoeffizienten als Maße der linearen Beziehung zwischen zeitlich verzögerten Beobachtungswerten einer Zeitreihe mit unterschiedlichen Lags l , $l = 1, \dots, T$ [Hansmann 1983, S. 65]. Die Autokorrelationsfunktion (AKF) als Folge der Koeffizienten jedes Lags beschreibt somit die Existenz und den Grad der autoregressiven Prozesse einer Zeitreihe [Hansmann 1983, S. 68].² Da diese nicht immer eindeutig sind, werden mehrere Kandidaten möglicher Prozesse identifiziert. Im Anschluss an die Modellidentifikation erfolgt die Schätzung der Parameter potenzieller Modelle durch nichtlineare kleinste Quadrate oder Maximum-Likelihood-Schätzer und die Beurteilung der statistischen Signifikanz ihrer Parameter und des Modells sowie die Analyse der Residuen [Makridakis et al. 1998, S. 359; Hansmann 1983, S. 74-95]. Abschließend erfolgt die Auswahl eines geeigneten Modells mit einer möglichst geringen Anzahl an Parametern und seine Anwendung zur Prognose [Makridakis et al. 1998].

Die Anwendung von ARMA-Modellen erfolgt unter der Annahme der Zeitreihe als Realisation eines stationären stochastischen Prozesses mit einer unkorrelierten, normalverteilten white-

¹ Der Begriff „moving average“ ist nicht mit dem gleichlautenden englischen Begriff für „gleitender Mittelwert der Beobachtungswerte“ zu verwechseln.

² Für eine statistische Bestimmung signifikanter Autokorrelations-Koeffizienten durch Portmanteau-Tests des Box-Pierce- oder Ljung-Box-Tests wird auf MAKRIDAKIS et al. [1998, S. 318-320] und Abschnitt 3.4.2.4 verwiesen.

noise-Störvariablen [Götz 1999, S. 118; Makridakis et al. 1980, S. 317]. Entsprechend erfordert ihre valide Anwendung die Stationarität von Mittelwert und Varianz der Zeitreihe [Box et al. 1994, S. 23; Schlittgen und Streitberg 1987], welche durch sogenannte „Unit-root“-Tests festgestellt werden kann [Makridakis et al. 1998]. Somit sind ARMA-Modelle ausschließlich auf stationäre, also nicht trendbehaftete und nichtsaisonale Zeitreihen anwendbar. Existiert ein mittelwert-stationärer Bedarfsverlauf mit instationärer Varianz, so muss die Zeitreihe durch entsprechende Transformation des Logarithmus oder eine BOX-COX-Transformation vorverarbeitet werden, um ein ARMA-Modell valide anwenden zu können. Die Erweiterung der ARMA-Modelle zur Abbildung Mittelwert-instationärer Bedarfsverläufe durch Integration der Zeitreihe in ARIMA-Modellen wird nachfolgend skizziert.

3.2.3.3 Verfahren bei trendbehaftetem Bedarfsverlauf

3.2.3.3.1 Naives Trendverfahren

Trendbehaftete Verläufe von Zeitreihen erfordern zur validen Prognose die explizite Berücksichtigung einer der Trendkomponenten in den Erklärungsmodellen, durch

$$y_t = f(L_t, T_t, E_t) \quad . \quad (3.57)$$

Dabei wird eine additive oder multiplikative Verknüpfung der Komponenten bei konstantem oder wechselndem Trend sowie bei konstantem oder wechselndem Niveau der Beobachtungswerte unterstellt. Der Prognosewert von Verfahren zur Prognose stationärer Zeitreihen, die irrtümlich zur Prognose von Zeitreihen mit Trend angewendet werden, „hinkt infolge seiner Konstruktion als gewogener Durchschnitt aller vergangenen Zeitreihenwerte dem aktuellen Trend hinterher“ [Hansmann 1983]. Entsprechend wurden spezifische Verfahren zur Berücksichtigung der regelmäßigen Komponente entwickelt.

Das Naive Verfahren NF1 kann auf trendbehaftete Zeitreihen zum NF2¹ erweitert werden, indem der letzte Trendwert als Veränderung im Niveau, also $y_t - y_{t-1}$ in die Zukunft prognostiziert wird:

$$\hat{y}_{t+1} = y_t + (y_t - y_{t-1}) = 2y_t - y_{t-1} \quad . \quad (3.58)$$

Für eine Mehrschrittprognose von einem Horizont h wird jeweils dasselbe Niveau und derselbe Trend linear in die Zukunft extrapoliert, indem $y_{t+1} = \hat{y}_{t+1}$ gesetzt wird bzw. die Trendkomponente h -fach addiert wird.

3.2.3.3.2 Doppelte Exponentielle Glättung

BROWN [1959] erweiterte das Prinzip der einfachen exponentiellen Glättung von der Prognose stationärer Zeitreihen um eine zusätzliche Glättung der Trendentwicklung zur Doppelten Exponentiellen Glättung (engl. *Double Exponential Smoothing*, D.ES).² Wird ein linearer Prozess aus einem Grundwert L_t zur Abbildung des Niveaus der Zeitreihe und einem Trendwert T_t als Veränderung des Grundwertes von Periode t zu Periode $t+1$ unterstellt [Scheer 1983, S. 111], der mit einer additiven Störvariablen überlagert ist, ergibt sich die funktionale Form zur Prognose von h Perioden in die Zukunft nach als

¹ Die Bezeichnung erfolgt zur Unterscheidung der Verfahren: NF1 gilt für stationäre, das später vorgestellte NF3 für saisonale Zeitreihen. In Abweichung zu MAKRIDAKIS [1998] wird nachfolgend das Verfahren des Naiven Trends als NF2 bezeichnet, und nicht das Verfahren der Naiven Saison.

² Obwohl das Verfahren der doppelten Glättung nur einen Parameter α besitzt, wird es korrespondierend zu der Linearen ES zur Prognose trendbehafteter Zeitreihen mit zwei Parametern auch als exponentielle Glättung zweiter Ordnung bezeichnet [Scheer 1983, S. 107].

$$\hat{y}_{t+h} = L_t + T_t \cdot h \quad . \quad (3.59)$$

Dies entspricht der konventionellen Modellierung linearer Trends durch Verfahren der linearen Regression und der Methode der kleinsten Quadrate mit der Zeit t als unabhängigen Variablen [Scheer 1983, S. 107]. Das Modell der D.ES zur Schätzung der Parameter L_t und T_t besteht aus den zwei Gleichungen

$$\hat{y}_t^{(1)} = \alpha \cdot y_t + (1 - \alpha) \cdot \hat{y}_{t-1}^{(1)} \quad , \quad (3.60)$$

$$\hat{y}_t^{(2)} = \alpha \cdot \hat{y}_t^{(1)} + (1 - \alpha) \cdot \hat{y}_{t-1}^{(2)} \quad , \quad (3.61)$$

wobei Gleichung (3.60) weitgehend mit der Gleichung (3.47) der S.ES unter Verwendung des Glättungsparameters α , mit $0 \leq \alpha \leq 1$, übereinstimmt und das Ergebnis als Ausgangswert in Gleichung (3.61) erneut geglättet wird. Aus den gleitenden Mittelwerten $\hat{y}_t^{(1)}$ und $\hat{y}_t^{(2)}$ ergeben sich durch Auflösung der Bestimmungsformeln die Schätzwerte für L_t und T_t als

$$L_t = 2 \cdot \hat{y}_t^{(1)} - \hat{y}_t^{(2)} \quad , \quad (3.62)$$

$$T_t = \left(\hat{y}_t^{(1)} - \hat{y}_t^{(2)} \right) \cdot \frac{\alpha}{1 - \alpha} \quad . \quad (3.63)$$

Aufgrund der unmittelbaren Anwendbarkeit der Gleichungen zur Prognose sei in Bezug auf ihre Herleitung auf HANSMANN [1983, S. 33-36] verwiesen. Alternativ erlaubt eine Kombination der Formeln (3.60) und (3.61) in (3.62) und (3.63) nach [Harrison 1967, S. 836; Schröder 1978, S. 52] eine unmittelbare Fortschreibung und Anwendung von L_t und T_t mit

$$L_t = y_t + (1 - (1 - \alpha)^2) \cdot (y_t - \hat{y}_t^{(2)}) \quad , \quad (3.64)$$

$$T_t = T_{t-1} + \alpha^2 \left(y_t - \hat{y}_t \right) \quad , \quad (3.65)$$

zum Einsetzen in Gleichung (3.59) [Scheer 1983, S. 110]. Die doppelte Exponentielle Glättung besitzt gegenüber herkömmlichen Verfahren der linearen Trendextrapolation durch regressionsanalytische Verfahren die Fähigkeit, lokale Veränderungen der Trendstruktur in Abhängigkeit des gewählten Glättungsparameters mehr oder weniger stark zu berücksichtigen [Scheer 1983, S. 109]. Zudem besteht die Möglichkeit, das Modell durch weitere Glättung auf polynomiale Trends höherer Ordnung zu erweitern [Scheer 1983]. Allerdings bedingt die Beschränkung auf einen α -Parameter die identische Glättung der Niveauwerte und der Trendwerte [Scheer 1983, S. 110]. Zur Berücksichtigung abweichender Glättungsintensitäten wurde von HOLT die Methode der Linearen Exponentiellen Glättung entwickelt, die nachfolgend skizziert wird.

3.2.3.3.3 Lineare Exponentielle Glättung

HOLT erweiterte die Grundidee der einfachen Exponentiellen Glättung nach BROWN (siehe Abschnitt 3.2.3.2.3) auf die separate Modellierung und Glättung trendbehafteter, saisonaler und anderer regelmäßiger Komponenten in einem unveröffentlichten Memorandum [Holt 1957]. Sowohl die nachfolgend dargestellte Lineare Exponentielle Glättung (L.ES) zur Prognose trendbehafteter Zeitreihen als auch die Saisonale Exponentielle Glättung nach HOLT-WINTERS gehen darauf zurück [Holt 2004a; Holt 2004b].

Zur Prognose trendbehafteter Zeitreihen postuliert HOLT die Verwendung von zwei Glättungsparametern α und β , ebenfalls mit $0 \leq \alpha, \beta \leq 1$, und drei ähnlichen Gleichungen auf der Basis der aus (3.59) bekannten linearen Trendgleichung

$$\hat{y}_{t+h} = L_t + T_t \cdot h \quad . \quad (3.66)$$

Er formuliert zwei Gleichungen: die erste zur Glättung des Niveaus der Zeitreihe unter Verwendung des α -Parameters zur Korrektur des Niveaus L_t um den Trendwert T_t der vorigen Beobachtungsperiode [Makridakis et al. 1998, S. 158]:

$$L_t = \alpha \cdot y_t + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad . \quad (3.67)$$

Die zweite Gleichung schätzt die Trendsteigung der Zeitreihe in t als Differenz zweier aufeinander folgender Zeitreihenniveaus L_t und L_{t-1} , die ähnlich der Formel der S.ES die Glättung der Trendkomponente über die Zeit erlaubt, mit

$$T_t = \beta \cdot (L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad . \quad (3.68)$$

Der Parameter β spezifiziert analog zur S.ES den Glättungsparameter der Trendänderung. In Anlehnung an die zwei Glättungsparameter wird das Verfahren auch als Exponentielle Glättung zweiter Ordnung bezeichnet. Für $\alpha = \beta$ kann die Methode der D.ES nach BROWN als Sonderfall von HOLT'S L.ES interpretiert werden [Makridakis et al. 1998]. Als Nachteil der L.ES-Verfahren sind die Notwendigkeit der Schätzung zweier Startwerte im Gegensatz zu nur einem bei S.ES und D.ES [Scheer 1983, S. 111] und die erhöhte Komplexität der Findung von zwei Parametern α und β durch Verfahren der Optimierung oder der schrittweisen Parametersuche (engl. *grid search*) zu nennen.

3.2.3.3.4 Exponentielle Glättung mit gedämpftem Trend

Bisherige Verfahren der Exponentiellen Glättung extrapolieren einen linearen Trend uneingeschränkt über den Prognosehorizont h in die Zukunft, was unter betrieblichen Anwendungsaspekten als wenig realistisch erscheint. So ist insbesondere durch die nichtlineare Annäherung negativer Trendverläufe an null sowie durch Annäherung an Sättigungsgrenzen des Absatzes im Markt ein degressiver Verlauf des Trends anzunehmen. Entsprechend erweiterte GARDNER und MACKENZIE [1989] die Grundidee der trendbehafteten Exponentiellen Glättung mit linearen Trends der D.ES und L.ES auf die Berücksichtigung additiver degressiver Trends, was sie als „dampened trend exponential smoothing“ durch einen „gedämpften Trend“ bezeichneten.

Das Verfahren stellt eine einfache Erweiterung des L.ES dar, indem die Formeln (1.44 ff.) der Trendkomponente um einen Dämpfungsfaktor θ , mit $0 < \theta \leq 1$, erweitert werden:

$$L_t = \alpha \cdot y_t + (1 - \alpha)(L_{t-1} + \theta T_{t-1}) \quad , \quad (3.69)$$

$$T_t = \beta \cdot (L_t - L_{t-1}) + (1 - \beta)\theta T_{t-1} \quad , \quad (3.70)$$

die kombiniert die Formel zur gedämpften Trendextrapolation in die Zukunft ergeben:

$$\hat{y}_{t+h} = L_t + \sum_{i=1}^h \theta^i T_t \quad . \quad (3.71)$$

Für $\theta = 1$ ergibt sich als Sonderfall das konventionelle Verfahren der L.ES. Auf eine Darstellung der verwandten, aber in der Praxis wenig verbreiteten Verfahren des „robust trend exponential smoothing“ [Grambsch und Stahel 1990] oder PARZENS „long-term trend exponential smoothing“ wird nachfolgend verzichtet.

3.2.3.3.5 ARIMA-Verfahren

Um die Klasse der ARMA-Verfahren auf trendbehaftete, nicht-stationäre stochastische Prozesse anwenden zu können, müssen die instationären Zeitreihen zunächst in mittelwert- und varianzstationäre Zeitreihen transformiert werden [Hansmann 1983, S. 72]. Dies erfolgt durch eine iterative Modellidentifikation, Datenvorverarbeitung und Transformation der Rohdaten im Rahmen der BOX-JENKINS-Methodologie. Eine Instationarität des Mittelwertes, z. B. durch lineare Trends oder Niveauverschiebungen, wird durch Bildung der ersten Differenz y'_t der Zeitreihe, auch als Integration der Zeitreihe bezeichnet, eliminiert. Der Grad der Integration kann in der Notation vergangener Beobachtungswerte oder in Analogie zur Formulierung der AR(p)- und MA(q)-Terme unter Verwendung des Backshift-Operators B dargestellt werden [Makridakis et al. 1998, S. 334]:

$$\begin{aligned} y'_t &= y_t - y_{t-1} \\ &= (1 - B)y_t \end{aligned} \quad (3.72)$$

Instationarität höherer Ordnung erfordert entsprechend die Bildung von Differenzen höherer Ordnung, z. B. zweiter Differenzen y''_t zur Bereinigung quadratischer Trends durch

$$\begin{aligned} y''_t &= y'_t - y'_{t-1} = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2} \\ &= (1 - B)^2 y_t \end{aligned} \quad (3.73)$$

bzw. n -fache Differenzbildung für eine polynomiale Instationarität n -ten Grades. Für nicht-polynomiale Trends logarithmischer, exponentieller oder unregelmäßig strukturierter Form sei dagegen auf die modifizierte Integration nach HANSMANN verwiesen [Hansmann 1983, S. 73]. Für eine Differenz der Ordnung d verkürzt sich die Zeitreihe aus n Beobachtungswerten von y_t, \dots, y_{t-n+1} entsprechend um d Beobachtungswerte auf $n-d$ Beobachtungswerte. Instationaritäten der Varianz des stochastischen Prozesses werden durch alternative Transformationen kompensiert.

Unter Kombination des Differenzen-Terms mit dem ARMA-Modell ergibt sich ein um die Integration (I) erweitertes ARIMA-Modell aus den drei Komponenten des AR(p)-Prozess, MA(q)-Prozess und dem Grad der Integration $I(d)$ in der Backshift-Notation als

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - B)^d y_t = c + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) e_t \quad (3.74)$$

Das Vorgehen der Modellierung erfolgt entsprechend der ARMA-Methodologie für stationäre stochastische Prozesse. Zur Kennzeichnung der iterativen Integration der Zeitreihen wird die ursprüngliche ARMA(p, q)-Notation um den Grad der Integration bzw. Differenzbildung d zu ARIMA(p, d, q) erweitert.¹

¹ Trotz umfangreicherer Notation bezeichnet z. B. ein ARIMA($p, 0, 0$), ARMA($p, 0$) oder AR(p) denselben Prozess. Dieser wird nachfolgend um saisonale Terme erweitert.

3.2.3.4 Verfahren bei saisonalem und trend-saisonalen Bedarfsverlauf

3.2.3.4.1 Naives Saisonales Verfahren

Wird eine stationäre Zeitreihe sowohl durch Trends als auch durch additive oder multiplikative saisonale Komponenten überlagert,

$$y_t = f(L_t, T_t, S_t, E_t) \quad , \quad (3.75)$$

so müssen beide regelmäßigen Komponenten sowie die Form ihrer additiven oder multiplikativen funktionalen Verknüpfung in der Modellbildung berücksichtigt werden.

In Anlehnung an das Naive Verfahren des letzten Beobachtungswertes kann ein saisonales Naives Verfahren mit zeitlicher Verzögerung einer Saisonlänge verwendet werden. „Naive Forecast 3 or NF3 ... goes beyond NF1 in that it considers the possibility of seasonality in the series“ [Makridakis et al. 1998, S. 47-48]. Das Prognoseverfahren berechnet die zukünftige Realisation der abhängigen Variablen \hat{y}_{t+h} in der Periode $t+h$ aus der um eine Saison s zurückliegenden Beobachtung, also $t-s+h$, durch

$$\hat{y}_{t+h} = y_{t-s+h} \quad . \quad (3.76)$$

Der Parameter s gilt als saisonaler Shift-Operator einer Saisonlänge, die nach dem Erhebungsintervall mit dem zeitlichen Lag l korrespondiert, z. B. durch einen jährlichen Lag von $l = 12$ Monaten bzw. einer Saison $s = l = 12$ bei monatlichen Beobachtungswerten. Das Verfahren NF3 unterstellt eine Extrapolation der Saisonalität der Zeitreihe, ist jedoch bei unkorrelierten Residuen ebenso auf stationäre Zeitreihen anzuwenden.

3.2.3.4.2 Saisonale Exponentielle Glättung

Das Verfahren zur Abbildung multiplikativer saisonaler Komponenten durch eine exponentiell geglättete Anpassung der Saisonindizes stellt eine Erweiterung des linearen Modells der exponentiellen Glättung nach HOLT in Formel (3.66) dar, das durch die Veröffentlichung von WINTERS große Aufmerksamkeit erhielt [Winters 1960]. Das Verfahren kontrolliert die Glättung des Niveaus L , des Trends T durch die bekannten Parameter α und β und der Saisonalität S gesteuert über den Parameter γ . Entsprechend wird das Modell auch als exponentielle Glättung dritter Ordnung [Scheer 1983, S. 122] oder Exponentielle Glättung nach HOLT-WINTERS bezeichnet.

Die Prognosegleichung der L.ES wird um den Saisonfaktor S_{t-s+h} der Periode $t-s+h$ erweitert, wobei s wie zuvor die Anzahl der Perioden eines Saisonzyklus beschreibt und $t-s+h$ jene Saisonperiode des Vergangenheitszeitraums identifiziert, in der der letzte Saisonfaktor beobachtet und berechnet wurde [Scheer 1983, S. 121]:

$$\hat{y}_{t+h} = (L_t + T_t \cdot h) S_{t-s+h} \quad . \quad (3.77)$$

WINTERS unterstellt somit einen linearen Trend und eine multiplikativ verknüpfte Saisonalität. Korrespondierend zur L.ES erfolgt die Bestimmung und Fortschreibung der Parameter für Niveau L_t und Trend T_t unter Saisonbereinigung nach

$$L_t = \alpha \cdot \frac{y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad , \quad (3.78)$$

wobei der Grundwert der Niveauveränderung in t durch den letzten vorliegenden Saisonfaktor aus $t-s$ bereinigt wird. Die Berechnung des Trendeinflusses erfolgt identisch wie in (3.80)

$$T_t = \beta \cdot (L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (3.79)$$

Der Saisonfaktor wird als geglätteter Quotient des aktuell beobachteten Absatzwertes y_t und des Grundwertes L_t in Analogie zur Grundüberlegung exponentieller Glättung berechnet, mit

$$S_t = \gamma \cdot \frac{y_t}{L_t} + (1 - \gamma)S_{t-s} \quad (3.80)$$

Somit berechnet die exponentielle Glättung nach HOLT-WINTERS zukünftige Werte durch Desaisonalisierung der geglätteten Niveau- und Trendwerte durch exponentiell geglättete Saisonindizes. Somit können auch Veränderungen der Saisonalität über die Zeit abgebildet werden, wenn auch zeitlich verzögert um eine Saisonlänge s .¹

Zur Modellierung additiver Saisonalität werden die Gleichungen (3.78) und (3.80) unter Konstanz der anderen Gleichung entsprechend modifiziert [Makridakis et al. 1998, S. 169] zu

$$\hat{y}_{t+h} = L_t + T_t \cdot h + S_{t-s+h} \quad (3.81)$$

und

$$L_t = \alpha \cdot (y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (3.82)$$

Das Verfahren der saisonalen Exponentiellen Glättung nach HOLT-WINTERS kann ebenfalls auf rein saisonale Zeitreihen ohne Trendkomponente der generellen Form

$$y_t = f(L_t, S_t, E_t) \quad (3.83)$$

übertragen werden. Hierzu werden einfach $\beta = 0$ und $T_0 = 0$ gesetzt. Entsprechend kann es ebenfalls als D.ES nach HOLT parametrisiert werden, indem $\gamma = 0$ und $L_{t-s} = 1$ gesetzt werden, sowie als S.ES nach Brown durch Kombination der vorangegangenen Parametereinstellungen.²

Die saisonale Variante der ES erfordert ebenfalls die Ex-ante-Schätzung oder Festlegung valider Startparameter L_0, T_0 sowie S_0, \dots, S_{s-1} für die Parameterberechnung [Scheer 1983, S. 122]. Daneben zeichnet sich das Verfahren, so wie alle Verfahren der Klasse der Exponentiellen Glättung, durch eine einfache Berechnungsweise und geringe Datenanforderungen aus, da lediglich die Glättungsparameter α, β, γ , die Saisonparameter S und die Werte L_t, T_t gespeichert werden müssen [Scheer 1983, S. 122]. Die Bestimmung der Glättungsparameter gestaltet sich durch die simultane Festlegung von α, β, γ , mit $0 \leq \alpha, \beta, \gamma \leq 1$, trotz der relativ geringen Sensitivität der Prognosequalität auf Parameterveränderungen komplex.³ Die

¹ Alternative Ansätze der Regression modellieren Saisonindizes oder kombinierte Trend-Saison-Gleichungen vielfach mit binären Dummy-Variablen durch Kleinste-Quadrate-Schätzer mit konstanter Gewichtung aller Beobachtungswerte [Scheer 1983, S. 113-120].

² Diese Fähigkeit zur Abbildung aller Verfahrensformen durch geeignete Parametrisierung erübrigt jedoch nicht die Frage der validen Modellselktion, da die Anwendung einer D.ES auf stationäre Zeitreihen oder einer saisonalen exponentiellen Glättung auf nicht-saisonale Zeitreihen erhebliche Verzerrungen und eine verringerte Prognosegenauigkeit bedingt.

³ Dabei sind die möglichen Parameterwerte zwischen 0 und 1 auf sinnvolle Intervalle von $0,01 \leq \alpha, \beta \leq 0,4$ [Götz 1999, S. 112] und Werten von γ bei wenigen vollständig vorhandenen Saisons auf $\gamma > 0,5$ zu beschränken. Optimale Parameterwerte außerhalb der genannten Grenzen können zudem auf eine Missspezifikation

Parameterbestimmung erfolgt entweder mittels Optimierungsverfahren des Gradientenabstiegs [Winters 1960], der nichtlinearen Optimierung [Scheer 1983, S. 124] oder eines strukturierten „grid search“ durch schrittweise Berechnung aller Parameterkombinationen mit anschließendem Vergleich und Auswahl des Kandidaten mit der besten Parameterkombination. Da das Verfahren nach HOLT-WINTERS in jeder Periode jeweils einen Saisonfaktor der Teilperiode neu berechnet und die $s-1$ anderen Werte unverändert bleiben, wirken sich Veränderungen der Saisonstruktur in t frühestens um 1 Saison s verzögert auf den Prognosewert in $t+s$ aus. Dies kann durch eine alternative Ermittlung der Saisonfaktoren als periodische Funktionen der Fourierkoeffizienten im Modell von HARRISON behoben werden [Harrison 1967; Scheer 1983, S. 124-131], das jedoch aufgrund der geringen empirischen Verbreitung – trotz seiner möglichen Relevanz bei kurzen Bedarfszeitreihen mit nur wenigen vollständigen Saisonzyklen – vernachlässigt wird [Hansmann 1980, S. 235; Lewandowski 1974, S. 111].

3.2.3.4.3 Saisonale ARIMA-Verfahren

Die Verfahrensklasse zur Prognose von ARIMA-Prozessen erlaubt auch eine Erweiterung zur Abbildung von saisonalen Prozessen. Analog zur zeitlichen Struktur nichtsaisonalen ARIMA(p,d,q)-Komponenten können die Beobachtungswerte y_{t-l} saisonale Strukturen der Autokorrelation, Moving Averages oder der Integration für $l = s, 2s, 3s, \dots, ns$ aufweisen.

Saisonale Instationaritäten des Mittelwertes oder der Varianz erfordern analog zu nichtsaisonalen Mustern die Bildung stationärer stochastischer Prozesse durch Integration mittels saisonaler Differenzen

$$\begin{aligned} y_t^{i's} &= y_t^i - y_{t-s}^i \\ &= (1 - B^s) y_t^i \end{aligned} \quad (3.84)$$

wobei s die Anzahl der Perioden pro Saison angibt, z. B. $s = 12$ für monatliche Zeitreihen, $s = 4$ für quartalsweise Zeitreihen etc. [Makridakis et al. 1998]. Der maximale Grad der saisonalen Integration D spezifiziert die saisonale Modellform $I(D)_s$.

Durch die Analyse der absoluten und der partiellen Autokorrelationskoeffizienten r_{t-s} zum saisonalen Lag $t-s$ erfolgt die Identifikation saisonaler Autokorrelationen und Moving-Average-Komponenten. Das ARIMA-Modell aus Gleichung (3.74) wird entsprechend um zusätzliche Terme der saisonalen autoregressiven Komponenten mit den Parametern Φ_P sowie der saisonalen Moving-Average-Komponenten mit den Parametern Θ_Q und der saisonalen Differenzen erweitert zu

$$\begin{aligned} & (1 - \phi_1 B - \phi_2 B^2 \dots - \phi_p B^p) (1 - \Phi_1 B - \Phi_2 B^2 \dots - \Phi_p B^p) (1 - B)^d (1 - B^s)_s y_t \\ &= c + (1 - \theta_1 B - \theta_2 B^2 \dots - \theta_q B^q) (1 - \Theta_1 B - \Theta_2 B^2 \dots - \Theta_Q B^Q) e_t \end{aligned} \quad (3.85)$$

Korrespondierend wird die Notation um die saisonalen Charakteristika auf ARIMA(p,d,q)(P,D,Q) $_s$ erweitert [Makridakis et al. 1998, S. 246].¹

Es wurde gezeigt, dass spezifische ARIMA-Modelle den Verfahren der Exponentiellen Glättung entsprechen [Chatfield und Yar 1991]. So repräsentiert ein ARIMA(0,1,1)-Modell

des Modells hindeuten, z. B. ein $\alpha = 0,99$ unterstellt einen instationären Prozess ähnlich dem „random walk“, der die Anwendung der einfachen Exponentiellen Glättung verbietet.

¹ Sie werden zur Unterscheidung z. T. als SARIMA-Verfahren bezeichnet, obwohl eine Trennung der saisonalen und nichtsaisonalen Modelle eines Verfahrens mit integrierter Methodologie wenig trennscharf erscheint.

das Verfahren der einfachen Exponentiellen Glättung S.ES mit $\theta = 1 - \alpha$. HOLT's L.ES ist äquivalent zu einem ARIMA(0,2,2)-Modell mit $\theta_1 = 2 - \alpha - \alpha\beta$ und $\theta_2 = \alpha - 1$. Die saisonale Exponentielle Glättung des additiven HOLT-WINTERS-Modells entspricht einem ARIMA(0,1, $s+1$)(0,1,0)_s-Modell, während die multiplikative HOLT-WINTERS ES kein ARIMA-Äquivalent aufweist.

Auch wenn die manuelle Identifikation der stochastischen Prozesse und die entsprechende Spezifikation eines ARIMA-Verfahrens wegen der iterativen Vorgehensweise und der aufwendigen nichtlinearen Techniken zur Parameterschätzung mit hohem Aufwand verbunden ist [Götz 1999, S. 119; Montgomery et al. 1990], existieren ausgewählte Softwarelösungen der Expertensysteme, z. B. Autobox, ForecastPro oder SAS Forecast Server. Diese ermöglichen eine vollautomatische Modellierung je Zeitreihe und somit den Einsatz in Prognosesystemen des Handels.

3.2.4 Objektive Prognoseverfahren der Kausalanalyse

3.2.4.1 Dynamische Regression

Zeitreihenanalytische Modelle erlauben keine Berücksichtigung vorhandener Informationen über exogene Einflüsse, selbst wenn diese deterministisch und bekannt sind und ohne Kosten der Informationsbeschaffung vorliegen, z. B. als Zeitreihen bekannter Kalendereffekte durch Urlaubszeiten, Feiertage mit wechselndem Datum, unternehmensinterne Preisänderungen, Werbemaßnahmen oder Streiks. Diesen deterministischen und i. d. R. a priori bekannten externen Effekten kommt bei der validen Bedarfsprognose eine hohe Bedeutung aufgrund der signifikanten Beeinflussung des Bedarfs und somit der Prognosegenauigkeit zu [Götz 1999, S. 118]. In der Praxis erfordert das Auftreten externer Einflüsse eine manuelle Korrektur der objektiven Prognosen um dominante externe Einflussfaktoren durch Disponenten und somit durch subjektive Verfahren, was erheblichen manuellen Aufwand bedingt und zusätzliche Fehlerquellen entstehen lässt. Im Zuge automatisierter Prognosen erfordert das Vorliegen von bedeutenden externen Einflüssen deren Berücksichtigung als kausalanalytische Modelle.

Eine kausale Modellierung als Wirkungsanalyse (vgl. Abschnitt 3.1.2.4) durch Verfahren der einfachen oder multiplen Regression vernachlässigt die Dynamik des Prognosegegenstandes in Form von Saisonalitäten, Trends oder Niveauverschiebungen sowie die Dynamik der kausalen Einflüsse auf die abhängige Variable. Dynamische Regressionsmodelle [Pankratz 1991] und Transferfunktionsmodelle (engl. *transfer function models* oder *autoregressive distributed lag models*) mit ARIMA-Fehlertermen [Box und Jenkins 1970] erlauben darüber hinaus die Modellierung dynamischer kausaler Effekte der abhängigen auf die unabhängigen Variablen. Der Effekt der unabhängigen Variable x_t muss dabei nicht gleichzeitig in t auftreten, sondern kann über mehrere Perioden $x_t, x_{t-1}, x_{t-2}, \dots$ bzw. $y_t, y_{t+1}, y_{t+2}, \dots$ verteilt sein. Beeinflussen zeitlich verzögerte Realisationen der abhängigen Variablen auch die Ausprägung der unabhängigen Variablen, mit

$$\begin{aligned} x_t &= f(x_{t-1}, \dots, y_t, \dots) + \varepsilon_t \\ y_t &= f(y_{t-1}, \dots, x_t, \dots) + \varepsilon_t \end{aligned} \quad (3.86)$$

erfordert dies die simultane Schätzung in Mehrgleichungsmodellen der multivariaten Autoregressiven Modelle bzw. Vektor-Autoregression (VAR), die hier jedoch nicht betrachtet

werden. Eine Einordnung der Modellformen der dynamischen Regression mit einer abhängigen Variablen in die kausalanalytische Modellierung erfolgt in Abbildung 3.19.¹

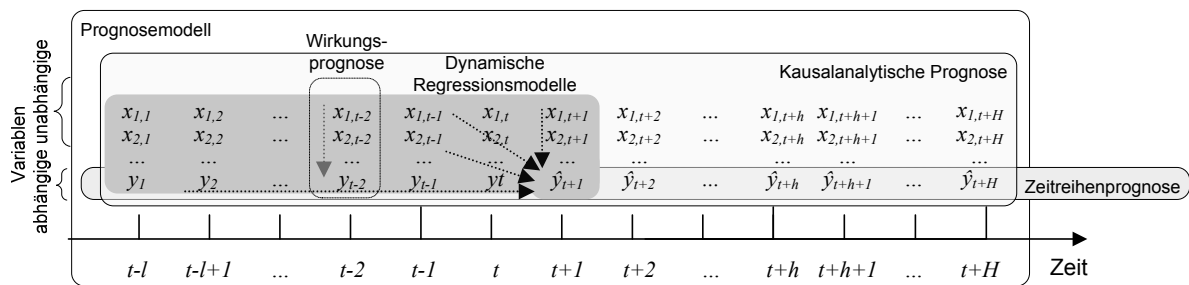


Abbildung 3.19: Prognosemodell der dynamischen Regression zur kausalen Prognose

Eine umfassende Darstellung der Klasse dynamischer Regressionsmodelle und VAR-Modelle übersteigt den Rahmen dieser Untersuchung. Daher wird an dieser Stelle auf die Literatur der Regression mit ARIMA-Fehlertermen [Makridakis et al. 1998], der dynamischen Regression [Diebold 2001; Harvey 1993] sowie der multivariaten Vektor-Autoregressiven-Modelle [Litterman 1986] verwiesen.

3.2.4.2 Verfahren der Interventionsmodellierung

Aufgrund der Komplexität umfassender Modelle der dynamischen Regression² werden selbst einfache Ansätze nur von wenigen Innovationsführern des Handels zur Prognose eingesetzt (vgl. Abschnitt 3.1.3.3). Die wenigen Ansätze verfolgen Sonderfälle der statischen Regression mit dynamischen ARIMA-Fehlertermen, die exogene Einflüsse durch binäre Erklärungsvariablen im Sinne von Interventionen auf den zugrundeliegenden stochastischen Prozess berücksichtigen. Diese sogenannten Interventionsmodelle [Box und Tiao 1975] werden in Form von ARIMA-Modellen mit exogenen Eingabewerten (ARIMAX) oder durch Regressionsmodelle mit AR-Fehlertermen (ARX) gebildet, die eine Berücksichtigung gleichzeitiger Einflüsse der exogenen Variablen x_i auf den Prognosegegenstand y_i in einer Periode t ohne Berücksichtigung dynamischer kausaler Effekte erlauben. Da sie ebenfalls die kausalanalytische Erweiterung von Neuronalen Netzen zur Prognose erlauben, werden sie hier gleichfalls betrachtet und ihre Modellform in Abbildung 3.20 eingeordnet.

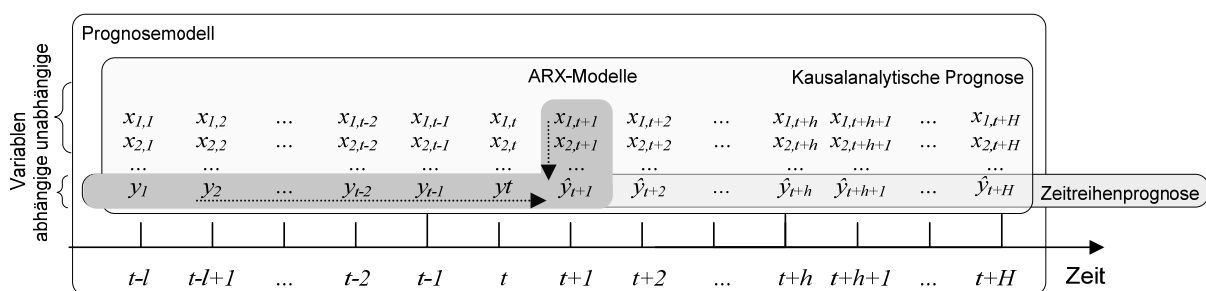


Abbildung 3.20: ARX(p)-Prognosemodell zur kausalen Prognose mit Interventionen

Ein ARX-Modell beschreibt die kausalanalytische Erweiterung eines Autoregressiven AR(p)-Prognosemodells der Zeitreihenanalyse, das den grundlegenden stochastischen Prozess für

¹ Statische Interdependenzen zwischen der abhängigen Variablen und den unabhängigen Variablen, z. B. des Absatzes auf Werbe- und Preisanstrengungen, erfordern theoretisch die simultane Bestimmung in ökonomischen Mehrgleichungssystemen. Aufgrund ihrer Komplexität wird darauf jedoch verzichtet.

² Modelle der dynamischen Regression stellen hohe Anforderungen an Experten in der manuellen Modellierung und sind bislang nur vereinzelt in Expertensystemen automatisiert, z. B. „Autobox“ [Reilly 2000], „TRAMO/SEATS“ (Time Series Regression with ARIMA Noise, Missing Observations and Outliers) oder „SCA“. Auch dort sind sie aufgrund des hohen Ressourcenbedarfs nur bedingt auf Massendaten wie etwa im Handel anwendbar.

zurückliegende Beobachtungen $y_t, y_{t-1}, \dots, y_{t-l}$ ohne kausale Interventionen beschreibt. Dieses Grundmodell wird durch eine unabhängige Interventionsvariable $x_{i,t}$ im Zeitpunkt t überlagert, die ausgewählte kausale Komponenten als Effekte eine bekannte externe Störung des zugrunde liegenden Prozesses darstellen [Götz 1999, S. 118; Makridakis et al. 1998, S. 389]. Als Intervention gilt „a one-off event which has an impact on the forecasting variable“ [Makridakis et al. 1998, S. 418], wodurch insbesondere Ausreißer, Werbeaktionen oder Kalendereffekte abgebildet werden können. Die aus (3.51) bekannte Formel eines schwach stationären AR(p)-Prozesses wird um r unabhängige Variablen $x_{r,t}$ ($r = 1, \dots, R$) mit Wirkung im Zeitpunkt t erweitert zu

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varphi_1 x_{1,t} + \varphi_2 x_{2,t} + \dots + \varphi_r x_{r,t} e_t \quad (3.87)$$

Der Parameter φ_r bezeichnet den Interventions-Regressionskoeffizienten, der die Intensität sowie die positive bzw. negative Wirkungsrichtung der Interventions-Variablen im Zeitpunkt t durch eine lineare Verschiebung des Niveaus der Prognose abbildet. Die Modelle werden analog zu anderen Modellen durch Verfahren der nichtlinearen kleinsten Quadrate geschätzt [Götz 1999, S. 120-123]. Eine Erweiterung zu ARIMAX-Modellen findet sich bei [Makridakis et al. 1998, S. 390-402].

ARX-Modelle der in (3.87) spezifizierten Form implizieren eine einmalige, externe Intervention auf die Systematik der Zeitreihe, im Sinne eines rein temporären Effektes im Zeitpunkt u ohne dynamische Effekte in Folgeperioden und einem unmittelbarem Rückgang auf den zugrunde liegenden stochastischen Prozess [Makridakis et al. 1998, S. 422], z. B. durch Ausreißer oder fehlende Beobachtungswerte. Die Modellierung der Intervention erfolgt als Zeitreihe von Binär- oder Dummyvariable in der Form

$$x_{i,t} = \begin{cases} 0 & t \neq u \\ 1 & t = u \end{cases} \quad (3.88)$$

Zeitlich verzögerte Effekte der Intervention in Form einer „decayed response intervention“ sowie Niveauverschiebungen durch „step-interventions“ im Sinne von

$$x_{i,t} = \begin{cases} 0 & t < u \\ 1 & t \geq u \end{cases} \quad (3.89)$$

oder „delayed response step-interventions“ mit zeitlicher Verzögerung der Niveauverschiebung [Makridakis et al. 1998, S. 418-423], erfordern dagegen die Erweiterung des Erklärungs- und Prognosemodells zu dynamischen Regressionsmodellen. Da diese die Modellierungskomplexität signifikant erhöhen werden sie nachfolgend vernachlässigt. ARX-Modelle repräsentieren jedoch einen effizienten Übergang zeitreihenanalytischer Modellierung zu komplexen, kausalanalytischen Modellansätzen der dynamischen Regressions- und multivariaten ARIMA-Modelle, und werden für deterministische Effekte im Verfahrensvergleich des Kapitel 6 berücksichtigt.

3.3 Beurteilung von Prognoseverfahren

3.3.1 Wirtschaftlichkeit und Prognosegenauigkeit

Die Entwicklung und Anwendung von Prognoseverfahren in der betrieblichen Planung erfordert eine Beurteilung ihrer Wirtschaftlichkeit im Vergleich zu konkurrierenden Verfahren¹, die ARMSTRONG [2001a] zufolge nur durch eine Kosten-Nutzen-Analyse erfolgen kann.

Die Beurteilung der Kosten eines Prognoseverfahrens betrachtet die Nutzung betrieblicher Ressourcen der Verfahrensanwendung im Unternehmen als fixe und variable Kostengrößen, z. B. durch abweichende Daten- oder Speicherplatzanforderungen, unterschiedliche Rechenzeiten, Personaleinsatz etc. [Hansmann 1983, S. 141f.]. Die Betrachtung umfasst dabei alle Phasen des Lebenszyklus von der Modellentwicklung bzw. Beschaffung, der Implementierung in Informationssystemen und der Parametrisierung über die Anwendung bis hin zur regelmäßigen Wartung im Sinne aller anfallenden Kosten von Investitionsgütern (engl. *total cost of ownership*). Die Kosten eines Verfahrens werden dabei nachhaltig durch seine Komplexität in Entwicklung, Implementierung und Anwendung im Sinne der „Schwierigkeit der Prognose“ [Armstrong 2001a], seine Flexibilität bei der Anpassung an veränderte Entscheidungssituationen und durch den Grad seiner Automatisierung beeinflusst.

Der Nutzen eines Verfahrens entsteht dagegen durch seine Prognosequalität gemessen anhand der Genauigkeit und Robustheit der Prognose, die durch eine verbesserte Einschätzung der Unsicherheit alternativer Zukunftsszenarien eine verbesserte Entscheidungsfindung erlaubt [Armstrong 2001a, S. 464]. Während strategische und taktische Effekte einer generell verbesserten Voraussicht und Marktkenntnis nur schwer quantifizierbar sind, kann der operationale Nutzen einer verbesserten Prognosequalität explizit durch Kostensenkungen, z. B. durch verringerte Warenbestände aufgrund der Dispositionsentscheidung, quantifiziert werden. Der quantifizierte Nutzen kann den Kosten konkurrierender Prognoseverfahren gegenübergestellt werden, somit kann eine wirtschaftliche Entscheidung über die Anwendung eines objektiven oder subjektiven Verfahrens sowie konkurrierender Verfahrenskandidaten eines Prognosemodells erfolgen. HANSMANN [1983, S. 141] fordert dementsprechend die gemeinsame Erörterung der Prognosegenauigkeit, der Prognosekosten und der Komplexität bzw. Benutzerfreundlichkeit der zur Wahl stehenden Verfahren.

Grundsätzlich kann keine positive Korrelation zwischen der Komplexität eines Verfahrens und seiner Genauigkeit festgestellt werden. Einfache Verfahren zeigen auf unterschiedlichsten Datensätzen häufig eine äquivalente oder sogar bessere Ergebnisqualität bei niedrigeren Implementierungskosten [Makridakis und Hibon 2000; Fildes et al. 1998]. Dies erfordert die individuelle Beurteilung des Nutzens geeigneter Modelle und Verfahren je Entscheidungsproblem und Datensatz. Da die vorliegende Arbeit den Vergleich etablierter Prognoseverfahren zur Warendisposition mit Verfahren der Neuronalen Netze losgelöst von der Investitionsentscheidung ihrer Implementierung im Betrieb betrachtet, rückt eine Beurteilung der Prognosequalität in den Vordergrund.²

¹ Auf eine Beurteilung von Prognosemodellen oder -verfahren nach alternativen Kriterien, insbesondere ihrer Anwendbarkeit, Komplexität, Flexibilität und Stabilität, wird zugunsten der Übersichtlichkeit der vorliegenden Arbeit verzichtet. Die Relevanz routinemäßiger, operativer Entscheidungsunterstützung erlaubt nachfolgend die Beschränkung der Diskussion auf die Prognosequalität einzelner Verfahren. Für eine Darstellung der Vielzahl alternativer Taxonomien relevanter Beurteilungsmaße und Kriterienkataloge sei auf [Brockhoff 1977, S. 52] verwiesen.

² Nach BROCKHOFF [1977, S. 51] ist die Verfahrensauswahl nach der Prognosequalität eines Verfahren von den folgenden Sachentscheidungen zu trennen, obwohl sie auf die Auswahl eines Prognoseverfahrens einwirken: der Art des Detaillierungsgrades notwendiger Daten, des Zeitraums der Datenbereitstellung, des zeitlichen Abstands der Entscheidungswiederholung, der Sensitivität der Entscheidung bezüglich der Datenarten, der Sensitivität des Unternehmensziels bezüglich der Entscheidungen und der beteiligten Instanzen inklusive ihrer Interessen und Methodenkenntnisse.

3.3.2 Beurteilung der Prognosegenauigkeit

3.3.2.1 Aspekte der Beurteilung

Wie aus Abschnitt 3.1.2 f. hervorging, können empirische Problemstellungen in unterschiedlichen Erklärungs- und Prognosemodellen abgebildet und mit einer Vielzahl konkurrierender Prognoseverfahren gelöst werden. Das resultierende Problem der Selektion des besten bzw. die Rangreihung alternativ anwendbarer Modelle und Verfahren für eine gewählte Informationsversorgung [Granger 1980, S. 12] erfordert die valide und reliable Beurteilung der relativen Qualität der Prognose, auch als Güte oder Genauigkeit der Prognose bezeichnet, durch objektive und robuste Beurteilungskriterien. Die nachfolgende Diskussion unterschiedlicher Fehlermaße erfolgt somit im Zusammenhang der Verfahrensselektion.

Die Beurteilung der Genauigkeit eines Prognoseverfahrens wird zunächst hinsichtlich des Zeitpunktes der Beurteilung unterschieden.¹ HANSMANN [1983, S. 14; ebenso Brockhoff 1977, S. 52] unterscheidet die Prognosequalität je nach dem Zeitpunkt der Beurteilung in eine Ex-ante- und Ex-post-Qualität. Die Ex-ante-Beurteilung erfolgt noch vor der Berechnung von Prognosewerten und befasst sich mit der Angemessenheit des zugrunde gelegten Erklärungsmodells bzw. der Datenversorgung [Brockhoff 1977, S. 53]. Sie beurteilt „die Auswahl der kausalen Variablen, die zugrunde gelegte Funktionsform ..., die Länge des Beobachtungszeitraumes sowie eventuelle nachhaltige Veränderungen der Zeitreihenstruktur im Beobachtungszeitraum“ [Hansmann 1983, S. 14]. Nach SCHWARZE [1973, S. 541ff.] werden zudem die stochastischen Eigenschaften der unabhängigen und abhängigen Variablen, die Spezifizierung der Zeitstabilitätshypothese, die Verfahren der Parameterschätzung und mögliche Prognosefehler der vorherbestimmten stochastischen, unabhängigen Variablen systematisch auf Fehlerquellen überprüft.² Die Ex-ante-Qualität beurteilt so die Kongruenz zwischen Realität, Datensatz, Erklärungs- bzw. Prognosemodell und Prognoseverfahren hinsichtlich ihrer Eignung zur Problemlösung. Sie erlaubt so erste Rückschlüsse auf die zu erwartende Güte des Verfahrens. Diese Problemstellungen der Ex-ante-Beurteilung der Prognosegenauigkeit sind in der ökonometrischen Literatur ausführlich behandelt worden [Schönfeld 1969; Hochstädter und Uebe 1970; Schneeweiß 1971; Brockhoff 1977]. Dennoch kann eine abschließende Beurteilung der Prognosequalität nur ex post, nämlich durch den Vergleich der prognostizierten Werte mit den tatsächlichen eingetretenen Werten, anhand der Prognosefehler erfolgen [Hansmann 1983, S. 14].

Die Ex-ante Beurteilung eines Prognoseverfahrens ist deutlich von den Konzepten einer Ex-ante- und der Ex-post-Prognose zu trennen, die die Gestaltung unterschiedlicher Untersuchungsdesigns für empirische Vergleichsstudien bezeichnet. Eine Ex-ante-Prognose repräsentiert eine echte Prognose im Hinblick auf unbekannte Datenwerte der Zukunft. ARMSTRONG [2001a] bezeichnet hingegen auch die Ex-post-Simulation eines Verfahrens bei Aufspaltung der Datenmengen in eine Parametrisierungs- und Test-Menge als Ex-ante-Evaluation.³ Diese Prognosesimulation wird zur Unterscheidung nachfolgend als „Quasi-ex-ante“-Prognose bezeichnet. Sie erfolgt durch die Prognose aus den historischen Werten einer Kalibrationsstichprobe auf ungesehene historische Daten einer Teststichprobe [Küsters 2006,

¹ Dabei wird grundsätzlich die quantitative Prognose von Mengeneinheiten im Sinne einer Regression mit metrischen Skalenniveaus vorausgesetzt. Somit werden nachfolgend qualitative Ansätze zur Beurteilung nominal skalierten Prognosen von Wendepunkten von Richtungsänderungen vernachlässigt [Schwarze 1980; Brockhoff 1977; Theil 1966; Hansmann 1983, S. 16-17].

² Mögliche Fehler und somit suboptimale Prognoseergebnisse können ex ante z. B. durch die näherungsweise Abbildung eines Prognosemodells mit linear verknüpften Variablen bei signifikanter Nichtlinearität oder durch Abbildung eines kausalen Wirkungszusammenhangs eines dynamischen Regressionsmodells durch ein ARIMA-Verfahren aus Gründen der Wirtschaftlichkeit entstehen und akzeptiert werden [Hansmann 1983].

³ Zur besseren Unterscheidung der ökonometrischen Begriffsverwendung und der Verwendung im Zuge empirischer Vergleichsstudien wird nachfolgend der Begriff „Quasi-ex-ante-Beurteilung“ verwendet.

S. 370ff.] (vgl. Abschnitt 3.3.3). Im Gegensatz dazu stellt eine Ex-post-Prognose alle historischen Daten uneingeschränkt zur Verfügung und erlaubt so nur die Beurteilung der Anpassungsgüte, nicht jedoch der Generalisierungsleistung des Verfahrens.

Die Ex-post-Bestimmung der Prognosequalität erfolgt üblicherweise durch dimensionslose, statistische Prognosefehler und daraus abgeleitete Gütemaße [Hansmann 1983, S. 14; Granger 1980; Makridakis et al. 1998; Schwarze 1980], die eine Beurteilung der Verfahren unabhängig von der Problemstellung und vom Anwendungsbereich erlauben. Es existiert eine Vielzahl alternativer Fehler und Fehlermaße, die den absoluten Betrag der Abweichung von Prognosewert und wahren Wert unterschiedlich gewichten. Die Fehlermaße können als metrische Werte unmittelbar hinsichtlich ihrer Distanz verglichen oder durch Skalierung in ordinale Gütekriterien durch Rangbildung geordnet und robust verglichen werden. Neben wenigen frühen Untersuchungen zur Beurteilung der Prognosequalität [Brockhoff 1977; Theil 1966; Schwarze 1980; Granger 1980] entstand eine rege Diskussion und eine Vielzahl wissenschaftlicher Beiträge im Zuge der Prognosewettbewerbe [Ahlburg 1992; Armstrong und Collopy 1992; Armstrong und Fildes 1995; Armstrong et al. 1998; Fildes 1992b; Fildes et al. 1998]. Die Erkenntnisse werden nachfolgend zur Auswahl und Anwendung statistischer Fehlermaße zusammengefasst.

Bei der Darstellung ist zu berücksichtigen, dass quantitative Ex-post-Gütemaße nicht nur in der Modell- und Verfahrensevaluation zur Bestimmung der Prognosegüte anhand der Prognosefehler Verwendung finden. Darüber hinaus werden sie nach SCHWARZE [1980] auch in der Steuerung von adaptiven Prognoseverfahren durch automatische Anpassung der Verfahrensparameter (siehe 3.2.3.2.3) sowie zur regelmäßigen Kontrolle der Verfahrensgüte in der laufenden Anwendung als Warnsignale, Indikatoren, zur Signalverfolgung (engl. *tracking signal*) oder als Alarmsignale (engl. *alert*) zur Identifikation von Datenänderungen eingesetzt [Schwarze 1980, S. 318].¹ Dieselben statistischen Fehlermaße werden vielfach in der Parametrisierung der Verfahren als Zielfunktionen zur Anpassung an die untersuchte Zeitreihenstruktur verwendet. Diese werden häufig durch eine Methodologie spezifiziert, so z. B. die Verwendung quadratischer Fehlermaße als Zielfunktionen der Kleinste-Quadrate-Schätzer, oder aber sind frei wählbar, z. B. die Verfahren des Gradientenabstiegs für ARIMA-Modelle. Wird ein Verfahren auf einem Fehlermaß parametrisiert und ausgewählt, z. B. dem *MSE*, jedoch auf einem anderen Fehlermaß oder einer Kombination von Fehlermaßen beurteilt, die z. B. als *MAE* oder winsorisierte *MAPE* eine abweichende Gewichtung der Fehler implizieren, entsteht eine Diskrepanz durch systematische Verzerrungen der Prognosegüte [Zellner 1986]. Dies ist bei der Auswahl der Fehlermaße zu berücksichtigen.

Die weitgehend heterogene Definition statistischer, deskriptiver und kostenbasierter Gütemaße und die Erweiterung der Kostenfunktionen zu Zielfunktionen der Parametrisierung von neuronalen Verfahren in nachfolgenden Kapiteln erfordert eine systematische Darstellung und Diskussion in Abgrenzung zu etablierten statistischen Fehlermaßen unter Verwendung einer einheitlichen Notation zur betrieblichen Prognose. Die Vielzahl möglicher Fehlermaße entzieht sich jedoch einer vollständigen Darstellung, so dass nachfolgend nur die wichtigsten Fehler analysiert werden.

¹ SCHWARZE [1980, S. 318] bezieht seine Ausführungen jedoch explizit auf Ex-post-Prognosefehlermaße, da er deskriptive oder kostenorientierte Gütemaße vernachlässigt. Seine Ausführungen können jedoch entsprechend verallgemeinert werden. Die von SCHWARZE dort angeführte Ex-post-Beurteilung eines einzelnen Prognoseverfahrens erscheint hingegen fragwürdig, da auf diese Weise selbst vor dem Hintergrund einer kostenorientierten Evaluation keine Entscheidungsalternativen generiert werden.

3.3.2.2 Statistische Prognosefehler

Der einfache Prognosefehler ist eine reelle Zahl, die ex post als Differenz aus dem tatsächlich eingetretenen Istwert y_{t+h} und dem prognostizierten Wert \hat{y}_{t+h} für einen Prognosezeitpunkt $t+h$ in Abhängigkeit des Prognoseursprungs t und des Prognosehorizontes h berechnet wird:

$$e_{t+h} = y_{t+h} - \hat{y}_{t+h} \quad .^1 \quad (3.90)$$

Im Fall vollständiger Übereinstimmung von Prognosewert \hat{y}_{t+h} und Istwert y_{t+h} nehmen alle Prognosefehler und darauf aufbauende Fehlermaße den Wert $e_{t+h} = 0$ als Maß einer exakten Prognose an. Ziel der Anwendung jedes Prognoseverfahrens ist somit die Minimierung aller ex post bewerteten Abweichungen [Wheelwright und Makridakis 1985, S. 22].²

Der Prognosefehler in (3.90) bezeichnet eine spezifische Metrik $d(y_t, \hat{y}_{t+h})$ als Distanz- oder Abstandsfunktion zur Bestimmung des Abstandes zwischen zwei Zahlen \hat{y}_{t+h} und y_{t+h} . Neben dem einfachen Prognosefehler können als weitere Metriken die der absoluten Fehler (*AE*), der quadratischen Fehler (*SE*) und der Wurzel der quadratischen Prognosefehler (*RSE*) verwendet werden, die jeweils den absoluten Betrag des Fehlers als nichtnegative, reelle Zahl messen. In Tabelle 3.4 werden die Prognosefehler einschließlich der relativen Fehler als prozentuale Maßzahl unter Angabe ihrer gebräuchlichen englischen Bezeichnung und Abkürzung abgebildet.

Tabelle 3.4: Prognosefehler zur Beurteilung der Prognosegenauigkeit

	Einfacher Prognosefehler	Absoluter Prognosefehler	Quadratischer Prognosefehler	Wurzel des quadrat. Prognosefehlers
Fehler	$E_{t+h} = e_{t+h} = y_{t+h} - \hat{y}_{t+h}$	$AE_{t+h} = e_{t+h} $	$SE_{t+h} = e_{t+h}^2$	$RSE_{t+h} = \sqrt[2]{e_{t+h}^2}$
	Error (3.91)	absolute error (3.92)	squared error (3.93)	root of squared error (3.94)
Relative Fehler	$PE_{t+h} = \frac{e_{t+h}}{y_{t+h}}$	$APE_{t+h} = \frac{AE_{t+h}}{y_{t+h}}$	-	-
	percent error (3.95)	absolute percent error (3.96)		

Die Metriken der Fehlerwerte weisen den Abweichungen unterschiedliches Gewicht zu, welches in der grafischen Abbildung 3.21 im Vergleich zum weit verbreiteten quadratischen Fehler *SE* der Kleinste-Quadrate-Schätzer sowie der Different der beiden Prognosefehler im Intervall [-2; 2] sichtbar wird.

¹ Einzelne Autoren definieren den Prognosefehler abweichend als $e_t = \hat{y}_t - y_t$ [u.a. Schwarze 1980; Armstrong 2001a, S. 470]. Die Auswirkung auf das Vorzeichen des Fehlers ist allerdings nur für relative, nicht für quadratische oder absolute Fehlermaße relevant. Die hier gewählte Notation entspricht dagegen der konventionellen Schreibweise in Anlehnung an die Methode der Kleinsten Quadrate [Hüttner 1986].

² Aufgrund der Unsicherheit in empirischen stochastischen Prozessen wird auch bei einer exakten Messung der Beobachtungswerte, korrekten Identifikation und Modellierung des Erklärungs- und Prognosemodells und fehlerfreien Verfahrensauswahl und -parametrisierung immer eine Reststreuung zwischen dem Prognosewert und der tatsächlichen Beobachtung auftreten [Wheelwright und Makridakis 1985, S. 22].

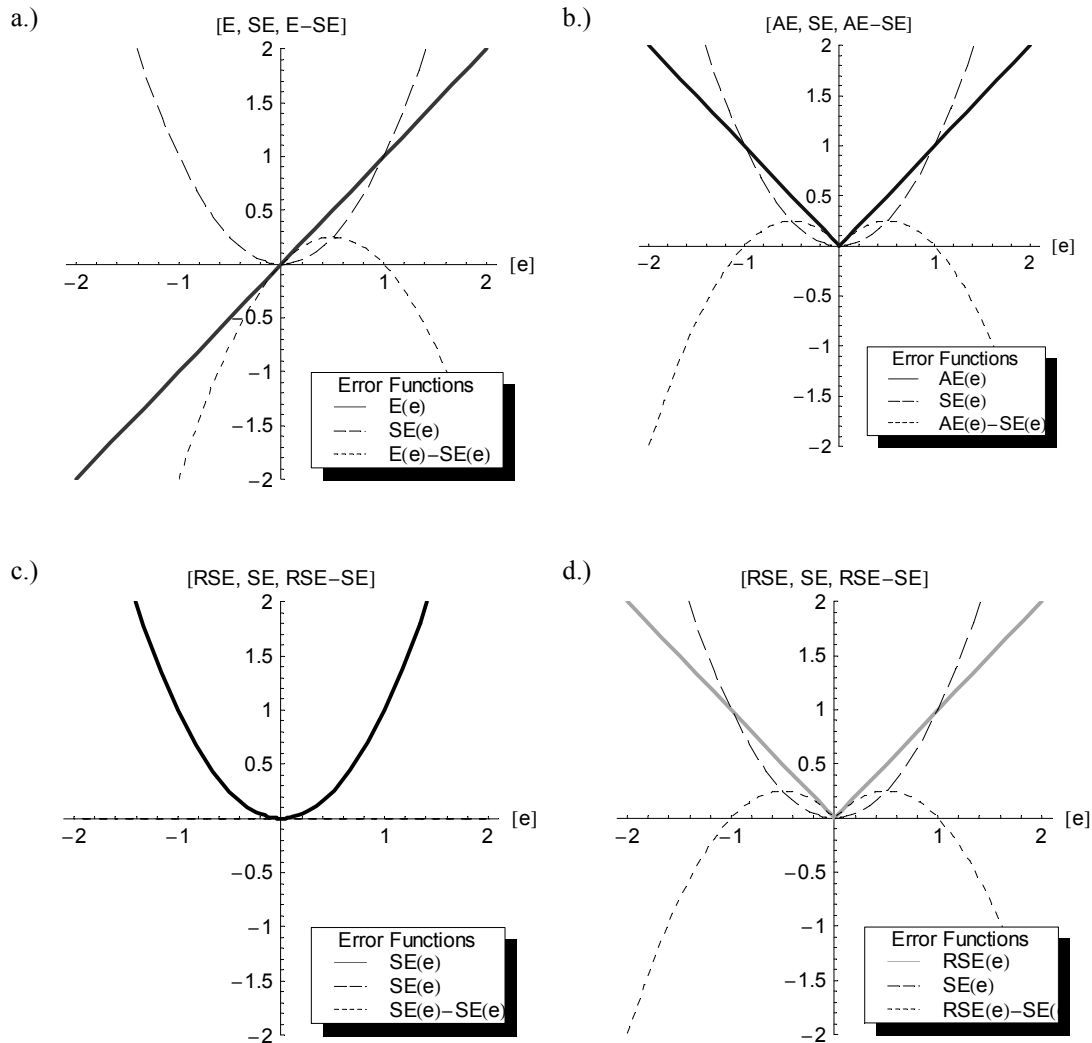


Abbildung 3.21: Fehlerverlauf des einfachen (a.), absoluten (b.), quadratischen (c.) und der Wurzel des quadratischen Prognosefehlers (d.)

In Abbildung 3.21 (a.) ist den Unterschied des einfachen Fehlerwertes E zum absoluten Betrag des Fehlers AE abgebildet, der nachfolgend noch diskutiert wird. Die Differenz der Fehlerwerte des Betrags $AE(e)$ und $SE(e)$ in Abbildung 3.21 (b.) zeigt, dass der quadratische Fehler SE höheren Abweichungen mit $e > 1$ überproportional mehr Gewicht zuweist als der absolute Fehler AE , während der SE geringen Fehlern $e < 1$ unterproportionales Gewicht zuweist. Dies bedingt eine überproportionale Gewichtung einer höheren Abweichung im Vergleich zu mehreren geringeren Abweichungen bei gleicher Fehlersumme und impliziert durch die asymmetrische Bewertung der Abweichungen eine implizite Kostenfunktion der Fehlerabweichungen. Die Wurzel des quadratischen SE in Abbildung (c.) entspricht bei einer einzelnen Beobachtung dem absoluten Fehler AE , während der typische Verlauf des populären quadratischen Fehlers SE in Abbildung d) abgebildet ist. Alle Prognosefehler sind symmetrisch um null und steigen monoton vom Ursprung. Konventionelle Prognosefehler weisen den einzelnen Abweichungen von Prognose und Beobachtungswert somit eine unterschiedliche Gewichtung und Bedeutung zu, was nachfolgend in der Modellierung expliziter Kostenfunktionen erneut aufzugreifen ist (siehe Abschnitt 5.2).

Die Notation der Prognosefehler kann nach der Unternehmensstruktur und dem jeweiligen Gegenstand der Analyse durch Summierung über m Prognosemethoden, s Zeitreihen unterschiedlicher Produkte oder SKUs ($s = 1, \dots, S$), p Absatzstellen ($p = 1, \dots, P$) für jeden Zeitpunkt

t und Prognosehorizont h erweitert werden [Arminger und Götz 1999, S. 11; Armstrong 2001a, S. 470], mit

$$SE_{t,m} = \sum_{p=1}^P \sum_{s=1}^S \sum_{t=l-h}^{T-h} e_{m,p,s,t+h} \quad (3.97)$$

Aus Gründen der Übersichtlichkeit wird darauf nachfolgend verzichtet.

3.3.2.3 Statistische Fehlermaße

Existiert eine zeitlich geordnete Folge von Prognosefehlern eines Prognoseobjektes oder Fehlerwerte mehrerer Prognoseobjekte, können statistische Fehlermaße abgeleitet werden.

Zeitreihen von Prognosefehlern entstehen durch die Aggregation einfacher Prognosefehler entlang der Zeitreihe $t = 1, \dots, T$. Werden die historischen Beobachtungen in Quasi-ex-ante-Experimenten in mehrere zeitlich zusammenhängende Datenmengen, entstehen entsprechend mehrere Zeitreihen zur separaten Beurteilung der Prognosegenauigkeit (siehe Abschnitt 3.3.3.2). Ferner entstehen Zeitreihen von Prognosefehlern durch Mehrschrittprognosen im Zeitpunkt t durch Aggregation der Prognosefehler über H Prognosehorizonte $t+h$ $h = 1, \dots, H$, für die Zeitpunkte $t+1, t+2, \dots, t+H$ [Schwarze 1980, S. 319].¹ Zudem entstehen Zeitreihen von Prognosefehlern bei einer rollierenden Prognoseevaluierung durch die Aggregation von Fehlerwerten über i Punktprognosen zu mehreren Prognoseursprüngen t , $t = t, t+1, \dots, i$, mit einem identischem Prognosehorizont $t+h$ für die Zeitpunkte $t+h, t+h+1, \dots, t+h+i$. Diese auch in Quasi-ex-ante-Simulationen verwendete Beurteilung über mehrere Prognoseursprünge kann ferner mit Mehrschrittprognosen kombiniert werden, wobei nach FILDES [1992a] zunächst eine Aggregation für einen Horizont h über alle Ursprünge i erfolgt und anschließend eine Aggregation über alle Horizonte.² Auf eine Erweiterung der Notation der Fehlermaße zur Abbildung mehrerer Zeitreihenursprünge und Horizonte wird aus Gründen der Übersichtlichkeit verzichtet.

Die unterschiedlichen Prognosefehler repräsentieren Schätzungen der ersten und zweiten Momente der Dichtefunktion der Prognosefehler in Form ihres Mittelwertes und der Varianz.³ Somit besteht auch die Gelegenheit, für eine Zeitreihe von Prognosefehlern alle bekannten statistischen Lage- und Streuungsparameter, Schiefe- und Wölbungsmaße zu bestimmen. Dennoch wird bis auf wenige Ausnahmen [Brown 1959; Gilchrist 1976] von einer Bestimmung und Analyse der statistischen Verteilung der Zeitreihen aller Prognosefehler abgesehen.⁴ Vielmehr wird vielfach versucht, die Verteilung durch ein einzelnes statistisches Fehlermaß der Varianz der Reihe zu beschreiben [Schwarze 1980, S. 319]. Eine Vernachlässigung der zentralen Lageparameter der Prognosefehler darf keinesfalls erfolgen, da diese eine erhebliche Bedeutung für nachfolgende Modelle der Bestellmengenrechnung besitzen und die einfache Annahme mittlerer Fehler von null nicht genügt.

¹ Die Prognose von n Variablen zu einem Zeitpunkt $t+h$ oder die Aggregation über unterschiedliche Verfahren erfordert hingegen alternative Gütemaße sowie abweichende Schritte der Aggregation.

² Das verfahrensneutrale Problem der Datenaufteilung in eine Kalibrationsstichprobe und Teststichprobe der Zeitreihe zur Vermeidung der Probleme des Overfittings eines Verfahrens an die Daten wird im inhaltlichen Zusammenhang der Verfahren der Neuronalen Netze im Abschnitt 4.3 näher betrachtet.

³ Daraus wird deutlich, dass ein Prognoseverfahren selbst unter Annahme normalverteilter Prognosefehler stets anhand von zwei Maßen der Lage und Streuung der Fehler, z. B. sowohl durch den ME und den MAE, bewertet beschrieben werden sollten.

⁴ Die resultierende Unschärfe wird vor dem Hintergrund der empirischen als auch der wissenschaftlichen Praxis vernachlässigt. Die systematische Analyse der Prognoserestiduen hinsichtlich der Verteilungsform, Stationarität sowie autoregressiver Strukturen ist jedoch in der Bestellmengenrechnung zur Warendisposition erforderlich, um systematische Fehlern in der Spezifikation von Daten, Modell und Verfahren zu identifizieren und in den Bestellmengen zu berücksichtigen (vgl. Abschnitt 2.3).

Um einen einheitlichen Vergleich der Fehlermaße zu den deskriptiven Gütekriterien der Warendisposition (vgl. Abschnitt 2.4.1) sowie ihre Erweiterung zu kostenorientierten Gütekriterien zu ermöglichen (vgl. Abschnitt 52.), werden einschlägig verwendete Fehlermaße der summierten mittleren und relativen Fehler in Tabelle 3.5 systematisch aufgeführt.¹ Nach einer Analyse der Eigenschaften summierter Fehlermaße werden anschließend mittlere und relative Fehlermaße untersucht.

Tabelle 3.5: Statistische Fehlermaße zur Beurteilung der Prognosegenauigkeit

	Einfache Fehlermaße	Absolute Fehlermaße	Quadratische Fehlermaße	Wurzel der quadrat. Fehlermaße
Summierte Fehler	$SE_l = \sum_{t=l-h}^{T-h} e_{t+h}$ sum of errors (3.98)	$SAE_l = \sum_{t=l-h}^{T-h} AE_{t+h}$ sum of absolute errors (3.99)	$SSE_l = \sum_{t=l-h}^{T-h} SE_{t+h}$ sum of squared errors (3.100)	$RSE_l = \sqrt{SSE_l}$ root squared error (3.101)
Mittlere Fehler	$ME_l = \frac{1}{T-l} \sum_{t=l-h}^{T-h} e_{t+h}$ average error (3.102)	$MAE_l = \frac{1}{T-l} \sum_{t=l-h}^{T-h} AE_{t+h}$ mean absolute error (3.103)	$MSE_l = \frac{1}{T-l} \sum_{t=l-h}^{T-h} SE_{t+h}$ mean squared error (3.104)	$RMSE_l = \sqrt{MSE_l}$ root mean squared error (3.105)
Relative Fehler	$MPE_l = \frac{1}{T-l} \sum_{t=l-h}^{T-h} PE_{t+h}$ mean percent error (3.106)	$MAPE_l = \frac{1}{T-l} \sum_{t=l-h}^{T-h} APE$ mean absolute percent error (3.107)	-	-

Einfache Fehlermaße der summierten Fehler (SE), der mittleren Fehler (ME) und der mittleren Fehler in Prozent (MPE) beschreiben das erste Moment der zentralen Lage der Verteilungsfunktion der Prognosefehler. Im Fall einer korrekten Spezifikation von Prognosemodell und -verfahren sind die positiven und negativen Prognoseabweichungen zufällig um null verteilt [Scheer 1983, S. 20]. Da positive und negative Abweichungen in einfachen Fehlermaßen mit dem Erwartungswert $E(E_{t+h}) = 0$ saldiert werden, nehmen SE , ME und MPE kleine Werte um null an und lassen somit keine Aussage über die erwartete Höhe der Prognosefehler im Sinne ihrer Streuung zu. Sie dienen vielmehr bei signifikanten Abweichungen von null bei einer Mehrzahl von positiven oder negativen Fehlern als Indikatoren einer systematischen Verzerrung (engl. *forecast bias*) durch regelmäßige Unter- oder Überprognose, die beispielsweise durch Fehler in der Modellspezifikation und -parametrisierung hervorgerufen werden können [Makridakis et al. 1998, S. 43]. Sie sind jedoch als Maß für die zu erwartende Höhe des Prognosefehlers ungeeignet.

Absolute und quadratische Fehlermaße verhindern eine Summierung positiver und negativer Fehler durch die Bildung des Betrags und beschreiben so die Varianz der Prognosefehler als zweites Moment ihrer Verteilungsfunktion. Quadratische Fehlermaße des MSE und des $RMSE$ dienen ebenfalls der Berechnung nichtnegativer Fehlermaße der Streuung. Im Gegensatz zur Gleichgewichtung aller Prognosefehler des MAE erhalten jedoch die Abweichungen unterschiedliche Gewichte. „The mean squared error ... penalizes a forecast much more for extreme deviations than it does for small ones“ [Wheelwright und Makridakis 1985, S. 23].

¹ Die Tabelle zeigt eine Auswahl statistische Fehlermaße, die zur Beurteilung von Prognoseverfahren in Theorie und Praxis verwendet werden und die Ableitung und Diskussion von kostenorientierten Gütekriterien erlauben. Eine Diskussion alternativer Fehlermaße, z. B. des in frühen Veröffentlichungen verwendeten Korrelationskoeffizienten oder des Bestimmtheitsmaßes R^2 [Meffert und Steffenhagen 1977, S. 199] wird aufgrund der offensichtlichen Probleme in der Verfahrensparametrisierung und Modellselektion verzichtet [Armstrong 2001a, S. 457-458]. Sie erhebt somit keinen Anspruch auf Vollständigkeit. Eine umfassende Übersicht alternativer Fehlermaße findet sich bei SCHWARZE [1980].

Sie implizieren somit eine Gewichtung der Relevanz unterschiedlicher Fehlermaße. Die Wahl eines Verfahrens nach einem quadratischen Fehlermaß impliziert somit die Präferenz mehrerer geringer Abweichungen der Prognosewerte von den tatsächlichen Werten gegenüber einer großen Abweichung [Makridakis et al. 1998, S. 23], z. B. durch die überproportionale Bedeutung der Vermeidung von Ausreißern auf die Prognosegüte.¹ Auf ein mögliches systematisches Über- oder Unterschätzen wird ebenfalls nicht eingegangen. Die Wurzel quadratischer Fehlermaße des *RMSE* erlaubt ihre Ex-post-Interpretation in der ursprünglichen Dimension der Beobachtungswerte, wobei lediglich die Prognosefehler *AE* und *RSE*, nicht aber darauf aufbauende Fehlermaße identisch sind. In der Literatur wird betont, dass größere Abweichungen im *MSE* bzw. *RMSE* ein höheres Gewicht erhalten, was zu einer überproportionalen Bewertung von Ausreißern führt. Es ist jedoch herauszustellen, dass Fehler unter eins, $e < 1$, durch die Quadrierung eine unterproportional geringere Gewichtung erhalten, was insbesondere im Falle einer Datenskalierung auf ein Intervall $[0;1]$ bzw. $[-1;1]$ zur Parametrisierung von Neuronalen Netzen zu einer Unterbewertung der Fehlerfläche führt. Der *MSE* bezeichnet dabei eben die empirische Varianz $\hat{\sigma}_e^2$ und der *RMSE* die empirische Standardabweichung $\hat{\sigma}_e$ der Fehlerverteilung, und wird als solches vielfach zur Beschreibung der Streuung der Prognosefehler verwendet.

Der Mittlere Absolute Prognosefehler (*MAE*), auch als Mittlere Absolute Abweichung (*MAA*) (engl. *mean absolute deviation, MAD*) bezeichnet, „verhindert das Saldieren positiver und negativer Abweichungen im Prognosezeitraum und gewichtet alle Prognosefehler gleich“ [Hansmann 1983, S. 15]. Er ist somit leicht interpretierbar und gilt in der Praxis als leicht erklärbar [Makridakis et al. 1998, S. 43]. Der *MAE* schätzt ebenfalls die Streuung der Prognosefehler um ihren Mittelwert aus den absoluten Abweichungen und kann als Approximation der Standardabweichung $\hat{\sigma}_e$ dienen, da gilt $\sigma_e = 1,25 \text{ MAE}$ [Brown 1956; Silver et al. 1998].

Aufgrund der Verzerrungen quadratischer Fehlermaße – insbesondere im Vergleich alternativer Verfahren – wurden im Rahmen der nach MAKRIDAKIS benannten M-Competitions [Makridakis 1993; Makridakis et al. 1982; Makridakis und Hibon 2000] sowie im Zuge des Verfahrensvergleichs der deutschen Gesellschaft für Operations Research [Schwarze 1980] konkurrierende Fehlermaße zur Modellselektion mit verringerter Sensitivität gegenüber Ausreißern diskutiert und einfache Erweiterungen postuliert. Seither werden verstärkt Fehlermaße angewendet, die auf dem Median als Zentralwert einer Anzahl von n geordneten Prognosefehlern basieren, bei dem die Werte mindestens der Hälfte der Fehler kleiner oder gleich sind, z. B. des *APE* (vgl. Formel 1.87)

$$MdAPE = \begin{cases} APE_{\frac{n+1}{2}} & n \text{ ungerade} \\ \frac{APE_{\frac{n}{2}} + APE_{\frac{n+1}{2}}}{2} & n \text{ gerade} \end{cases} \quad (3.108)$$

Um die Auswirkungen von Extremata auf Fehlermaße zu vermeiden, postuliert ARMSTRONG [2001a, S. 471] die Anwendung ausreißerunempfindlicher Fehlermaße durch gekappte Prognosefehler (engl. *winsorised errors*), die Extremwerten einen konstanten maximalen e_{max} oder minimalen Fehlerwert e_{min} zuweisen, z. B. für den winsorisierten *SE*

¹ Beispielhaft zählt ein *MAE* von 2 doppelt so stark wie eine Prognoseabweichung von 1, während der quadrierte Fehler von 2 als *MSE* viermal so stark zählt wie der einfache Fehler. Im Extremfall bedeutet dies, dass ein Verfahren, welches bis auf wenige Ausreißer exakt prognostiziert, einem Verfahren, das das zugrunde liegende Muster weniger gut approximiert, aber robust gegen Ausreißer ist, unterlegen ist. Im Falle einer Vorverarbeitung der Zeitreihen durch Ex-ante-Eliminierung von statistischen Ausreißern trifft dieses entsprechend nicht zu.

$$WSE = SE_{t+h} \begin{cases} e_{\min} & \text{für } SE_{t+h} < e_{\min} \\ SE_{t+h} & \text{für } e_{\min} \leq SE_{t+h} \leq e_{\max} \\ e_{\max} & \text{für } SE_{t+h} > e_{\max} \end{cases} \quad (3.109)$$

Winsorisierte Prognosefehler können analog zu einfachen Fehlern zu Fehlermaßen aggregiert und anschließend über die Prognosehorizonte h und Zeitreihen s summiert werden [Armstrong 2001a, S. 470-472]. Dabei ist zu beachten, dass die Vernachlässigung oder Kappung einzelner Beobachtungswerte in den Spitzen der Verteilung nicht länger die Interpretation der tatsächlichen Verteilungsform der Prognosefehler erlaubt. Handelt es sich bei den Extremwerten um tatsächlich realisierte Abverkaufsmengen und keine Mess- oder Datenfehler und sind gerade diese Abverkaufsspitzen von Interesse, muss eine derartige Korrektur der Fehlermaße unterbleiben, um eine Interpretation der Prognosequalität nicht systematisch zu verzerren [Armstrong 2001a, S. 457].

Im Vergleich summierter, mittlerer und relativer Fehlermaße sind zusätzliche Aspekte in der Anwendung zu berücksichtigen. Summierte Fehlermaße erlauben keinen validen und reliablen Vergleich über Zeitreihen unterschiedlichen Niveaus, unterschiedlicher Länge oder über unterschiedliche Zeitintervalle. Weisen Zeitreihen Längendiskrepanzen auf, führt die Summierung über eine unterschiedliche Anzahl von Fehlerwerten zu fehlerhaften Vergleichen.¹ Auch zum Vergleich konkurrierender Verfahren sollten summierte Fehlerwerte nicht eingesetzt werden, da konkurrierende Verfahren unterschiedlich große Stützbereiche zur Berechnung von Prognosewerten benötigen und so die Fehlersumme der kürzeren Zeitreihe bei konstanten, absoluten Fehlern pro Periode sinkt.² Zudem erlauben sie keinen Vergleich von Zeitreihen mehrerer Prognoseobjekte mit abweichendem Niveau der Zeitreihenwerte. Zum Vergleich von Fehlermaßen auf Zeitreihen mit einer unterschiedlichen Anzahl von Beobachtungswerten, Prognoseursprüngen oder Prognosehorizonten kommen daher mittlere Fehlermaße zur Anwendung.

Das Ausmaß summierter sowie mittlerer Fehlermaße hängt direkt vom Niveau der Beobachtungswerte ab. Sie erlauben somit keinen Vergleich der Prognosegüte über Produkte unterschiedlicher Wertebereiche oder eines Produktes in unterschiedlichen mengenmäßigen oder zeitlichen Aggregationsstufen, so z. B. von täglichem und monatlichem Absatz [Makridakis et al. 1998, S. 43].³ Um eine Vergleichbarkeit zu gewährleisten, werden relative bzw. prozentuale Fehlermaße des Mittleren Prozentualen Fehlers (*MPE*) oder des Mittleren Absoluten Prozentualen Fehlers (*MAPE*) ermittelt, die sich aufgrund ihrer guten Interpretierbarkeit hoher Popularität erfreuen [Makridakis et al. 1998, S. 45].⁴ Prozentuale Fehlermaße sind in ihrer Bewertung nicht vollständig symmetrisch, da sie im Intervall $[0; \infty]$ definiert sind und somit Fehler durch Unterschätzungen gegenüber solchen durch Überschätzungen ungleich gewichten [Armstrong 1978, S. 348]. Zur Kompensation dieser Asymmetrie wurden

¹ In der Praxis des Bestandsmanagements werden i. d. R. Prognosefehler für die Zeitreihen unterschiedlicher Produkte oder Marken einer Produktgruppe an einer Absatzstelle ermittelt, die aufgrund der variierenden Lebensdauer eines Produktes im Sortiment typischerweise eine abweichende Länge aufweisen.

² Betrachtet man z. B. zwei gleitende Mittelwerte der Länge $i=[2;20]$, reduziert sich die Zeitreihe zur Berechnung der Ex-post-Fehlermaße auf $t-i$, also $t-2$ bzw. $t-20$ Beobachtungswerte, von denen keine Fehlerinformationen aggregiert werden können.

³ Dies gilt deshalb, weil z. B. ein mittlerer *MAE* von 100 Mengeneinheiten für ein Produkt mit einem durchschnittlichen Absatz von 1000 Verkaufseinheiten eine andere Prognosegüte darstellt als für ein Produkt mit einem Absatz von 10 Verkaufseinheiten.

⁴ Der Informationsgehalt eines dimensionslosen *MAPE* oder *RMSPE* von 5 % erscheint deutlich höher als der eines dimensionsbehafteten *MAE* von 10 oder eines *MSE* von 100. Dabei ist zu beachten, dass sie nur bei metrisch skalierten Variablen mit bedeutungsvollem Nullpunkt eingesetzt werden können und die Zeitreihen nur Fehlerwerte $e_t \neq 0$ aufweisen dürfen, da der prozentuale Fehler sonst nicht berechnet werden kann [Makridakis et al. 1998, S. 45].

verschiedene Gütemaße mit unterschiedlichem empirischem Erfolg und unterschiedlichen interpretatorischen Problemen entwickelt, z. B. der „adjusted MAPE“ [Armstrong 1978; Armstrong und Grohman 1972; Schwarze 1980, S. 324]

$$adjMAPE_t = \frac{1}{T-l} \sum_{t=l-h}^{T-h} \frac{AE_t}{0.5(y_t + \hat{y}_t)} \quad , \quad (3.110)$$

der dem „unbiased absolute percent error“ nach MAKRIDAKIS [1993] ähnelt. Bei Zeitreihen von Produkten mit geringem Absatzvolumen und Warenumschlag können jedoch keine reliablen Entscheidungen auf der Basis prozentualer Fehlermaße getroffen werden, da die Probleme der Asymmetrien durch systematische Überprognosen und mangelnden Definition bei Nullmengen überwiegen. Die Auswahl der Fehlermaße erfordert somit die Berücksichtigung der Datenstruktur und der zu unterstützenden Entscheidungssituation.

Bei der Ermittlung der Fehlermaße findet die Zeit keine direkte Berücksichtigung. Summierte, mittlere und prozentuale Fehlermaße weisen allen Fehlerwerten der Zeitreihe ein identisches Gewicht zu, sodass weiter zurückliegende Fehler eine identische Gewichtung erhalten wie Fehler der vorigen Periode. Eine Betrachtung dynamisierter bzw. geglätteter Fehlermaße zur Abbildung heteroskedastischer Zeitreihen findet sich bei [Schwarze 1980, S. 338; Scheer 1983, S. 23-24; Lewandowski 1974, S. 194ff.]. Alle dargestellten Fehlermaße gewichten positive und negative Abweichungen durch Unter- oder Überschätzungen der Prognosewerte mit gleichem Betrag identisch, weshalb sie auch als symmetrische Fehlermaße bezeichnet werden.

Es bleibt festzuhalten, dass jedes der statistischen Fehlermaße unterschiedliche Verzerrungen der Interpretation bedingt und somit weitgehende Unstimmigkeit in der wissenschaftlichen Diskussion bezüglich der Validität und Reliabilität einzelner Fehlermaße herrscht. In der Theorie wird vielmehr die Beurteilung anhand einer Auswahl mehrerer Fehlermaße zur Verfahrensauswahl und Entscheidungsbildung postuliert, was aber ein Entscheidungsproblem der Gewichtung zur Aggregation der einzelnen Maße bedingt und das originäre Problem lediglich verlagert [Armstrong 2001a, S. 460].¹ Nachfolgend wird daher eine Auswahl empirisch und theoretisch etablierter Fehlermaße im Verfahrensvergleich eingesetzt.

3.3.2.4 Komparative statistische Fehlermaße

Als Ergänzung zu statistischen Fehlermaßen werden nachfolgend komparative Maße erläutert. Keines der statistischen Fehlermaße erlaubt die Ableitung einer generellen Aussage bezüglich der Güte eines Verfahrens oder gar der Optimalität der Prognose. Nach HANSMANN „fehlt .. ein kritischer Wert, der zwischen guten und schlechten Prognosen diskriminiert“ [Hansmann 1983, S. 16]. Aus diesem Grund wurden komparative Fehlermaße entwickelt, bei denen die Performanz eines Verfahrens mit der Genauigkeit eines Benchmarks verglichen wird [Makridakis et al. 1998, S. 46].²

Als einfaches komparatives Fehlermaß kann die Differenz zwischen dem *MAE* oder *MAPE* des zu beurteilenden Verfahren und einem Naiven Verfahren gemäß der Zeitreihenstruktur verwendet werden. „The difference ... provides a measure of the improvement attainable through use of that more sophisticated forecasting method. ... [and] ... a basis for evaluating the relative accuracy of those results“ [Makridakis et al. 1998, S. 47]. Die *U-statistic*, im Deutschen als THEIL'scher Ungleichheitskoeffizient (U) bezeichnet [Theil 1966, S. 18], stellt

¹ Die Kombination mehrerer Fehlermaße stellt zudem die Frage der Diskrepanz zwischen der Beurteilung eines Verfahrens auf einer Auswahl von Fehlermaßen und seiner Parameterfindung durch Minimierung eines einzelnen, meist quadratischen Fehlers.

² Sie werden auch als „vergleichende“ oder „normierte Fehlermaße“ bezeichnet.

ein normiertes Fehlermaß dar, das nach einem kritischen Wert U zwischen guten Prognosen mit $U < 1$ und schlechten Prognosen mit $U > 1$ diskriminiert:

$$U = \sqrt{\frac{\sum_{t=1}^{n-1} \left(\frac{\hat{y}_{t+1} - y_t}{y_t} - \frac{y_{t+1} - y_t}{y_t} \right)^2}{\sum_{t=1}^{n-1} \left(\frac{y_{t+1} - y_t}{y_t} \right)^2}} = \sqrt{\frac{\sum_{t=1}^{n-1} \left(\frac{\hat{y}_{t+1} - y_{t+1}}{y_t} \right)^2}{\sum_{t=1}^{n-1} \left(\frac{y_{t+1} - y_t}{y_t} \right)^2}}, \quad (3.111)$$

bzw. nach alternativer Definition [Hansmann 1983, S. 16] mit

$$U_2 = \sqrt{\frac{\sum_{t=1}^{t+h} (y_{t+h} - \hat{y}_{t+h})^2}{\sum_{t=1}^{t+h} (y_{t+h} - y_{t+h-1})^2}}. \quad (3.112)$$

Im Fall einer fehlerfreien Prognose gilt entsprechend $U = 0$, bei Prognosefehlern die mit den Fehlern der Naiven Prognose identisch sind $U = 1$. Eine Diskussion der unterschiedlichen Berechnungsvarianten des U findet sich in [Armstrong 1978, S. 349; sowie Schwarze 1980, S. 333-335]. Der Vergleich der Prognosegüte eines Verfahrens mit der Naiven Prognose NF1 erscheint empirisch relevant, da die Naive Prognose häufig bereits vergleichbar gute Prognoseergebnisse liefert [Armstrong 1984; Schnaars 1986].

Der Ungleichheitskoeffizient beruht auf der Berechnung quadratischer Fehlermaße mit den vorangehend beschriebenen Effekten der Übergewichtung großer Fehler. ARMSTRONG [2001a] postuliert entsprechend den Prognosefehler des Relativen Absoluten Fehlers (engl. *relative absolute error, RAE*), mit

$$RAE = \frac{|y_{t+h} - \hat{y}_{t+h}|}{|y_{t+h} - y_{t+h-1}|} \quad (3.113)$$

als Grundlage zur Berechnung weiterführender Fehlermaße, insbesondere des geometrischen Mittels des RAE über unterschiedliche Zeitreihen s , $s = 1, \dots, S$ [Armstrong 2001a, S. 471], mit

$$GMRAE_{m,h} = \left[\prod_{s=1}^S RAE_{m,h,s} \right]^{\frac{1}{S}}. \quad (3.114)$$

Weiterhin findet das einfache komparative Fehlermaß der prozentualen Verbesserung (engl. *percent better, PB*), weite Verbreitung zur Beurteilung von Verfahren über mehrere Zeitreihen s hinweg:

$$PB_{m,h} = \frac{\sum_{s=1}^S j_s}{S} \quad \text{mit} \quad j_s = \begin{cases} 1 & \text{wenn } |y_{t+h} - \hat{y}_{t+h}| < |y_{t+h} - y_{t+h-1}| \\ 0 & \text{sonst} \end{cases}. \quad (3.115)$$

Die Anwendung komparativer Fehlermaße repräsentiert bereits einen Verfahrensvergleich hinsichtlich der Prognosequalität und kann entsprechend auf Verfahren beliebiger Komplexität ausgedehnt werden. Die von SCHWARZE [1980] geforderte Nutzung komparativer

Fehlermaße zur Beurteilung der Prognosegüte eines einzelnen Verfahrens entspricht somit der Beurteilung durch einen Verfahrensvergleich, was für die Überprüfung etablierter sowie neuartiger Prognoseverfahren in der Erprobung und laufenden Anwendung relevant erscheint. Ein Verfahrensvergleich zwischen mehreren konkurrierenden Prognoseverfahren anhand von Ungleichheitskoeffizienten erscheint jedoch problematisch, da Prognosefehler zwischen $0 \leq U \leq +\infty$ nichtlinear abgebildet werden und somit keine valide Rangreihung der Verfahren zulassen [Hansmann 1983, S. 16]. Ferner repräsentieren komparative Fehlermaße nicht die eigentlichen Entscheidungskosten und erlauben keine Nutzung zur Parametrisierung der Verfahren im Sinne einer Minimierung von U zur Findung optimaler Parameter der Prognosefunktion.

Unter Betrachtung alternativer Ansätze statistischer Fehlermaße der Prognose bleibt somit das Problem der Auswahl eines adäquaten statistischen Fehlermaßes oder der Kombination mehrerer Fehlermaße zur validen und reliablen Modellselektion bestehen. Es wird deutlich, dass die Qualität einer Prognose letztlich im inhaltlichen Zusammenhang der betrieblichen Entscheidungssituation zu beurteilen ist. Wie in Abschnitt 2.4 dargestellt, erfolgt die Beurteilung der Dispositionsgüte anhand deskriptiver Fehlermaße der Fehlmengen und der Überbestandsmengen sowie des Auftretens von Fehlmengensituationen und Überbestandssituationen. Diese Fehlermaße sind hoch asymmetrisch, indem sie lediglich einseitige Dispositionsfehler beurteilen, und erfordern die Überprüfung der Zielerreichung anhand von zwei komplementären Zielgrößen des Servicegrades und der Bestandshöhe. Dies steht im Gegensatz zu den symmetrischen statistischen Fehlermaßen der Prognosefehler, die symmetrisch Über- und Unterprognosen identisch bewerten und keine unterschiedliche Gewichtung implizieren. Die Diskrepanz der deskriptiven Fehlermaße der Disposition und der statistischen Fehlermaße für das Teilproblem der Prognose verhindert eine kohärente Parametrisierung und Auswahl von Verfahren im Sinne der originären Problemstellung und Zielsetzung der Warendisposition. Diese Diskrepanz ist nachfolgend durch die Bewertung der Prognosefehler und der Dispositionsfehler mit den betrieblichen Kosten und ihrer Abbildung in einer integrierten, asymmetrischen Kostenfunktion zu überwinden (vgl. Kapitel 5).

3.3.3 Durchführung von empirischen Vergleichsstudien

3.3.3.1 Erkenntnisse aus Vergleichsstudien

Grundsätzlich kann keinem etablierten Prognoseverfahren oder -modell eine universelle Überlegenheit für alle Anwendungsbereiche, Entscheidungssituationen und Datenstrukturen unterstellt werden, sodass eine situative Beurteilung für das jeweilige Entscheidungsproblem erfolgen muss. Zudem erfordern neuartige Prognoseverfahren die systematische Beurteilung ihrer Prognosegenauigkeit im Vergleich zu etablierten Verfahren. Aus diesen Gründen sind systematische Vergleichsstudien erforderlich, um valide und reliable Erkenntnisse über die relative Güte der einzelnen Verfahren für die spezifische Entscheidungssituation zu gewinnen. Die Testsituation ist dabei weitestgehend an die tatsächliche Entscheidungssituation anzunähern [Armstrong 2001e, S. 710]. Dabei ist auf besondere Gestaltungskriterien der Verfahrensvergleiche zu achten, die nachfolgend dargestellt werden und in der Ableitung einer Methodologie zur Anwendung von Prognosen in Abschnitt 3.4 sowie bei der Beurteilung der Neuronalen Netze in Kapitel 6 zu berücksichtigen sind.

Der Vergleich konkurrierender Prognoseverfahren auf synthetischen oder empirischen Zeitreihen durch einen sogenannten Prognosewettbewerb (engl. *forecasting competition*) besitzt eine lange Tradition in den Wirtschaftswissenschaften. Einen Überblick über die Historie empirischer Verfahrensvergleiche findet sich in [Makridakis und Hibon 2000; Fildes und Makridakis 1995]. Aufbauend auf frühen Wettbewerben zum Vergleich der Prognosegüte konkurrierender Verfahren der Zeitreihenanalyse [Reid 1969; Reid 1972; Newbold und

Granger 1974; Makridakis und Hibon 1979] wurden von MAKRIDAKIS et al. mehrere Prognosewettbewerbe der M-Competition mit 1001 Zeitreihen [Makridakis et al. 1982] und der M3-Competition mit 3003 Zeitreihen organisiert [Makridakis und Hibon 2000].¹ Aus diesen Vergleichsstudien wurden vielfältige Erkenntnisse zur relativen Prognosequalität von objektiven Prognoseverfahren gewonnen, die hier nach MAKRIDAKIS und HIBON [2000] zusammengefasst werden:

- Verfahren mit höherer Komplexität sind nicht genauer als einfache Verfahren.
- Unterschiedliche Fehlermaße führen zu unterschiedlichen Rangordnungen der Verfahren.
- Kombinierte Prognosen sind genauer als die individuellen Prognosen der kombinierten Verfahren, und von vergleichbarer Genauigkeit gegenüber dem Rest der Methoden.
- Die Leistung eines Verfahrens hängt von der Ausdehnung des Prognosehorizontes ab.

Diese Schlussfolgerungen wurden mehrfach durch umfassende Vergleichsstudien verifiziert [Fildes 1992a] und durch erweiterte Untersuchungen wie die M3-Competition und durch nachfolgende Studien von [Meese und Geweke 1984; Fildes et al. 1998] bestätigt. Nach FILDDES [1998] können die Erkenntnisse wie folgt erweitert werden:

- Charakteristika der Zeitreihen beeinflussen die relative Leistung von Prognoseverfahren
- Verfahrensvergleiche sollten aus mehreren Zeitreihen-Ursprüngen durchgeführt werden, um Fehler durch die Stichprobenziehung der Datenmengen zu verhindern.

Über diese Erkenntnisse hinaus wurden durch die intensive wissenschaftliche Diskussion der Versuchsaufbauten einzelner Studien wertvolle Erkenntnisse für die generelle Durchführung von Vergleichsstudien gewonnen. Diese Erkenntnisse im Sinne einer „research methodology“ [Ord et al. 2000b] spezifizieren ein systematisches Vorgehen zur validen und reliablen Beurteilung bestehender und neuartiger Prognoseverfahren und gelten nachfolgend als Maßstab einer wissenschaftlichen Beurteilung von Prognoseverfahren.

3.3.3.2 Durchführung von Vergleichsstudien

TASHMAN [2000] fasst die Erkenntnisse zum Design empirischer Vergleichsstudien zusammen. Demnach sind vor allem die Beurteilung von Verfahren auf einer Test-Stichprobe, die Verwendung robuster, unverzerrter Fehlermaße als Gütekriterien, die rollierende Beurteilung auf mehreren Prognoseursprüngen und der Vergleich zu etablierten Benchmark-Verfahren zu berücksichtigen.

Die Beurteilung der Prognosegüte auf ungesehenen, historischen Beobachtungswerten (engl. *out-of-sample*) bestimmt maßgeblich die Validität aller Verfahrensvergleiche: „The performance of a model on data outside that used in its construction remains the touchstone for its utility in all applications“ [Fildes und Makridakis 1995]. Die Güte der Anpassung (engl. *goodness of fit*) des Verfahrens an die historischen Datenwerte lässt keine verlässlichen Rückschlüsse auf die Eignung eines Verfahrens zur Prognose zukünftiger Werte im Sinne einer Generalisierung zu. So können bei Vorliegen von n Beobachtungswerten durch Parametrisierung einer Polynomfunktion $n-1$ Grades alle Punkte exakt mit einem Fehler $e_t = 0$ interpoliert werden. Dieses erlaubt jedoch keine Prognose in die Zukunft, was als „over-fitting a model to a data series“ [Makridakis et al. 1998, S. 46] bezeichnet wird.² Daher zeigen Gütemaße auf den Daten auf denen das Modell konstruiert und das Verfahren parametrisiert wird (engl. *within-sample*) geringere Prognosefehler und Konfidenzintervalle an, als tatsächlich erreichbar sind [Chatfield 1993]. Ferner weisen Verfahren mit den besten „Within-

¹ Die M2-Competition stellte hingegen einen Wettbewerb zur Evaluation subjektiver Prognoseverfahren durch Expertenteams dar und kann demzufolge hier vernachlässigt werden [Makridakis et al. 1993].

² Auch wenn diese Eigenschaft für datengetriebene, semi-parametrische Verfahren wie die Neuronale Netze als Problem intensiv diskutiert wird, besteht das Problem des „overfittings“ ebenso für etablierte statistische Verfahren der Regression (zu Problemen der Parametrisierung von Neuronale Netzen siehe Abschnitt 4.3).

sample“-Fehlern häufig eine vergleichsweise schlechtere Leistung auf ungesehenen Daten auf [Bartolomei und Sweet 1989; Pant und Starbuck 1990]. Das Ziel der Evaluation ist jedoch die Beurteilung geeigneter Prognoseverfahren hinsichtlich ihrer Prognosegüte auf vorher unbekanntem Daten der Zukunft im Sinne der Generalisierbarkeit. Die Bestimmung der Güte konkurrierender Prognoseverfahren erfordert somit die Beurteilung auf unbekanntem Daten und die Aufteilung der Beobachtungswerte in disjunkte Datenmengen.

Die Quasi-ex-ante-Beurteilung¹ von Verfahren teilt die verfügbare Menge von t historischen Beobachtungen vom Zeitpunkt 1, 2, ..., t in zwei disjunkte Datenbereiche:² eine „in-sample“-Kalibrationsstichprobe³ von 1, 2, ..., $t-n-1$ der Vergangenheitsdaten, und eine „out-of-sample“-Teststichprobe der Länge n von $t-n, t-n+1, \dots, t$ im Sinne von quasi-Zukunftsdaten. „The initialization set is used to estimate any parameters and to initialize the method. Forecasts are made for the test set“ [Makridakis et al. 1998, S. 46]. Die Modelle und Verfahren werden somit allein unter Verwendung der Kalibrationsstichprobe gebildet und anschließend zur Simulation einer Ex-ante-Beurteilung auf den vorher zurückgehaltenen Beobachtungen der Teststichprobe hinsichtlich ihrer Prognosegenauigkeit beurteilt.⁴ Da die Daten der Testmenge dem Prognoseverfahren in der Parameterfindung vorenthalten werden, wird diese auch als „Hold-out“-Menge bezeichnet. In Abbildung 3.22 wird die Aufteilung der historischen Daten in die disjunkten Datenmengen visualisiert.

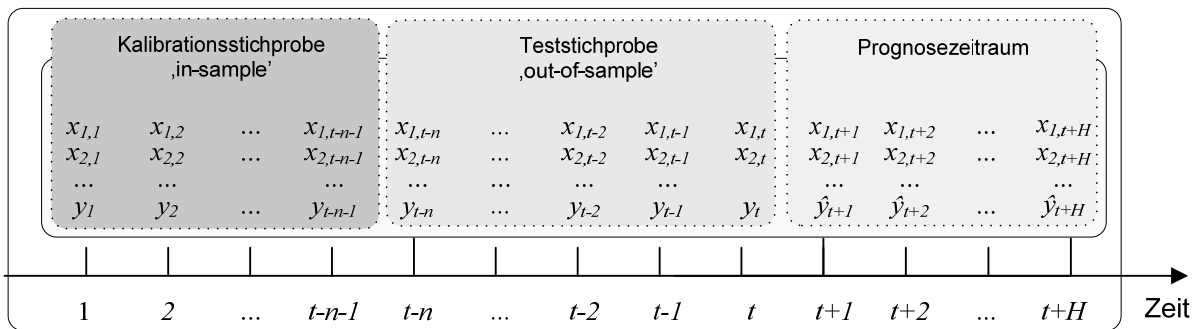


Abbildung 3.22: Kalibrations- und Teststichproben zur Beurteilung von Prognoseverfahren

Die Güte des Prognoseverfahrens auf der Kalibrationsstichprobe schätzt somit die Fähigkeit des Verfahrens, die zugrunde liegende Modellform aus den Daten im Sinne einer möglichst genauen Anpassung an die Beobachtungswerte zu approximieren. Die Prognosegüte auf der Teststichprobe zeigt hingegen seine Fähigkeit, die Modellform auf bisher nicht verwendeten, unbekanntem Daten in die Zukunft zu extrapolieren und somit im eigentlichen Sinne der Prognose zu generalisieren. Wurde der Testdatensatz nicht in der Phase der Modellparametrisierung benutzt, gelten diese Prognosen nach MAKRIDAKIS et al. [1998, S. 46] als „genuine forecasts“.⁵ KÜSTERS [2006] bezeichnet diese Vorgehensweise als „Prognosesimulation mit Kalibrations- und Teststichprobe“ und unterscheidet sie von einer reinen Ex-post-Evaluation

¹ Zur Abgrenzung von dem etablierten Begriffs einer tatsächlichen „Ex-ante-Beurteilung“ von Prognosen (vgl. Abschnitt 3.3.1) wird sie nachfolgend als „Quasi-ex-ante-Beurteilung“ bezeichnet.

² Der Einfachheit halber werden hier sequentielle Datenmengen unterstellt. Gleiches gilt jedoch auch für eine zufällige Stichprobenziehung ohne oder mit Zurücklegen.

³ Sie wird auch als auch als Parametrisierungsmenge, Stützbereich oder „initialisation set“ bezeichnet.

⁴ Die Bedeutung der Quasi-ex-ante-Beurteilung eines Prognoseverfahrens im Gegensatz zur Beurteilung seiner Anpassungsfähigkeit an historische Werte erscheint insbesondere für datengetriebene Verfahren der Neuronalen Netze relevant, die in der Lage sind, durch das Training jede beliebige, stetige Funktion zu einem beliebigen Grad in-sample zu approximieren. Daher wird die Quasi-ex-ante-Beurteilung auf zwei oder mehreren disjunkten, historischen Datensätzen zur Beurteilung der Prognosegüte und Generalisierungsfähigkeit von Neuronalen Netzen in Abschnitt 4.3 erneut aufgegriffen.

⁵ Auf die Möglichkeit einer Erhöhung der Messvalidität der Verfahrensvergleiche durch eine mehrfache Berechnung der Prognoseverfahren an unterschiedlichen Zeitpunkten t_1, t_2, \dots als Prognoseursprung kann hier aus Platzgründen nur hingewiesen werden.

und der „Track-Record-Analyse“ durch eine retrospektive ex post Evaluation eines ex ante bestimmten und empirisch eingesetzten Prognoseverfahrens [Küsters 2006, S. 372].

Durch die Aufteilung der historischen Beobachtungswerte in die Teilmengen ergeben sich Probleme bei der Ziehung repräsentativer Stichproben zur Sicherstellung der Repräsentativität und somit Generalisierungsfähigkeit. Die Höhe der Fehlermaße eines Verfahrens wird stark durch den letzten beobachteten Zeitreihenwert y_t als Ursprung der Prognose über den Prognosehorizont beeinflusst. Durch Zufallseinflüsse können die Prognosen aus einem Prognoseursprung t über den gesamten Prognosehorizont $t+1, t+2, \dots, t+H$ verzerrt werden. Eine Berechnung durchschnittlicher Fehlermaße über mehrere Ursprünge t vermeidet die starken Zufallseinflüsse einzelner Prognoseursprünge [Schwarze 1980, S. 320]. Die Beurteilung von Prognoseverfahren sollte daher mehrere Prognoseursprünge t umfassen, um durch rollierende Zeitreihenursprünge (engl. *rolling origins*) die Reliabilität und somit Generalisierbarkeit der Ergebnisse zu erhöhen [Fildes et al. 1998]. Entsprechend wurden neben der einfachen Aufteilung in zwei Teilmengen durch Kreuzvalidierung (engl. *cross-validation*) auf einer einzigen Testmenge auch Ansätze zur k -fachen Kreuzvalidierung mit mehreren Testmengen von unterschiedlichen Zeitreihenursprüngen entwickelt, die eine bessere Einschätzung der Validität und Reliabilität der Ergebnisse erlauben. Die Beurteilung der Prognosegüte für eine k -fache Kreuzvalidierung über k Datenmengen wird als einfacher Mittelwert über die Zeitreihenursprünge ermittelt. Die Prognosegüte wird über den Prognosehorizont von $t+h$ beurteilt, indem je Verfahren h iterative Einschnitt-Punktprognosen oder eine einmalige Mehrschritt-Punktprognose mit Horizont h durchgeführt wird.

Es ist jedoch anzumerken, dass auch Vergleichstudien mit nur einem Prognoseursprung, wie z. B. die M1- und M3-Wettbewerbe, durch die Überprüfung auf einer hohen Anzahl homogener Zeitreihen bereits valide und reliable Aussagen treffen können. Dies dient insbesondere bei Prognoseverfahren mit einem hohen Berechnungsaufwand, z. B. den Neuronalen Netzen, zu einer Vereinfachung des Experimentaufbaus.¹

In der Beurteilung der Prognosegüte sollte eine Auswahl unverzerrter, meist prozentualer Fehlermaße verwendet werden, die robust gegen vereinzelte hohe Fehler z. B. durch Ausreißer sind. Insbesondere sind quadratische Fehlermaße des MSE und $RMSE$ sowie das Bestimmtheitsmaß R^2 für eine Beurteilung ungeeignet [Armstrong 2001e, S. 715ff.]. Da Verfahrensvergleiche in der wissenschaftlichen Literatur unabhängig von einem zugrunde liegenden Entscheidungsproblem berechnet werden, kann keine Bewertung der Prognosen anhand der resultierenden Fehlerkosten erfolgen. Sind diese im konkreten Anwendungsfall jedoch bekannt, erlauben sie die unverzerrte, realistische Beurteilung der Prognosegüte. Diese Gestaltungsrichtlinien sind auch außerhalb von Prognosewettbewerben gültig, u. a. in der empirischen Bewertung neuartiger Prognoseverfahren wie der NN im Vergleich zu etablierten Verfahren. Sie werden in den empirischen Experimenten im Abschnitt 6 entsprechend berücksichtigt.

3.4 Anwendung von Prognoseverfahren

3.4.1 Methodologien der Prognoseprozesse

Die Anwendung von Prognoseverfahren in Theorie und Praxis erfordert eine systematische Erarbeitung von Prognosen² von der Festlegung des Prognosegegenstands über die Identifika-

¹ In Abhängigkeit des Verfahrens sind ferner unterschiedliche zeitliche Ordnungen der Stichprobenziehung möglich. So erfordern Verfahren der Exponentiellen Glättung und ARIMA sequentielle Stichproben wie in Abbildung 3.22. Verfahren der Neuronale Netze erlauben darüber hinaus auch Stichproben in zufälliger zeitlicher Reihenfolge durch Aufteilung in disjunkte Datenvektoren. Nachfolgend wird vereinfachend von sequenziellen Stichproben ausgegangen.

² Dieses wird vereinzelt auch als „Prognostik“ bezeichnet, siehe u.a. WEBER [1991, S. 7].

tion zulässiger und effizienter Modellformen und Verfahren bis zur Berechnung und Beurteilung der Prognoseergebnisse.

Somit kommt der Formulierung einer Methodologie als systematische Vorgehensweise und konzeptioneller Rahmen zur Prognose eine besondere Bedeutung zu.¹ Korrespondierend mit der Vielfalt betrieblicher Entscheidungsprozesse und Prognoseverfahren existiert eine Vielzahl prototypischer Phasenmodelle zur Modellierung von Prognoseverfahren [Gilchrist 1976, S. 12]. Ein allgemeingültiges Vorgehensmodell kann vor dem Hintergrund der Vielzahl möglicher Datenkonstellationen, Prognosehorizonte, Prognosemodelle und -verfahren nicht existieren. Nachfolgend werden unterschiedliche Methodologien der Prognose mit objektiven Verfahren diskutiert und ein verfahrensübergreifendes Prozess- und Referenzmodell der Prognose in Anlehnung an die Erkenntnisse empirischer Verfahrensvergleiche hergeleitet.²

Als Konsens aller Methodologien werden die drei Phasen der Prognosevorbereitung, der Prognoseerstellung und der Prognoseanwendung unterschieden, die iterativ durchlaufen werden. Vielfach beschränken sich die Vorgehensweisen auf die verfahrenstechnischen Anforderungen der Analyse, Selektion und Anwendung in der Prognoseerstellung einzelner Methoden. MAKRIDAKIS et al. [1998] betrachten ausschließlich Aspekte der Erstellung von Prognosen: „The application of a general class of forecasting methods involves two basic tasks: analysis of the data series and selection of the forecasting model (i. e. the specific methods within that class) that best fits the data series“ [Makridakis et al. 1998, S. 312]. Somit geht der Prognose immer die Analyse der Zeitreihen auf systematische Muster voraus. Anschließend erfolgt die Selektion eines Prognoseverfahrens, das die identifizierten Muster abbilden kann [Makridakis et al. 1998, S. 312]. Das populäre Vorgehensschema der BOX-JENKINS-Methodologie [Box und Jenkins 1970; Box et al. 1994] umfasst die Schritte Identifikation, Parameterermittlung, Test und Anwendung. Es berücksichtigt jedoch ausschließlich die Verfahrensklasse der ARIMA-Modelle und vernachlässigt den Vergleich alternativer Datensätze, Erklärungsmodelle und Prognoseverfahren. WEBER [1991, S. 6] zerlegt den Prognoseprozess in die drei Stufen Vorbereitungs-, Ausführungs- und Beurteilungsphase. HANSMANN [1983, S. 13] schlägt dagegen ein achtstufiges Ablaufschema vor, das an ein „pragmatisches“ Vorgehensmodell nach BROCKHOFF [1977, S. 49] in 13 Stufen mit rekursiven Schleifen und Sprüngen angelehnt ist. Beide spezifizieren die Durchführung von Prognosen von ihrer erstmaligen Entwicklung bis hin zur routinemäßigen, automatisierten Anwendung und unterscheiden zwischen der Entwicklung von Erklärungs- und Prognosemodellen und den darauf angewandten Prognoseverfahren. HRUSCHKA [1980, S. 2] bildet dagegen das Verhalten eines rational handelnden betrieblichen Prognostikers in mehreren Phasen nach. GILCHRIST [1976, S. 12] strukturiert das Ablaufschema des „scientific forecasting“ ebenfalls nach den Phasen eines Prognostikers, vernachlässigt jedoch die Schritte der Problemdefinition sowie der Modellevaluation.

In einem integrativen Ansatz zur Kombination der methodischen und expertenorientierten Betrachtungsweisen charakterisiert ARMSTRONG [2001c] den Prozess zur Erstellung von Prognosen durch sechs Hauptaufgaben eines Prognostikers. Zusätzlich ordnet er in den Prozess 139 Prinzipien, Ratschläge und methodische Verfahrensanweisungen zur validen und reliablen Prognose ein [Armstrong 2001e]. Die Phasen des Prognoseprozesses sind in Abbildung 3.23 in Anlehnung an die prozessorientierte Darstellung der Prozesse zur Warendisposition in Abschnitt 2.1 dargestellt. Die Prinzipien, die als Konsens wissenschaftlicher

¹ Dies wird auch durch die Notwendigkeit der Automatisierung des Prognoseprozesses in der Warendisposition vieler Produkte durch softwaregestützten Planungssystemen

² Dies ist auch erforderlich, da für Verfahren der Neuronalen Netze eine etablierte Methodologie zur Modellierung in Anwendungen der Punktprognose fehlt. Diese ist in einem nachfolgenden Abschnitt zu entwickeln und in die hier dargestellten Phasenkonzepte zu integrieren.

Experten der Prognostik erarbeitet wurden, werden in der Diskussion der einzelnen Teilprozesse oder Verfahren aufgegriffen.¹

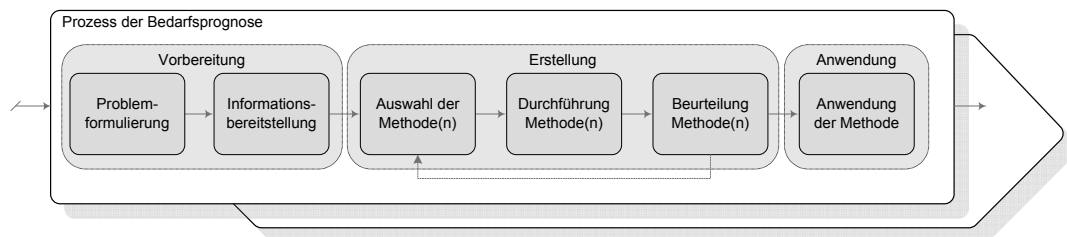


Abbildung 3.23: Phasen des Prognoseprozesses²

Der Prognoseprozess sollte die Anforderungen aller, auch neuartiger Prognoseverfahren berücksichtigen. Insbesondere sollte die Modellierung moderner nicht-parametrischer und semi-parametrischer Verfahren wie der Neuronalen Netze in das Phasenschema integrierbar sein. Dies erfordert eine Berücksichtigung der Interdependenz von Prognosemodell, Prognoseverfahren und der Datenvorverarbeitung, die für statistische Prognoseverfahren weitgehend vernachlässigt wird [Hansmann 1983; Hüttner 1986; Mertens und Deistler 1994; Weber 1991]. Unterschiedliche Prognoseverfahren stellen in der Datenvorverarbeitung abweichende Anforderungen hinsichtlich der Datenkodierung, Ausreißerbehandlung, Saison- oder Trendbereinigung, die je nach der situativen Beschaffenheit der Daten für einzelne Verfahren optional oder obligatorisch sein. JOERENSEN [1996; , zitiert nach Thiesing 1998, S. 80-83] unterscheidet dementsprechend zwischen modellbezogenen, fehlerbezogenen und sachbezogenen Methoden der Vorverarbeitung, die hier den unterschiedlichen Phasen des Prognoseprozesses zugeordnet werden. Der sechsstufige Prognoseprozess nach ARMSTRONG wird dementsprechend um die Berücksichtigung verfahrensspezifischer Komponenten im Prognoseprozess erweitert.

Die Untersuchung und Erweiterung der einzelnen Phasen in den folgenden Abschnitten wird auf die Berücksichtigung objektiver Prognoseverfahren zur Bedarfsprognose beschränkt.³ Aus ihnen werden Prozess- und Funktionsdiagramme abgeleitet, die einen prozessorientierten Vergleich von herkömmlichen statistischen Verfahren mit neuartigen Prognoseverfahren der Neuronalen Netzen zulassen sowie ihre Einordnung in die Prozesse der Bedarfs- und Bestellmengenrechnung der Warendisposition

3.4.2 Phasen des Prognoseprozesses

3.4.2.1 Problemformulierung

Die einleitende Phase der Problemformulierung (siehe Abbildung 3.23) verbindet den Prognoseprozess mit dem betrieblichen Entscheidungsprozess. Sie besteht aus der allgemeinen Zieldefinition und der Strukturierung und Festlegung der Prognoseparameter.

Die Spezifikation der Zielsetzung gestaltet sich für operative Anwendungen durch den direkten Bezug zur betrieblichen Entscheidung meist unproblematisch. So erfordert z. B. die Warendisposition eine möglichst genaue und unverzerrte Bedarfsprognose durch möglichst geringe Prognosefehler und ohne systematische Über- oder Unterschätzungen für die valide Berechnung der Bestellmengen. Dagegen gilt die Festlegung der Parameter des Prognosepro-

¹ Die Prinzipien in ARMSTRONG [2001e] reflektieren die gegenwärtigen, wissenschaftlichen Erkenntnisse zur Prognostik. Allerdings sind nicht alle Prinzipien für die Anwendung quantitativer Verfahren relevant, sodass nachfolgend nur eine Auswahl der Prinzipien zur Prognose mit zeitreihenanalytischen und kausalen Verfahren in der Entwicklung einer robusten Methodologie berücksichtigt wird.

² Eigene Darstellung in Anlehnung an die „stages of forecasting“ nach ARMSTRONG [2001c]

³ Dies erfolgt unter Vernachlässigung der Interaktion der Prognose mit subjektiven Expertenurteilen durch nachträgliche Korrekturen zur Einbeziehung zusätzlicher Informationen.

zesses selbst für wohlstrukturierte Entscheidungsprobleme als schwierig [Armstrong 2001e, S. 680; Makridakis et al. 1998, S. 13]. Zunächst sind die Freiheitsgrade des Entscheidungs- und Prognoseproblems zu erheben und in kontrollierbare, endogene Größen und nicht kontrollierbare, exogene Größen zu unterteilen. Anschließend erfolgt die Gestaltung der Entscheidung und der Dekomposition des Problems in Teilprobleme zur Untersuchung möglicher Einflussfaktoren und ihrer Interdependenzen [Armstrong 2001e].

Entscheidungen über den minimalen und maximalen Prognosehorizont $t+h$ und über die Prognosezeiträume in Abhängigkeit von der Vorlaufzeit des Entscheidungsproblems werden aus den Planungshorizonten der Entscheidungsfindung abgeleitet. Im einfachen Fall einer $t+1$ erfolgt die Vorhersage des Bedarfs der jeweils nachfolgenden Periode $t+h$ mit $h = 1$. Dagegen erfordern multiple Prognosehorizonte von $h = 1, 2, \dots, H$ die Berechnung der Prognosen und ihre Auswertung über unterschiedliche Prognoseursprünge, -horizonte und -zeiträume. Darüber hinaus können zeitliche Vorlauffristen von t bis zum ersten Prognosehorizont $t+h$ existieren, da in dieser Zeit die Erkenntnisse der Prognose keine betrieblichen Entscheidungen beeinflussen können, z. B. durch fixierte Ablaufpläne der Produktionsplanung oder fester Lieferzeiten zu disponierender Ware. Sie sollten somit nicht in der Beurteilung der Prognose berücksichtigt werden.¹ Darüber hinaus können je Prognoseobjekt unterschiedliche, überlappende Prognosehorizonte für unterschiedliche operative, taktische oder strategische Entscheidungen erforderlich sein. Die Bestimmung erfordert so die Erhebung der Interaktion der Prognosen mit allen nachfolgenden Entscheidungen entlang der betrieblichen Wertschöpfungskette.

Ferner sind Entscheidungen zur Länge der Prognosezyklen in Anlehnung an die mögliche Entscheidungsintervalle und -rhythmen wie der Bestellrhythmen [Scheer 1983]. Auch sind zulässige Aggregationsgrade der zu prognostizierenden Größen in den Hierarchien der Sortimente, Artikel oder Standorte aus dem Entscheidungsproblem abzuleiten und mögliche Prognoseobjekte zu bestimmen. Dies erlaubt eine Koordination der Datenaggregation und Entscheidungsgranularität auf eine einheitliche Bezugsgröße.

Die Problemdefinition erfordert somit ein intensives Verständnis des Entscheidungsproblems, der Interaktion mit anderen Entscheidungen, der betroffenen Organisationseinheiten und der Unternehmensteilfunktionen für die Beurteilung der möglichen Auswirkungen der Prognose [Makridakis et al. 1998, S. 13]. Die Phase schließt mit einer Bestimmung des Prognosegegenstandes in zeitlicher und mengenmäßiger Art. Diese grundsätzlichen Überlegungen definieren die Struktur des Entscheidungsproblems und geben den Entscheidungsrahmen für nachträgliche Phasen vor. Ausgehend von der Phase der Problemformulierung werden somit erste Hypothesen über relevante Erklärungsmodelle und Prognosevariablen aufgestellt, womit die nachfolgende Prozessphase der Informations- und Datenbeschaffung sinnvoll bestimmt und eingegrenzt wird [Brockhoff 1977].²

3.4.2.2 Informationsbereitstellung

3.4.2.2.1 Identifikation und Sammlung der Daten

In der Phase der Informationssammlung werden alle Tätigkeiten zur Transformation von Daten in entscheidungsrelevante Informationen für den Prognoseprozess subsumiert. Sie beinhaltet die verfahrensunabhängigen Teilschritte der Identifikation, Sammlung und

¹ So erfolgt die Disposition in Filialen des Einzelhandels vielfach unter Berechnung täglicher Prognosen für den Prognosezeitraum der nachfolgenden Woche, z. B. an einem Mittwoch t mit einem Prognosehorizont $h = 5, 6, \dots, 11$ zur Vorhersage der Absatzzahlen des darauf folgenden Montags ($t+5$) bis zum Sonntag ($t+11$).

² Aufgrund der Bedeutung der ersten Phase des Prognoseprozesses, der Subjektivität ihrer Spezifikation und der Veränderbarkeit der Entscheidungsvariablen und -parameter im Zeitablauf sind ihre Ergebnisse regelmäßig im Zuge der taktischen und strategischen Planung kritisch zu hinterfragen.

Aufbereitung bzw. Vorverarbeitung von Daten [Armstrong 2001e, S. 219-226; Gilchrist 1976; Armstrong 2001b], und wird in Abbildung 3.24 dargestellt.

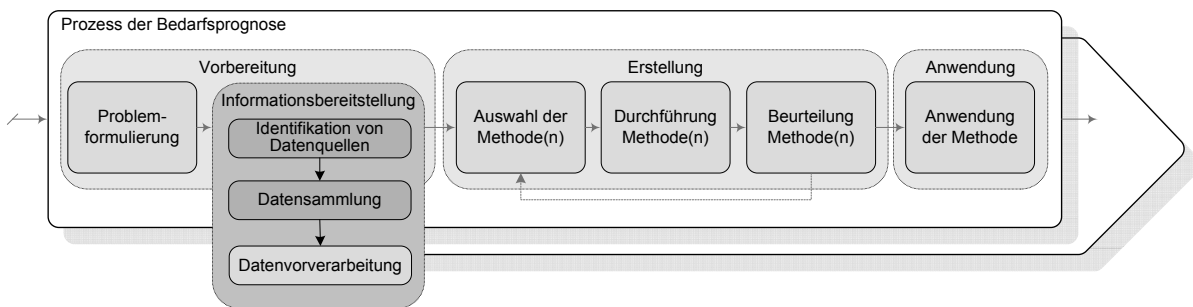


Abbildung 3.24: Informationsbereitstellung im Prognoseprozess

Die Identifikation potenziell informativer Primär- und Sekundärdatenquellen wird meist durch den Anwendungskontext, den Theorierahmen und die Datenverfügbarkeit für den Prognosegegenstand bestimmt.¹

Die Identifikation von Datenquellen für ein bestimmtes Prognoseobjekt beinhaltet die Suche nach unterschiedlichen Datenquellen innerhalb und außerhalb des Unternehmens. Dabei werden unternehmensinterne Daten unterschiedlicher Skalierung erhoben: quantitative Daten metrischer Skalierung umfassen Beobachtungen des historischen Bedarfs oder Absatzes, der Produktionsmengen, Warenbestände, vergangener und zukünftiger Preise oder Marketing-Budgets. Das Fehlen von Datenwerten, z. B. historischer Werbeaktionen, Preissenkungen oder Lieferengpässen, oder das Auftreten von Ereignissen kann als nominal skalierte Informationen durch binäre Dummy-Variablen erhoben werden. Die Erhebung nominaler Informationen ermöglicht auch die Abbildung zukünftiger Realisationen kontrollierbarer Aktivitäten, deren Eintreten zum Zeitpunkt der Prognose bekannt ist, deren Auswirkungen auf das Prognoseobjekt jedoch noch unbekannt sind, z. B. bei zukünftigen Werbeaktionen.

Die Suche umfasst ferner unternehmensexterne Datenquellen der politisch-rechtlichen, soziokulturellen und technologischen Umwelt, die bei langfristigem Prognosehorizont zusätzliche Informationen liefern können [Gilchrist 1976, S. 12]. Daten können auch qualitativer Art sein, in Form von Erwartungen und Expertenwissen, die entweder zur Berücksichtigung im Prognosemodell quantifiziert und kodiert werden oder extern durch eine anschließende Korrektur der Prognosewerte berücksichtigt werden.

Im Anschluss erfolgt eine Überprüfung des Informationsgehalts, der Validität und der Reliabilität der Daten, um systematische Verzerrungen in der Prognose zu vermeiden. Dabei entsteht das Entscheidungsproblem der Trennung relevanter von irrelevanten Informationen durch eine Ex-ante-Datenselektion.² Nach GILCHRIST [1976, S. 13] wird aus der Vielzahl potenzieller Datenquellen zunächst eine Auswahl von Daten mit hoher Relevanz für das Prognose- und Entscheidungsproblem, hoher Reliabilität und hoher Aktualität gebildet, welches er als minimale Informationsversorgung (engl. *minimum information set*) bezeichnet. Erst nach der Analyse dieser minimalen Datenmenge sollte der Datenrahmen sukzessive durch zusätzliche Informationen erweitert werden, die einen höheren Informationsgehalt für die Prognosesituation versprechen, jedoch geringere Relevanz oder Reliabilität besitzen. Dies entspricht der Forderung von GRANGER [1980], nicht eine einzige Informationsversorgung, sondern eine Auswahl unterschiedlicher Datenmengen I_g zu definieren. Entsprechend werden

¹ Neben der Suche von Informationen des Prognosegegenstandes kann auch nach Datenquellen mit ähnlichen oder analogen Informationen für verwandte Prognoseobjekte zur Erweiterung des Datenversorgung gesucht werden [Armstrong 2001e, S. 684], was besonders für die Prognose neuer Produkte relevant ist.

² Die Datensammlung initiiert zudem weitere Entscheidungsprozesse in Bezug auf die Frage, ob und in welcher Form und Frequenz weitere Daten sowie zusätzliche Prädiktorvariablen erhoben werden sollen.

mehrere plausible Datenszenarien für ein Erklärungsmodell gebildet, welche die iterative Evaluation und Selektion alternativer Modelle und Verfahren zur Verifikation bzw. Falsifikation vermuteter Datenzusammenhänge erlauben.¹ In der Absatzprognose besteht eine minimale Informationsversorgung zunächst aus der Zeitreihe der abhängigen Variablen, die eine Prognose durch Zeitreihenmodelle erlaubt. Diese wird sodann sukzessive um zusätzliche Informationen deterministischer exogener Variablen (Feiertage, Urlaubszeiten etc.) sowie stochastischer Einflussgrößen (Wetter, Marketing-Aktivitäten des Unternehmens sowie seiner Mitbewerber, Indikatoren der Konsum- und Marktentwicklung etc.) erweitert.²

Diese unterschiedlichen Datenmengen dienen als Kandidaten für eine spätere Modellierung durch Verfahren für unterschiedliche Erklärungsmodelle. Die Bereitstellung der Daten erfolgt somit verfahrensunabhängig unter Bereitstellung mehrerer plausibler Datenmengen, was eine einheitliche Beurteilung unterschiedlicher Prognosemodelle und -verfahren auf abweichenden Datenmengen ermöglicht. Die Identifikation relevanter Datenmengen ist zu einem hohen Grad abhängig von der konkreten Entscheidungssituation. Sie erfordert die subjektive Beurteilung durch Experten [Gilchrist 1976, S. 13] und ist daher nicht automatisierbar.

3.4.2.2 Datenvorverarbeitung

Die Phase der Datenvorverarbeitung³ überführt die gewonnenen Rohdaten in ein alternatives Format aus Skalenniveau, Verteilung, Aggregationsgrad und Beobachtungsintervall „mit dem Ziel .., aus der Originalzeitreihe geeignete Eingabedaten für das Prognoseverfahren zu erzeugen“ [Thiesing 1998, S. 80]. Sie wird in Abbildung 3.26 in den Prognoseprozess nach ARMSTRONG eingeordnet.

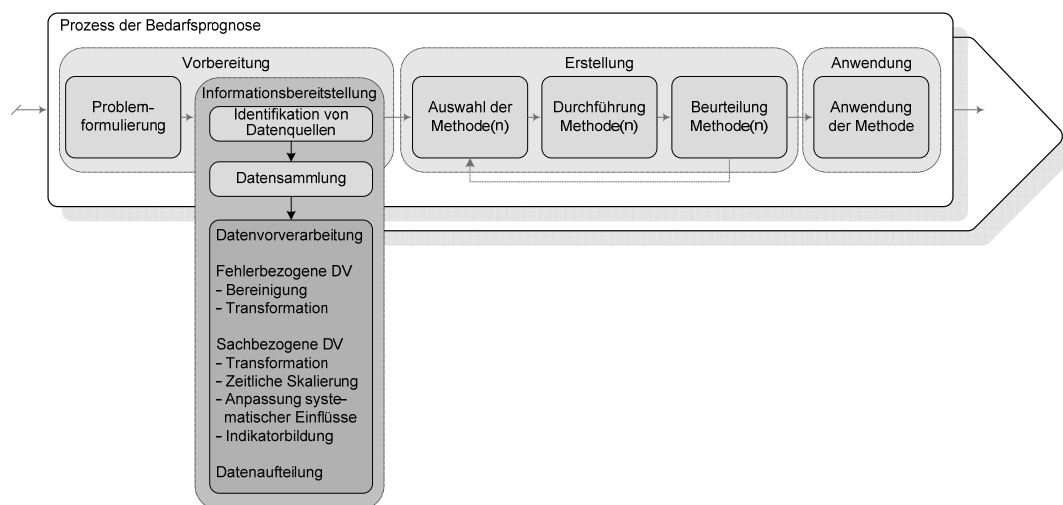


Abbildung 3.25: Verfahrensunabhängige Datenvorverarbeitung im Prognoseprozess

Es existiert eine Vielzahl von Methoden der Datenvorverarbeitung, die mit unterschiedlichen Zielsetzungen angewendet werden. Prognoseverfahren stellen unterschiedliche Anforderungen an Datenkodierung, Ausreißerbehandlung, Desaisonalisierung oder Trendbereinigung. Da je nach Beschaffenheit der Daten die Anwendung von Methoden der Vorverarbeitung für ein

¹ Eine alternative Vorgehensweise gestatten Neuronale Netze, in dem alle relevanten und irrelevanten Informationen durch eine Pre-Processing-Schicht an das Verfahren anzulegen und es eigenständig die Trennung relevanter von irrelevanten Werten erlernen zu lassen [Neuneier und Zimmermann 1998].

² Diese Vorgehensweise erscheint auch bei der Vermutung nichtlinearer Zusammenhänge vorteilhaft, da die Identifikation nichtlinearer Informationen nicht anhand linearer statistischer Tests erfolgt, sondern eine Beurteilung des Datensatzes im Modell erfordert. GILCHRIST [1976] geht darauf jedoch nicht ein.

³ Sie wird im Englischen als „data preprocessing“ oder „data preparation“ bezeichnet.

Prognoseverfahren optional oder sogar obligatorisch sein kann¹, ist der Umfang der Vorverarbeitung verfahrensspezifisch im Prognoseprozess zu berücksichtigen [Mertens und Deistler 1994]. Entsprechend unterscheidet JOERESSEN [1996; zitiert nach Thiesing 1998, S. 79] modellbezogene, fehlerbezogene und sachbezogene Methoden der Datenvorverarbeitung. Die modellübergreifenden Methoden der fehlerbezogenen und sachbezogenen Vorverarbeitung dienen der Bereitstellung einer einheitlichen Datenbasis für alle Modelle und Verfahren und werden nachfolgend dargestellt. Die verfahrensspezifischen Aktivitäten der Vorverarbeitung werden dagegen in die Phase der Prognoseerstellung eingeordnet und in Abhängigkeit der von den Verfahren diskutiert.² Alle Aktivitäten der Datenvorverarbeitung erfordern dabei eine hinreichende Dokumentation der Korrekturen und Transformationen [Armstrong 2001b].

Ziel der fehlerbezogenen Methoden der verfahrensübergreifenden Datenvorbereitung ist die Erzeugung einer regelmäßigen Zeitreihe aus den Zeitreihendaten durch Korrektur jener Unregelmäßigkeiten, die die Systematik der Zeitreihe verfälschen. „The advice to clean data applies to all quantitative methods, not just extrapolation“ [Armstrong 2001b, S. 222-223]. Korrekturen beinhalten vor allem die Rekonstruktion fehlender Zeitreihenwerte sowie die Korrektur jener extremen Beobachtungswerte der Ausreißer, die auf Fehler in der Datenerfassung zurückzuführen sind. In der Praxis wird aufgrund der Kenntnis der konkreten Zeitreihe die Identifikation von Ausreißern als erkennbar falsch aufgenommene Zeitreihenwerte vereinfacht, die häufig durch Mess- und Erfassungsfehler bei Eingabe oder Nacherfassung sowie durch Dateninkonsistenzen bei Datentransfers oder -konsolidierungen in Computersystemen entstehen [Armstrong 2001b, S. 222]. Hierunter fallen nicht Extremwerte, die der tatsächlichen Verteilung des Bedarfs angehören, z. B. im Zuge von Absatzspitzen durch erfolgreiche Werbeaktivitäten. Die Rekonstruktion fehlender Beobachtungswerte (engl. *missing value imputation*) einer Zeitreihe erfolgt i. d. R. durch Einsetzen der zentralen Lageparameter der Beobachtungswerte oder durch funktionale Anpassung eines Prognosemodells unter Verwendung der Prognosewerte für den fehlenden Wert. Ferner erfolgt eine Anpassung der Zeitreihen für strukturelle Brüche in der Historie [Armstrong 2001b; Thiesing 1998].³ Die fehlerbezogenen Ansätze werden somit auch als Methoden der Datenbereinigung (engl. *data cleansing*) bezeichnet [Armstrong 2001b, S. 222].

Sachbezogene Methoden umfassen die Transformation der Daten im Hinblick auf eine Behandlung systematischer Einzeleinflüsse. Hierunter fällt z. B. die Korrektur zur Abbildung sinnvoller, äquidistanter Zeitabstände der Beobachtungswerte, z. B. durch Interpolation von unregelmäßig erfassten Beobachtungen, die Anpassung der Daten an die wechselnde Zahl der Kalendertage oder Arbeitstage je Monat, die Umrechnung wöchentlicher in monatliche Zeitintervalle durch 4-4-5-Wochenkalender sowie die Korrektur unvollständiger Kalenderwochen zu Jahresende und -anfang, um einzelne Muster der Zeitreihen deutlicher hervorzuheben. Die temporale Aggregation in ein größeres Zeitintervall oder auch die Aggregation von Artikeln auf Marken, Artikelgruppen oder regionale Bereiche erlaubt ferner die Vermeidung von unregelmäßigen Beobachtungen fehlender Werte des sporadischen Bedarfs oder ungenauer Erhebungen, unter Rückkoppelung zu der Bestimmung des Prognosegegenstandes in der vorigen Phase der Problemformulierung [Armstrong 2001e].

¹ So erfordert z. B. die Berechnung statistischer ARIMA-Verfahren die Trendbereinigung zur Einhaltung der Prämissen der Normalverteilung und Stationarität. Die Integration trendbehafteter Zeitreihen könnte für Neuronale Netze hingegen zu einer verringerten Prognosegüte führen, da nutzbare Informationen der Niveauveränderungen aus der Zeitreihenstruktur verloren gingen. Dies ist somit separat zu evaluieren.

² Modellspezifische Verfahren der Datenvorverarbeitung zur Modellierung von Neuronalen Netzen werden in Abschnitt 4.3 detailliert untersucht. Sie dienen hier lediglich als Beispiele im Kontext des Gesamtprozesses.

³ Strukturbrüche die nicht dem datengenerierenden, stochastischen Prozess zuzuschreiben sind entstehen beispielsweise durch eine erneute Indizierung prozentualer Zeitreihen oder exogener Einflüsse der Inflation.

Darüber hinaus kann eine Transformation der Beobachtungswerte von Werten in [GE] in Mengeneinheiten [ME] sowie eine Transformation von Werten [GE] in dimensionslose prozentuale Veränderungen [-] erfolgen, um einzelne Muster der Zeitreihen zu eliminieren oder stärker hervortreten zu lassen. Ferner erlauben sachbezogene Methoden die Rekonstruktion der tatsächlichen Zielgröße aus den Beobachtungswerten, z. B. aus der Schätzung der vollständigen Verteilung des historischen Bedarfs aus den Zeitreihen des Absatzes mit Fehlmengen, die einseitig das obere Quantil der kumulierten Wahrscheinlichkeitsfunktion abschneiden. Die Transformation als Mittel der Anpassung einer Variablen an die etablierte ökonomische Theorie, z. B. per Logarithmierung zur Abbildung konstanter Elastizitäten anstelle direkter Beziehungen, ist ebenfalls der Zielsetzung einer numerisch günstigen Modellformulierung zuzuordnen [Armstrong 2001e].

Die Indikatorbildung beinhaltet die Transformation von relevanten Zeitreihen in Indikatoren im Sinne einer Datenreduktion [Gilchrist 1976]. Die Umformung in Indikatoren dient häufig der Aufbereitung der Daten in eine für das Modell geeignete Form, z. B. im Sinne der Dimensionsreduktion durch die Hauptkomponentenanalyse [Thiesing 1998, S. 84], oder auch einer Erhöhung der Dimensionalität, z. B. durch Ableitung mehrerer Zeitreihen der ersten und höheren Differenzen als zusätzliche Informationsträger der Veränderung der Zeitreihen [Neuneier und Zimmermann 1998]. THIESING [1998, S. 80] betrachtet für die Verfahren der Neuronalen Netze auch das Hervorheben oder Ausblenden einzelner Zeitreihenkomponenten wie Trends, Saisonalitäten oder einzelner kausaler Sondereinflüsse sowie die Korrektur von Strukturbrüchen, Ausreißern und fehlenden Werten als sachbezogene Methoden. Der Einsatz dieser Methoden muss in Anbetracht der unterschiedlichen Sensitivität jedoch verfahrensspezifisch erfolgen und wird nachfolgend der modellbezogenen Vorverarbeitung zugeordnet.

Der abschließende Schritt der Datenvorverarbeitung erfordert die Aufteilung der Beobachtungswerte in disjunkte Datenmengen zur Parametrisierung und Beurteilung anhand eines Quasi-ex-ante-Vergleichs. Die Vorgehensweise bei der Datenaufteilung folgt den Empfehlungen empirischer Vergleichsstudien in Abschnitt 3.3.3 und wird hier nicht erneut abgebildet. Die Parametrisierung erfolgt auf der Grundlage der Kalibrationsstichprobe, die Beurteilung der Prognosegüte auf einer "Out-of-sample-Teststichprobe, wobei alle nachfolgenden Prozessschritte der Datenanalyse ausschließlich die Kalibrationsstichprobe verwenden dürfen. Erfolgt z. B. eine Skalierung oder Ausreißerkorrektur unter Kenntnis der Beobachtungswerte der Testmenge, wären die Prämissen einer Ex-ante-Prognose verletzt. Bei der Datenaufteilung muss ferner die Notwendigkeit einer ausreichend großen Datenmenge für jede Datenstichprobe berücksichtigt werden, auch bei mehrfachen Prognoseursprüngen und Prognosehorizonten unterschiedlicher Länge.

3.4.2.3 Auswahl der Verfahren

3.4.2.3.1 *Explorative Datenanalyse*

Nach der Vorverarbeitung können die Daten auf Strukturen der Zeitreihenkomponenten sowie kausale Interaktionen mit anderen Zeitreihen analysiert werden, darauf aufbauend sodann adäquate Prognosemodelle und -verfahren ausgewählt werden.

Zur Identifikation der systematischen Komponenten der Saisonalität, Trends und Zyklen existiert eine Vielzahl statistischer und heuristischer Parametertests. Zunächst kann eine einfache grafische Evaluation durch Zeitreihen- und Saisondiagramme visuelle Hinweise geben. Ansätze der Dekomposition von Zeitreihen geben Aufschluss über die Interaktion einzelner Komponenten metrischer Zeitreihen und eine Klassifikation einerseits in regelmäßige Zeitreihenverläufe stationärer, trendbehafteter, saisonaler oder trendsaisonaler Form, andererseits in unregelmäßige Verläufe sporadischer Beobachtungen. Zudem wird die Identifikation irregulärer Komponenten der Ausreißer, Niveauverschiebungen und Struktur-

brüche begünstigt und es werden Anhaltspunkte für die Bedeutung der einzelnen Komponenten in Bezug auf die Gesamtvariabilität der Zeitreihe geliefert.

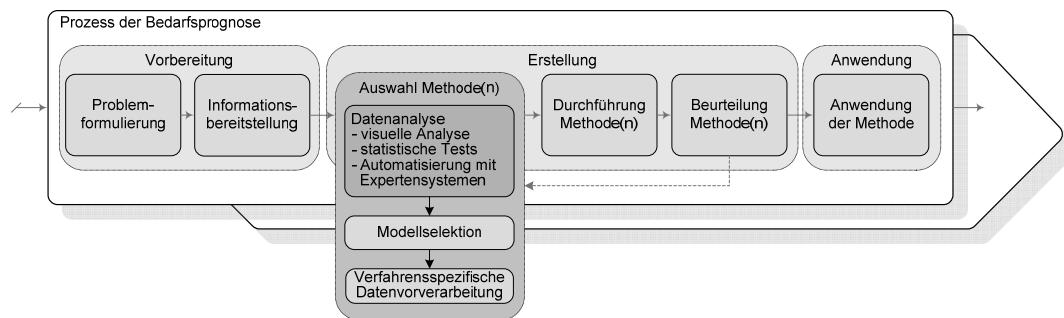


Abbildung 3.26: Datenanalyse im Prognoseprozess

Zudem erlaubt die Berechnung statistischer Momente der Lage und Streuung auf unterschiedlichen Teilbereichen einer Zeitreihe Aussagen über ihre Verteilung und Stationarität [Makridakis et al. 1998, S. 14]. Eine Analyse der Korrelogramme der absoluten und partiellen Autokorrelationskoeffizienten bei iterativer Integration der Zeitreihen ermöglicht Rückschlüsse auf die Struktur des datengenerierenden Prozesses hinsichtlich Trends und Saisonalität und erlaubt erste Rückschlüsse auf eine adäquate ARIMA-Modellform. Darüber hinaus erfordert die Identifikation sich überlagernder Komponenten multipler Saisonalitäten mit unterschiedlicher Zykluslänge eine Analyse durch Verfahren der Spektralanalyse [Hansmann 1983, S. 53-62].

Diesen explorativen Methoden der Zeitreihenanalyse stehen automatisierbare, statistische Tests der Zeitreihenkomponenten entgegen, die aus der ökonometrischen Zeitreihenanalyse stammen. Zur Beurteilung der Zeitreihe hinsichtlich der Verlaufsform unregelmäßiger oder sporadischer Verläufe genügt i. d. R. eine einfache Entscheidungsregel bezüglich des Anteils der Nullperioden der Zeitreihe. Darüber hinaus wurden in der statistischen Zeitreihenanalyse vielfältige Tests zur Identifikation von Zeitreihenmustern entwickelt. Unit-Root-Tests bestimmen das Vorliegen eines stochastischen Trends durch die Signifikanz der Parameter unterschiedlicher, linearer Regressions- bzw. Autoregressionsgleichungen in den Varianten des Dickey-Fuller-Tests [Dickey und Fuller 1979], des Augmented-Dickey-Fuller-Tests [Said und Dickey 1984] oder des Phillips-Perron-Test der Zeitreihe zur Analyse mittelwertstationärer Verläufe. Die Analyse systematischer Komponenten erfolgt auch durch Runs-Tests bzw. Wald-Wolfowitz-Tests, welche die Zufälligkeit einer Folge zeitlich geordneter Beobachtungswerte aus den linearen Autokorrelationskoeffizienten untersuchen. Neben dem einfachen Test einer signifikanten Autokorrelation im Lag $p = 1$ AR(1) durch den Durbin-Watson-Tests [Durbin und Watson 1971; Durbin und Watson 1951; Durbin und Watson 1950] untersuchen Portmanteau-Tests wie der Ljung-Box-Test, der Breusch-Godfrey-Test oder die Box-Pierce-Statistic auch die Signifikanz der Koeffizienten für Zeitfenster beliebiger Länge [Makridakis et al. 1998].

Lineare Interaktionen zwischen Zeitreihen können durch Regressions- und Kreuzkorrelationsanalysen nach Pearson sowie durch rangbasierte Korrelationen nach Spearmans rho analysiert und entsprechend in der Modellierung berücksichtigt werden [Lyerly 1952]. Der Chow-breakpoint Test überprüft die Zeitreihen auf strukturelle Brüche anhand mehrerer lokaler, linearer Regressionsmodelle [Chow 1960]. Viele der Tests können auch auf den Residuen von Prognoseverfahren eingesetzt werden, ebenso der Ramsey-RESET-Test bei einer Modellfehlspezifikation des linearen Regressionsmodells [Ramsey 1969] oder der Chow-Forecast Test zum Testen der Konstanz der Verfahrens-Parameter. Zur Bestimmung der Form der Wahrscheinlichkeitsverteilung überprüft der verteilungsfreie Kolmogorov-Smirnov-Test, ob die

Stichprobe der Beobachtungswerte einer zuvor angenommenen Wahrscheinlichkeitsverteilung folgt [Massey 1951]. Darüber hinaus erlauben der Shapiro-Wilk-Test oder der Anderson-Darling-Test die Überprüfung der Verteilungsfunktion auf die überwiegend unterstellte Normalverteilung [Shapiro und Wilk 1965]. Expertensysteme wie z. B. die Softwaresysteme „ForecastPro“ [Goodrich 2000] oder „Autobox“ [Reilly 2000], imitieren Experten durch eine regelbasierte Kombination dieser Tests in einem virtuellen Prognoseprozess zur Identifikation der Zeitreihenform und zur Auswahl eines passenden Prognoseverfahrens.

Allerdings ist das verbreitete statistische Instrumentarium zur validen und reliablen Analyse von Mustern und Interaktionen auf lineare Erscheinungen beschränkt. Nichtlineare Effekte, wie sie in der betrieblichen Realität regelmäßig auftreten, verzerren die Erkenntnisse der Tests und die Güte linearer Prognoseverfahren. Sie können häufig nur durch iterative Datentransformation, Schätzung der Modelle und neuerliche Analyse erkannt und durch Umformungen für die Anwendung linearer Verfahren entfernt werden.

3.4.2.3.2 Modell- und Verfahrensselektion

Ausgehend von den Erkenntnissen der explorativen Datenanalyse kann anhand der identifizierten Muster einer Zeitreihe oder der Interaktion zwischen Zeitreihen anschließend eine Auswahl konkurrierender Erklärungsmodelle als Hypothesen postuliert werden, die die vorhandene Datensituation und Systematik hinreichend beschreiben und somit zu ihrer Prognose in die Zukunft genutzt werden können. Abbildung 3.27 ordnet die Phasen der Modellselektion und der Verfahrensselektion in den Prognoseprozess ein.

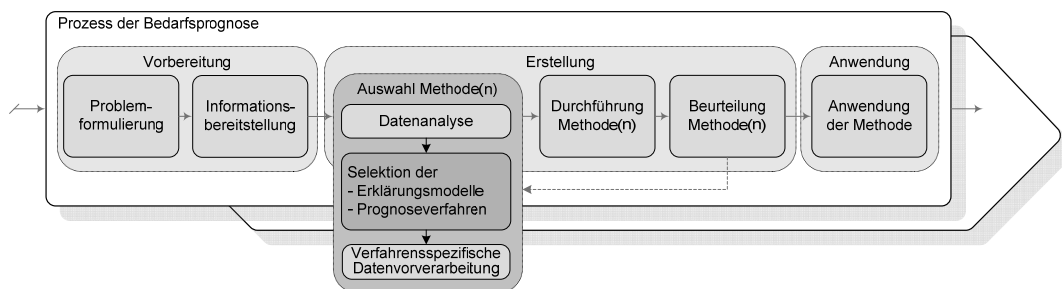


Abbildung 3.27: Modellselektion und Verfahrensselektion im Prognoseprozess

Die objektiven Modelle werden gemäß der beobachteten Interdependenz von Daten und Prognosegegenstand gebildet und in einem mathematischen Modell spezifiziert [Gilchrist 1976]. Die Auswahl von Modellen vollzieht sich in Abhängigkeit der verfügbaren Daten, der Kenntnis des Anwendungsbereiches und der erwarteten Konsistenz der Daten und Erklärungsmodelle [Armstrong 2001e, S. 693]. Nach BROCKHOFF [1977] erfolgt zunächst die Formulierung eines Erklärungsmodells des Prognosegegenstandes, bei kausalen Modellansätzen unter Bestimmung der potenziell zu berücksichtigenden Variablen, der Art ihrer funktionalen Beziehung, der Richtung der Beziehungen und der Stärke der Beziehungen. Für zeitreihenanalytische Modelle erfolgt die Bestimmung der Modellform durch die identifizierten systematischen Zeitreihenkomponenten von Trend und Saisonalität sowie ihrer funktionalen Verknüpfung. Anschließend wird eine Hypothese über die Bedeutung des Erklärungsmodells als Prognosemodell unter Berücksichtigung der Zeitstabilitätshypothese, z. B. im Hinblick auf vergangene oder zukünftige Strukturbrüche, formuliert. Wie unter 3.1.2 dargestellt, existiert jedoch eine Vielzahl alternativer Modelle, uni- oder multivariater Zusammenhänge der Kausal- und Zeitreihenanalyse, die in Abhängigkeit der unterschiedlichen Datenmengen zur Anwendung kommen können. Somit kann eine Vielzahl von Erklärungsmodellen auf der Basis einer oder mehrerer Informationsversorgungen gebildet werden. Abbildung 3.28 gibt eine Handlungsempfehlung zur Modellsektion in Abhängigkeit der Datengrundlage nach den „Principles of Forecasting“ [Armstrong 2001d].

Das Diagramm verdeutlicht, dass selbst bei einer Beschränkung auf objektive Daten und auf formale, quantitative Erklärungsmodelle zur Anwendung im Handel, die auf der rechten Seite des Baumdiagramms dargestellt werden, unterschiedliche multivariate Erklärungsmodelle der Kausalanalyse sowie diverse univariate Modelle der Zeitreihenanalyse in Form von Extrapolativen Methoden, Neuronalen Netzen sowie Regelbasierten Modellen anwendbar sind.¹ MAKRIDAKIS et al. schlagen bei Vorliegen von mehr als einem realistischen Erklärungsmodell die Beschränkung auf „one or two leading contenders for subsequent analysis“ vor [Makridakis et al. 1998, S. 15]. Somit entsteht ein Problem der Modellselektion, weil konkurrierende Kandidaten unterschiedlicher Erklärungsmodelle und ggf. Datenmengen miteinander verglichen werden müssen. Entsprechend können Prognosemodelle unterschiedlicher Klassen basierend auf unterschiedlichen Datenstrukturen gebildet werden und somit lineare Modelle der Zeitreihenanalyse gegenüber linearen und nichtlinearen Modellen der Kausalanalyse evaluiert werden. Können z. B. keine kausalanalytischen Interdependenzen zwischen den Variablen, wohl aber Autokorrelationen innerhalb der Zeitreihe der abhängigen Variablen beobachtet werden oder liegen nur Beobachtungswerte der abhängigen Variablen vor, kann die Modellevaluation auf univariate Zeitreihenmodelle beschränkt werden.

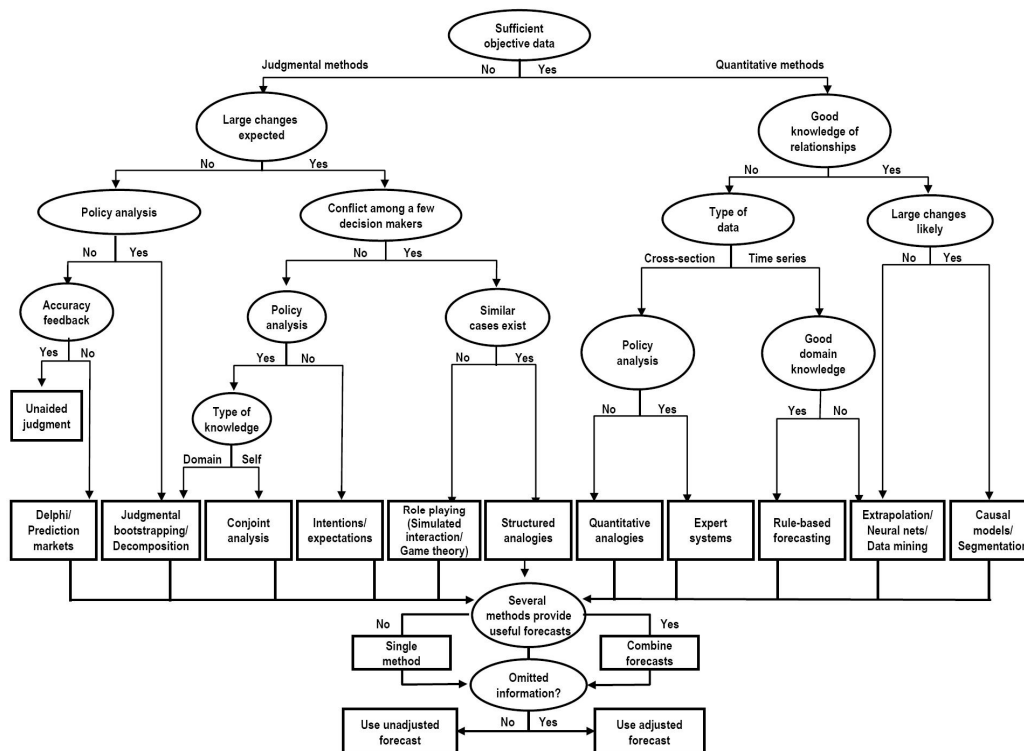


Abbildung 3.28: Baumdiagramm zur Modellselektion²

Anschließend ist für jedes der Prognosemodelle ein angemessenes Prognoseverfahren zur Schätzung auszuwählen. Existiert mehr als ein Verfahren für das jeweilige Modell, muss im Sinne der Verfahrensselektion eine sinnvolle Teilmenge relevanter Prognoseverfahren ausgewählt werden.³ So kann ein univariates Modell einer trendbehafteten Zeitreihe durch

¹ Dabei ist darauf hinzuweisen, dass Armstrong [2001d] in diese Darstellung sowie grundsätzlich nicht trennscharf zwischen Modellen und Verfahren unterscheiden. So sind Neuronale Netze insbesondere in der Lage, sowohl lineare als auch nichtlineare zeitreihenanalytische und kausalanalytische Prognosemodelle abzubilden, was die Darstellung vernachlässigt.

² Darstellung entnommen aus [Armstrong 2001d].

³ Dieses wird im Allgemeinen als „relevant set“ der Verfahren bezeichnet und korrespondiert mit den „information sets“ der Datenversorgung. Eine systematische Evaluierung erfordert somit einen umfassenden Experimentaufbau mit mehreren Kandidaten der Verfahren und Informationssätze in Verfahrensvergleichen.

unterschiedliche Verfahren (z. B. Exponentielle Glättung, ARIMA oder Neuronale Netze) prognostiziert werden. Alternativ kann die Zeitreihe durch Transformationen in eine stationäre Zeitreihe umgewandelt und gemäß ihrer stationären Struktur durch Verfahren der einfachen exponentiellen Glättung, AR-, MA- oder ARMA-Verfahren, gleitende Mittelwerte oder Neuronale Netze prognostiziert werden. Die Anwendung einer Verfahrensvariante der Exponentiellen Glättung oder der ARIMA-Modelle erfordert dabei die korrekte Modellselektion in Abhängigkeit des zugrunde liegenden Musters der Zeitreihe, um systematische Fehler durch Fehlspezifikation des Modells zu vermeiden. Somit bildet sich in Abhängigkeit der Daten eine Menge alternativer Prognoseverfahren zur Lösung eines Prognosemodells, die anschließend auf den Daten parametrisiert werden und ex ante hinsichtlich der resultierenden Prognosequalität zu evaluieren sind [Thiesing 1998, S. 68 f.].

Unter Berücksichtigung theoretischer und praktischer Prinzipien sind wohlstrukturierte, objektive Prognoseverfahren mit transparenter Methodologie den subjektiven Verfahren bei identischen Prognosefehlern aufgrund der Effizienz, Robustheit und Automatisierbarkeit in der Anwendung vorzuziehen. Innerhalb der objektiven Modelle sind gemäß der Forderung MEFFERTS [1980, S. 147] nach einem kausallogisch befriedigenden Modell aller relevanten Einflussfaktoren der Absatzmärkte (vgl. Abschnitt 3.1.2.4) die kausalanalytischen Verfahren den zeitreihenanalytischen Verfahren vorzuziehen [ebenso Armstrong 2001e; Armstrong 1985; Allen und Fildes 2001]. Während diese theoretische Überzeugung die Anwendung komplexer, nichtlinearer kausalanalytischer Verfahren wie der Neuronalen Netze begünstigen würde, sind einfache Verfahren aus Gründen der Wirtschaftlichkeit in Anwendungen zu bevorzugen, solange nicht empirische Beweise dafür vorliegen, dass komplexere Verfahren den zusätzlichen Aufwand rechtfertigen. Somit müssen die alternativen Prognosemodelle und Prognoseverfahren hinsichtlich ihrer Kosten einerseits und den durch sie erzielbaren Zugewinnen durch eine gesteigerte Prognosegüte andererseits beurteilt werden [Armstrong 2001e, S. 693]. Dies erfordert einen objektiven Vergleich mit anderen Verfahren, wodurch einzelne Verfahrenskandidaten iterativ modifiziert und bei suboptimaler Prognosequalität für eine Anwendung verworfen werden [Brockhoff 1977]. Da eine Beurteilung der Güte konkurrierender Erklärungsmodelle ihre Schätzung durch ein Prognoseverfahren erfordert, entsteht das Problem der kombinierten Modellselektion und Verfahrensselektion in einem empirischen Verfahrensvergleich (vgl. Abschnitt 3.3.3).

Neben der Güte als Selektionskriterium können nach ARMSTRONG [2001e, S. 695] auch weitere Kriterien zur Einschränkung der Modell- und Verfahrensmengen verwendet werden. Insbesondere kann die Beurteilung der Verfahren durch objektive Experten sowie anhand bereits bestehender Erkenntnisse hinsichtlich ihrer Prognosequalität erfolgen, z. B. anhand von Ex-ante-Verfahrensvergleichen in ähnlichen Anwendungen. Grundsätzlich erfordert die Existenz mehrerer geeigneter Prognosemodells oder Prognoseverfahren jedoch immer die Beurteilung durch einen systematischen empirischen Verfahrensvergleich [Armstrong 2001e].

3.4.2.3.3 *Verfahrensspezifische Datenvorverarbeitung*

Verfahrensbezogene Methoden der Datenvorverarbeitung zielen auf eine obligatorische Gewährleistung der Prämissen oder eine günstige Formulierung der Eingabedaten im Hinblick auf statistische, numerische und algorithmische Belange des jeweiligen Prognoseverfahrens.¹ Abbildung 3.29 ordnet die verfahrensspezifischen Vorverarbeitungen in die Phasen der Prognoseerstellung in Abgrenzung zu den fehler- und sachbezogenen Methoden der Prognosevorbereitung ein.

¹ Im Gegensatz zu JOERESSENS [1996] Unterscheidung werden sie daher auch als verfahrensspezifische und nicht als modellspezifische Vorverarbeitungen bezeichnet.

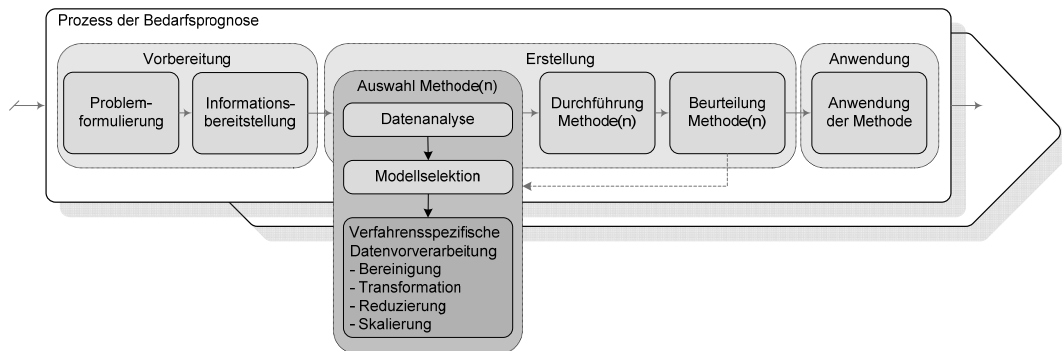


Abbildung 3.29: Verfahrensspezifische Datenvorverarbeitung im Prognoseprozess

Die Algorithmen zur Berechnung einzelner Prognoseverfahren setzen vielfach eine obligatorische Datenvorverarbeitung zur Einhaltung statistischer Prämissen voraus, z. B. für Regressionsverfahren in Form der Normalverteilung der Beobachtungswerte, der Linearität, der Stationarität von Mittelwert und Varianz, der Homoskedastizität, der Korrektur von Ausreißern etc. Zeitreihen der kausalanalytischen Modellierung sind auf Multikollinearität zu untersuchen und entsprechend auszuschließen. Dies erfordert vielfach eine iterative Vorverarbeitung insbesondere kausalanalytischer Ansätze, die nicht verfahrensübergreifend anwendbar ist.

Die Anwendung von ARIMA-Verfahren erfordert das Vorliegen der Daten als stationären stochastischen Prozesses ohne Trend, weshalb die iterative Differenzbildung bzw. Integration der Zeitreihe eine zentrale Stellung in der BOX-JENKINS-Methodologie einnimmt (vgl. Abschnitt 3.2.3.3.5). So ist die Trendbereinigung zur Sicherstellung der Stationarität einer Zeitreihe für einige statistische Verfahren wie z. B. ARMA oder S.ES erforderlich, während trendorientierte Verfahren wie D.ES oder L.ES bei Fehlen eines Trends zu einer Überinterpretation des Rauschens als Trend neigen und somit eine Trendbereinigung negative Auswirkungen hätte. Ferner erfordern lineare Regressionsverfahren, ARIMA oder Exponentielle Glättung die Transformation nichtlinearer in lineare Zusammenhänge, um eine Prognose mittels linearer Prognoseverfahren zu erlauben¹, während nichtlineare Verfahren wie die Neuronale Netze eine derartige Transformation nicht benötigen.²

Die Güte statistischer Prognoseverfahren kann ferner durch eine geeignete optionale Datenvorverarbeitung verbessert werden. Vorverarbeitungen in Richtung auf eine numerisch günstige Modellformulierung beinhalten z. B. die Skalierung oder Normierung der Beobachtungswerte auf ein Intervall, die Trend- oder Saisonbereinigung sowie die Ausreißerbereinigung zur Steigerung von Effizienz, Validität und Reliabilität eines Verfahrens bzw. zur Verringerung numerischer Probleme [Thiesing 1998, S. 81].³ Als Ausreißer gelten extreme Beobachtungswerte, die nicht dem originären stochastischen Prozess entspringen oder seine Struktur verfälschen. Auch wenn einzelne Verfahren als besonders Ausreißersensibel gelten⁴ kann ihre Korrektur die Prognosequalität auch anderer Verfahren erhöhen [Armstrong 1985, S. 459-461]. Eine Korrektur erfolgt z. B. durch sogenanntes Winsorizing, also das Ersetzen des Ausreißers durch einen determinierten Maximalwert mittels lokaler oder globaler Mittelwerte bzw. Mediane der Zeitreihe. Alternativ werden die Werte mit den

¹ Mit einigen Ausnahmen, z. B. des Verfahrens der Exponentiellen Glättung mit gedämpften Trends oder Multiplikativer Saisonalität, die eine Approximation nichtlinearer Zeitreihenverläufe zulassen.

² Neuronale Netze erlauben durch die hohe Flexibilität in der Modellierung die Verwendung redundanter Informationen und versprechen selbst bei multikollinearen Problemen robuste Ergebnisse [Crone et al. 2006e]. Die empirische Güte und Robustheit neuartiger Verfahren wie der Neuronale Netze ist jedoch im Einzelfall zu prüfen, da theoretische Fähigkeiten noch keine Aussagen über die resultierende Prognosegüte zulassen.

³ Sie lassen sich jedoch nicht immer von der obligatorischen Vorverarbeitung trennen, sofern unterschiedliche Theorien zur sinnvollen Parametrisierung einzelner Verfahren vorliegen.

⁴ Beispielsweise eine lineare Regression mit Ausreißern zum Beginn oder zum Ende der Zeitreihen.

geschätzten Prognosewerten ersetzt.¹ Darüber hinaus können exogene Erklärungsvariablen als metrische, ordinale oder nominale Eingabewerte skaliert werden und so die Prognosegüte beeinflussen.²

3.4.2.4 Durchführung der Verfahren

Die ausgewählten Verfahrenskandidaten werden anschließend durch Bestimmung der freien Parameter auf den transformierten Daten berechnet, im Sinne einer Anwendung des Erklärungsmodells als Prognosemodell mit vorgegebenen Messvorschriften und Daten [Brockhoff 1977].

Die Parametrisierung erfolgt üblicherweise computergestützt mittels einfacher Heuristiken, statistischer Verfahren oder Ansätzen der mathematischen Optimierung.

Die Parameterfindung erfolgt i. d. R. durch Veränderung der Verfahrensparameter zur Minimierung der Abweichungen zwischen der Prognose und den Beobachtungswerten der Zeitreihe, gemessen anhand einer explizit formulierten Zielfunktion der statistischen Fehlermaße oder einer implizit durch ein Verfahren oder ein Softwarepaket angewendeten Zielfunktion, meist auf der Grundlage Kleinst-Quadrat-Schätzer. Die anzuwendende Methode der Parameterfindung wird in Abhängigkeit von der funktionalen Form und der Komplexität des Prognosemodells gewählt. Zur Schätzung eines einzelnen oder weniger Parameter, z. B. bei Verfahren der Exponentiellen Glättung, können einfache Ansätze wie die vollständige Enumeration aller Parameterkombinationen oder die Rastersuche (engl. *grid search*)³ in definierten Intervallschritten für Kombinationen der freien Parameter verwendet werden. Die Parameterkonstellation mit dem geringsten Fehler auf den Parametrisierungsdaten wird zur Prognose ausgewählt. Alternativ werden Ansätze der linearen Optimierung einer Zielfunktion der statistischen Prognosefehler angewendet [Makridakis et al. 1998, S. 176]. Für die rekursive Parameterbestimmung in ARIMA-Verfahren sowie für nichtlineare Modellformulierungen müssen dagegen komplexe, nichtlineare Optimierungsverfahren wie z. B. lokale Gradientenabstiegsverfahren Anwendung finden, die nicht immer optimale Parameterkonstellationen garantieren. Vor dem Hintergrund verfügbarer Rechenleistung stellt die Berechnung der Parameter selbst umfangreicher Prognosemodelle keinen Engpassfaktor in der Verfahrensanwendung der Prognose dar.

Die Parametrisierung der Verfahren erfordert zunächst ihre Initialisierung mit Startwerten durch geeignete Schätzverfahren, die je nach Prognoseverfahren erheblichen Einfluss auf die Prognosequalität haben können [Makridakis et al. 1998, S. 174]. Eine Übersicht und Diskussion alternativer Ansätze zur Bestimmung initialer Werte für Verfahren der Exponentiellen Glättung findet sich in [Hansmann 1983, S. 29; Makridakis et al. 1998, S. 174-175; Scheer 1983]. Diese zählen ebenfalls zu den Freiheitsgraden in der Modellierung und sind entsprechend als freie Verfahrensparameter zu berücksichtigen.

¹ Die vollautomatische statistische Identifikation von Ausreißern als auch eine grafische Analyse von Ausreißern – sowie ihres Pendant, der „Einreißer“ [Reilly 2000] – birgt erhebliche Probleme und kann zu weiteren Störungen der Zeitreihe führen.

² In Bezug auf die Verfahren der Neuronalen Netze wird vielfach noch diskutiert, ob eine Trend- oder Saisonbereinigung, eine Ausreißerkorrektur oder eine Skalierung und Kodierung von exogenen Variablen für eine valide und reliable Prognose überhaupt erforderlich ist [Zhang und Qi 2005]. Im Gegensatz zu statistischen Verfahren erfordern Neuronale Netze jedoch die Skalierung der Beobachtungswerte auf die Intervalle ihrer Verarbeitungsfunktionen [0; 1] bzw. [-1; 1]. Ihre spezifischen Erfordernisse werden in Abschnitt 4.3. näher untersucht.

³ Es existieren vielfältige Heuristiken zur Definition und schrittweisen Verfeinerung des Rasters im *grid search*, auf die an dieser Stelle nicht näher eingegangen werden soll.

3.4.2.5 Beurteilung der Verfahren

Die Evaluation konkurrierender Prognoseverfahren erfolgt unter Nutzung der unter 3.3.2 dargestellten Gütemaße der Prognosequalität. Dabei sind die Kriterien zur Durchführung valider und reliabler Verfahrensvergleiche aus Abschnitt 3.3.3. zu berücksichtigen. Abbildung 3.30 ordnet die Phase der Beurteilung in den Prognoseprozess ein.

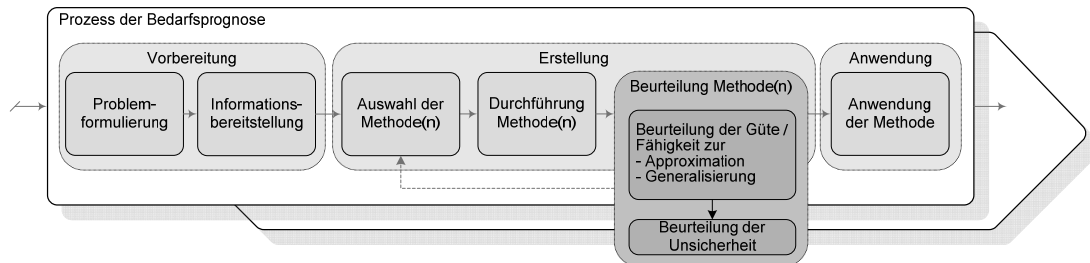


Abbildung 3.30: Beurteilung der Verfahrensgüte im Prognoseprozess

Die Güte jedes Prognoseverfahrens wird getrennt auf der Kalibrationsstichprobe (zur Bestimmung der Approximationsfähigkeit) und der Teststichprobe (zur Bestimmung der Generalisierungsfähigkeit durch die Quasi-ex-ante-Prognose auf unbekanntem Daten) bestimmt. Für U Prognoseursprünge unterschiedlicher Zeitpunkte und einem maximalen Prognosehorizont von H werden anschließend die Prognosefehler jeder Beobachtungsperiode über alle $h = 1, \dots, H$ und alle $u = 1, \dots, U$ zu statistischen Fehlermaßen aggregiert.

Bei Verwendung mehrerer Informationsversorgungen, Prognosemodelle, Prognoseverfahren, Gütemaße, Prognoseursprünge und Horizonte erfolgt die Beurteilung der Verfahrensgüte als systematischer Vergleich in umfangreichen Ergebnistabellen unter Berechnung statistischer Signifikanztest zur Feststellung signifikanter Unterschiede der Prognosefehler. Die Güte eines Verfahrens sollte stets als relative Güte im Vergleich zu konkurrierenden Prognoseverfahren beurteilt werden [Armstrong 2001e, S. 709]. Bei Überprüfung nur eines einzigen Modells oder Verfahrens kommt den relativen Fehlermaßen im Vergleich mit der Naiven Prognose eine besondere Bedeutung zu. Ferner sollten die Messvorschriften zur Beurteilung der Verfahren vor der Beurteilung festgelegt und die wahren Entscheidungskosten in einer Kosten-Nutzen-Analyse überprüft werden.¹

Die gefundenen Ergebnisse sollten einer Replikation der Verfahrensdurchführung sowie einer Replikation der Beurteilung standhalten, um Fehler im Prognoseprozess aufzuzeigen [Armstrong 2001e]. Nach der Bestimmung des Verfahrens mit der höchsten Prognosegüte wird dieses schließlich in der empirischen Anwendung eingesetzt.

3.4.2.6 Anwendung der Verfahren

Im letzten Schritt des Prognoseprozesses wird das Prognoseverfahren in der betrieblichen Planung auf unbekannte Daten der Zukunft angewendet.

Dazu wird das selektierte Verfahren zunächst auf dem gesamten Datenmaterial von Parametrisierungsmenge und Testmenge bis zum Gegenwartszeitpunkt parametrisiert und anschließend für die unbekannte Zukunft prognostiziert. Es erstellt entsprechend echte Ex-ante-Prognosen für unbekannte Prognosehorizonte der Zukunft, die erst nach Eintreten des wahren Beobachtungswertes überprüft werden können. Dies erfordert die dauerhafte Überwa-

¹ Neben der Beurteilung der Prognosegüte kann eine Beurteilung der Unsicherheit der Prognose anhand statistischer Konfidenzintervalle über den Prognosehorizont erfolgen [Armstrong 2001e, S. 718-722]. Diese dienen als Indiz für die Robustheit der Prognosegüte einzelner Verfahren und können auf alternative Informationsmengen, Prognosemodelle und Prognoseverfahren zurückgeführt werden, um die Auswahl iterativ zu verfeinern.

chung und Kontrolle der Prognoseergebnisse durch eine zyklische Evaluation der Prognosequalität aller vorangegangenen Prozessschritte. Die laufende Ex-post-Kontrolle der Prognoseergebnisse erlaubt zudem die Gewinnung neuer Hypothesen [Hansmann 1983, S. 13] und somit neuer Ansätze zur Verbesserung des Prognoseprozesses.

Die Anwendung der objektiven Prognoseverfahren erfolgt fast ausschließlich durch computergestützte Planungssysteme der Prognose. Dabei kommen spezielle Systeme zur Prognoserechnung zum Einsatz, insbesondere sogenannte „Forecasting Support Systeme“ als Expertensysteme mit speziellen Methodenbanken der Prognose zur Auswahl und Durchführung unterschiedlicher Verfahren [Preßmar 1995] sowie sogenannte „Advanced Planning Systems“ als Informations- und Planungssysteme des Supply Chain Managements.

3.4.3 Fehlerquellen im Prognoseprozess

Die Prognosegenauigkeit einzelner Erklärungsmodelle bzw. Prognoseverfahren ist in der empirischen Anwendung nicht allein auf die technische Güte der Modelle und Verfahren zurückzuführen. Vielmehr existieren vielfältige Fehlerquellen bei der Prognoseanwendung, die durch suboptimale Methodologien oder systematische Fehlentscheidungen in der Auswahl, Parametrisierung und Beurteilung der Verfahren die Prognosequalität verzerren. SCHNEEWEIß [1971, S. 76 ff.] nennt vier Gruppen von möglichen Fehlerquellen, deren gemeinsame Wirkung den Prognosefehler ausmachen:

1. Fehlspezifikation des Erklärungsmodells
2. Fehler in der Spezifizierung der Zeitstabilitätshypothese (Strukturbrüche)
3. Fehler in den unabhängigen Variablen
 - a. Messfehler der deterministischen unabhängigen Variablen
 - b. Prognosefehler in der Prognose stochastischer unabhängiger Variablen
4. Schätzfehler in den geschätzten Parameterwerten

SCHWARZE [1973] erweitert die Systematik möglicher Fehlerquellen in der Spezifikation des Erklärungsmodells: die Auswahl der beteiligten Variablen, die Festlegung der Funktionsform, die Festlegung der stochastischen Eigenschaften der Störvariablen, die Festlegung der stochastischen Eigenschaften der prädeternierten und der unverzögerten endogenen Variablen und die Festlegung der Beobachtungsperiode.¹ Die einzelnen Problemstellungen wurden in der ökonomischen Literatur aufgegriffen und ausführlich behandelt [Brockhoff 1977, S. 53; Hochstädter und Uebe 1970; Schneeweiß 1971; Schönfeld 1969]. Diese potenziellen Fehlerquellen können den einzelnen Phasen des Prognoseprozesses in Abbildung 3.23 zugeordnet werden und so in der Absatzprognose berücksichtigt werden.²

Fehler bei der Prognoseerstellung entstehen durch eine fehlerhafte Selektion geeigneter Erklärungs- und Prognosemodelle sowie passender Prognoseverfahren. Die Modellbildung muss interdependent mit der fehlerhaften oder unvollständigen Identifikation der Muster und Komponenten bei zeitreihenanalytischer Modellierung oder mit der fehlerhaften Identifikation von kausalen Wirkungseinflüssen zwischen Zeitreihen bei kausalanalytischer Modellierung erfolgen. So entstehen vielfach Fehler durch die Anwendung eines Prognoseverfahrens, das

¹ Auch bei korrekter Identifikation eines Erklärungsmodells mit allen relevanten Modellkomponenten und unter fehlerfreier Schätzung durch ein adäquates Verfahren verbleibt eine stochastische Residualstreuung, die den minimal erzielbaren Prognosefehler des Prognoseobjektes durch Zufälligkeiten auf dem Absatzmarkt bezeichnet. Da diese jedoch unbekannt ist, gestaltet sich die Suche nach Verbesserungspotenzialen im Prognoseprozess sowie durch neuartige Verfahren als iterativer Prozess schrittweiser Verbesserungen.

² Fehlerquellen in der Prognosevorbereitung beruhen überwiegend auf einer fehlerhaften Formulierung des Entscheidungsproblems und der Determinanten des Prognoseproblems, also der Prognosehorizonte, Frequenzen und Zielsetzungen. Da die Phasen der Bestimmung des Entscheidungsproblems sowie der Informationssammlung und -aufbereitung für alle Prognoseverfahren einheitlich erfolgen, können sie kein Verfahren selektiv begünstigen. Somit werden sie in der Betrachtung vernachlässigt.

die Struktur des eigentlichen Erklärungsmodells gar nicht abbilden kann, z. B. durch ein Verfahren des S.ES auf eine saisonale Zeitreihe oder des L.ES zur Abbildung von Trends auf eine stationäre Zeitreihe mit Niveauverschiebungen. Vor diesem Hintergrund ist auch die Abbildung nichtlinearer Zusammenhänge des Erklärungsmodells durch lineare Verfahren zu nennen, die durch die Verwendung geeigneter Vorverarbeitungen oder nichtlinearer Verfahren systematische Fehler hervorruft. Fehlerquellen in der verfahrensspezifischen Datenvorverarbeitung entstehen dagegen, wenn zwar die Struktur des Erklärungsmodells korrekt identifiziert wird, es aber an einer geeigneten Abbildung durch obligatorische oder optionale Vorbearbeitung mangelt, z. B. bei fehlender Integration instationärer Zeitreihen und dadurch verzerrten Parameterschätzern der ARIMA und der Regressionsverfahren.

Systematische Fehler in der Beurteilung der Verfahren entstehen durch die Verwendung suboptimaler Fehlermaße. Diese können durch die Berücksichtigung der Empfehlungen zur Auswahl valider Fehlermaße in der Durchführung empirischer Vergleichsstudien (vgl. Abschnitt 3.3.3) vermieden werden. Können diese nicht vollständig umgesetzt werden, z. B. bei einer Beurteilung aufgrund nur eines Prognoseursprungs, muss dies in der Beurteilung der Reliabilität der Ergebnisse berücksichtigt werden. Die Beurteilung der Prognosegenauigkeit ruft auch dann Fehler hervor, wenn tatsächlich die Entscheidungsqualität des Verfahren anhand asymmetrischer Entscheidungskosten beurteilt werden müsste.

Die prozessorientierte Betrachtung der Bedarfsprognose verdeutlicht die Komplexität der Entscheidungen, die Beurteilung unterschiedlicher Verfahrensrisiken im Prognoseprozess und die Notwendigkeit eines systematischen Vergleichs von Prognoseverfahren für jede Anwendung und jeden Datensatz. Daraus werden nachfolgend Erkenntnisse abgeleitet, inwiefern das Prognoseverfahren der Neuronalen Netze zu einer systematischen Vermeidung von Fehlern in der Spezifizierung von Modell und Verfahren beitragen kann. Diese werden im nachfolgenden Kapitel analysiert.

4 Neuronale Netze als Instrument der Prognose

In Kapitel 4 werden die Prognoseverfahren der künstlichen Neuronalen Netze (NN) als Klasse ämathematisch-statistischer Verfahren vorgestellt. In den Hauptgliederungspunkten werden zu diesem Zweck die methodischen Grundlagen und die Prinzipien der Informationsverarbeitung in den Phasen der Modellbildung, der Parametrisierung und der Anwendung analysiert.¹

Die Modellierung von NN bietet erhebliche Freiheitsgrade in der mathematischen Formulierung der Signalverarbeitung in einzelnen Neuronen, der Informationsverarbeitung in Netzwerken von Neuronen und der Parametrisierung durch Lernverfahren. Daher werden diese zunächst eingehend und losgelöst vom Anwendungsbezug der Prognose untersucht, um die Funktion und Bedeutung der Zielfunktionen im Training von NN herauszustellen. Die systematische Analyse der NN, ihrer historischen Entwicklung, ihres biologischen Vorbildes und ihrer technischen Realisierung verdeutlicht die Unabhängigkeit der Zielfunktionen im Training von NN von der Zielsetzung, der Architektur und den Lernverfahren. Dies erlaubt nachfolgend eine Erweiterung der Theorie von NN auf alternative Zielfunktionen hinsichtlich betrieblicher Kosten und somit auf die simultane Prognose und Disposition durch NN.

In einem zweiten Schritt wird die Analyse auf die Anwendung von NN zur betrieblichen Absatzprognose konzentriert. Zunächst wird dazu die theoretische Eignung von NN zur Approximation, Generalisierung und Vorhersage aus Zeitreihen von verrauschten Daten untersucht. Die Vielzahl von Anwendungen von NN zur Prognose erfordert eine strukturierte Literaturanalyse zur Beschränkung auf relevante Anwendungen der Regression und Absatzprognose. Abschließend werden typische Architekturen von NN zur Absatzprognose diskutiert und die Informationsverarbeitung von Neuronen und Netzwerken von Neuronen im Vergleich zu statistischen Prognoseverfahren erörtert.

4.1 Methodische Grundlagen von Neuronalen Netzen

4.1.1 Definition und Abgrenzung

„A neural network is a system composed of many simple processing elements operating in parallel whose function is determined by networks structure, connection strengths, and the processing performed at computing elements or nodes. ... Neural network architectures are inspired by the architecture of biological nervous systems, which use many simple processing elements operating in parallel to obtain high computation rates” [DARPA 1989, S. 60].

Der allgemeine Begriff der Neuronalen Netze bezeichnet somit kein Prognoseverfahren, sondern vielmehr ein interdisziplinäres Forschungsgebiet von Systemen, Theorien, Modellen und Verfahren hinsichtlich biologischer Neuronen und Neuronenverbände. „The field of neural networks is now being investigated by ... biologists, psychologists, cognitive scientists, computer scientists, mathematicians, engineers, physicists, philosophers, psychiatrists and psychotherapists, and social scientists” [Levine 2000, S. xi]. Neurobiologie, Neurologie und Neurolinguistik nutzen die Prinzipien der Modellierung von Neuronalen Netzen, um die theoretischen Grundlagen der Informationsverarbeitung im Gehirn von Lebewesen zu klären [Churchland 1986; Ramsey et al. 1991]. Psychologen und Kognitionswissenschaftler hoffen durch die Analyse von Neuronalen Netzen auf ein besseres Verständnis kognitiver Phänomene, während Sozialwissenschaftler auf ein besseres Verständnis menschlicher Entscheidungsprozesse und Philosophen gar auf Einsichten in epistemologische Fragen des Geistes und des Verstandes zielen [Levine 2000, S. xi]. Auch die Mathematik, die Statistik

¹ Die informationstechnische Realisierung von NN erfolgt computergestützt in Form von speziellen Hardware- oder Softwaresimulatoren, auf die hier nicht näher eingegangen wird.

und die Physik analysieren theoretische Probleme dynamischer Systeme in computergestützten Modellen künstlicher Neuronaler Netze. „Over the years, the terms cybernetics, neural networks, brain theory, artificial intelligence, and neural modelling (among others) have been applied to this general area” [Schwartz 1990b, S. ix]. Die vorliegende Arbeit vernachlässigt nachfolgend allerdings jene Disziplinen, deren Fokus auf der Analyse kognitiver Funktionen und biologischer neuronaler Strukturen beruht.¹

Nach LEVINE [2000, S. 2] wird die industrielle Anwendung von Neuronalen Netzen durch die Bezeichnung als „künstliche neuronale Netze“ (engl. *artificial neural networks*) abgegrenzt, während eine allgemeine Bezeichnung als „Neuronale Netze“ (engl. *neural networks*) sowohl künstliche als auch biologische, theoretische und angewandte Modelle umfasst.² HECHT-NIELSEN [1990] unterteilt die Erforschung von künstlichen NN in drei Untersuchungsgebiete: Architektur und Theorie, Implementierung sowie Anwendung. Die theoretische Forschung ist beschränkt auf das Studium und die Entwicklung formaler, implementierungsunabhängiger mathematischer Definitionen neuronaler Netzwerkarchitekturen, die Entwicklung von Theorien zur Funktions- und Arbeitsweise dieser Architekturen und ihre experimentelle Überprüfung an einfachen Anwendungsproblemen.³ [Hecht-Nielsen 1990, S. 11]. Der Bereich der Implementierung betrachtet die Umsetzung der mathematischen Architekturen in Software einerseits und physischer Hardware durch Neurocomputer andererseits, als Voraussetzung für die Anwendung von NN in praktischen Problemstellungen [Hecht-Nielsen 1990, S. 11]. Die Erkenntnisse aus der theoretischen Erforschung von NN liefern die Basis für eine Vielzahl von Anwendungen der Mustererkennung, Signalverarbeitung, technischen Regelung und Kontrolle, Kategorisierung und Prognose in der Praxis [Haykin 1999, S. XII].

Die anwendungsorientierte Forschung, insbesondere in den Ingenieurwissenschaften und der Informatik, untersucht die theoretischen Prinzipien der neuronalen Informationsverarbeitung als Inspiration zum Design neuer Apparate, Modelle oder Methoden. „A *neural network* is a machine that is designed to *model* the way in which the brain performs a particular task or function of interest; the network is usually implemented by using electronic components or is simulated in software on a digital computer“ [Haykin 1999, S. 2]. Dies erlaubt auch eine Definition von künstlichen Neuronalen Netzen als „parallel, distributed, adaptive information processing systems that develop information processing capabilities in response to exposure to an information environment“ [Hecht-Nielsen 1990, S. 2]. Diese informationsverarbeitenden Systeme bestehen „aus einer großen Anzahl einfacher Einheiten (Zellen, Neuronen) ...“, die

¹ Dies umfasst Fragestellungen der Neuroscience als eigenständiger wissenschaftlicher Disziplin mit dem Ziel eines besseren Verständnisses von Gehirn und Verstand [Hecht-Nielsen 1990, S. 12], der Computational Neuroscience und der Bio-Informatik im Schnittpunkt von Neurobiologie und Informatik [Churchland 1989; Schwartz 1990a; Feng 2004; Hecht-Nielsen und McKenna 2003; Perkel 1990], des HEBB'schen Konnektionismus zur Untersuchung kognitiver Funktionen in neurobiologisch inspirierten Teilmodellen des Gehirns [Levine 2000, S. 2; Hebb 1949] und des Neurocomputing im Schnittpunkt von Neurobiologie, Kognitions- und Ingenieurwissenschaften, auch als „neural computation“ oder „brainlike computation“ bezeichnet, mit dem Ziel, Computer mit gehirnähnlichen Strukturen und Funktionalitäten zu entwerfen [Anderson 1995, S. vii].

² Aus der abgekürzten Englischen Bezeichnung als ‚neural nets‘ wird die Deutsche Bezeichnung der ‚Neuronalen Netze‘ abgeleitet, wobei die irreführende Bezeichnung als ‚Netz‘ im Sinne des ‚Netzwerk‘ zu interpretieren ist. In Anlehnung an den etablierten Sprachgebrauch wird nachfolgend auf die verbreitete Bezeichnung als ‚Neuronale Netze‘ zurückgegriffen. Alternativ werden sie nach HAYKIN [1999, S. 2] auch als ‚connectionist networks‘ sowie nach Rumelhart et al. [1986] auch als ‚parallel distributed processing systems‘ in Anlehnung an ihre Struktur der Verarbeitungseinheiten bezeichnet: „A neural network is a parallel, distributed information processing structure consisting of processing elements (which can possess a local memory and can carry out localized information processing operations) interconnected via unidirectional signal channels called connections“ [Hecht-Nielsen 1990, S. 2-3]. Ein NN entspricht somit einer Multiple-Instruction-Multiple-Data-(MIMD) Struktur der parallelen Informationsverarbeitung.

³ Er bezeichnet diese als „toy problems“ im Sinne von „problems that are simple to state and that require no significant application domain knowledge to understand and appreciate“ [Hecht-Nielsen 1990, S. 11]. Die Abgrenzung theoretischer und anwendungsbezogener Forschung erfolgt somit durch das empirische Wissen der Anwendungsdomäne.

sich Informationen in Form der Aktivierung der Zellen über gerichtete Verbindungen (connections, links) zusenden“ [Zell 2000, S. 23].¹

In der Betriebswirtschaftslehre wird die Anwendung von künstlichen Neuronen Netze in den Teildisziplinen Wirtschaftsinformatik, Finanzwissenschaften und Operational Research als Verfahren der Informationsverarbeitung zur betrieblichen Entscheidungsunterstützung untersucht.² Vor dem Hintergrund betriebswirtschaftlicher Entscheidungsprobleme und der Prognose hat sich keine eindeutige Definition für Neuronale Netze etabliert. Die nachfolgende Analyse betrachtet künstliche neuronale Netze (NN) als Klasse mathematisch-statistischer Verfahren, deren Funktionsweise dem Vorbild der neurobiologischen Informationsverarbeitung in Nervenzellen und Nervenverbänden des Menschen nachempfunden wurde und die aufgrund ihrer Fähigkeiten zur nichtlinearen, datengetriebenen und semi-parametrischen Informationsverarbeitung zur Prognose in der betrieblichen Entscheidungsunterstützung verwendet werden können.³

4.1.2 Motivation zur Analyse von Neuronalen Netzen

„Historically, the interest in neural networks has two roots: the desire to understand the principles on which the human brain works, and the wish to build machines that are capable of performing complex tasks for which the sequentially operating, programmable computers conceived by Babbage and von Neumann are not well suited“ [Müller et al. 1995, S. 18]. Die Untersuchung und Anwendung von NN begründet sich somit in der Leistungsfähigkeit und Funktionsweise von biologischen neuronalen Zellen und Zellverbänden.

Das zentrale Nervensystem des Menschen ist eine bemerkenswert effiziente, effektive und robuste Struktur zur Informationsverarbeitung. Es besteht aus Milliarden individueller Nervenzellen, die in einem engen Netzwerk verbunden sind und auf die nahezu alle menschlichen Handlungen zurückgeführt werden können [Gupta und Rao 1994a, S. 6]. Es besitzt die Fähigkeit zur Wahrnehmung, Intuition und Inferenz aus verrauschten Beobachtungen, den Grundlagen des abstrakten Denkens in Hierarchien [Müller et al. 1995, S. 22]. Ferner kann es bestimmte Informationsverarbeitungen, z. B. Mustererkennung und Informationsextraktion aus verrauschten Daten, um ein Vielfaches schneller berechnen als digitale Computer [Anderson 1995, S. vii; Müller et al. 1995, S. 22].⁴ Derartige Probleme der Informationsverarbeitung sind nicht durch herkömmliche Ansätze der Computertechnologie lösbar, so dass einer naturanalogen Modellierung der Probleme durch Verfahren der NN eine besondere Bedeutung zukommt.

¹ Alternativ erlauben NN auch eine Interpretation im Kontext der Systemtheorie als „gerichteter, gewichteter Graph ...“, wobei die Kanten die gewichteten Verbindungen zwischen den Neuronen darstellen“ [Zell 2000, S. 72] und für jeden Knoten ein Aktivierungszustand als Funktion der eingehenden Gewichte und Zustände verbundener Knoten definiert ist [Müller et al. 1995, S. 13].

² Die Fragestellungen bezüglich der biologischen Plausibilität und Kongruenz der internen Repräsentation und Verarbeitung der Informationen treten bei der Betrachtung von NN als Methode der betriebswirtschaftlichen Entscheidungsfindung in den Hintergrund. Sie werden daher nachfolgend vernachlässigt.

³ Nachfolgend werden lediglich jene Entwicklungen hervorgehoben, die für die Anwendung ausgewählter Architekturen in der Prognose relevant erscheinen. Eine detaillierte Darstellung und Diskussion historischer Netzwerk-Konzepte findet sich bei [Hecht-Nielsen 1990]. Eine Sammlung aller wichtigen Veröffentlichungen geben ANDERSON et al. in [Anderson und Rosenfeld 1988; Anderson et al. 1988]. Details zu den Ursprüngen der Neurosciences finden sich in [Rall 1990].

⁴ So führt das menschliche Gehirn routinemäßig die visuelle Erkennung, z. B. von bekannten Gesichtern in einer unbekanntem Umgebung, in 100 bis 200 Millisekunden durch, während Aufgaben mit deutlich geringerer Komplexität auf konventionellen Computern mehrere Tage benötigen [Haykin 1999, S. 1]. Selbst das nur pflaumengroße Neuronale Netzwerk der Fledermaus ist in der Lage, nach der Ortung beweglicher Objekte aus Distanz, relativer Geschwindigkeit, Größe und räumlichen Eigenarten des Ziels die relevanten Informationen zu extrahieren und durch Berechnung des Azimuts, des Winkels und der Höhe das Ziel zu verfolgen und zu fangen, was durch herkömmliche Technik nur schwer zu erreichen wäre [Haykin 1999, S. 1]. Somit ist der Wunsch nach einem System zur Informationsverarbeitung auf dem Niveau eines Wirbeltiers, einer Maus oder auch nur eines Insekts eine noch unlösbare Herausforderung an gegenwärtige Architekturen.

Das das Nervensystem benötigt bei 1% bis 2% des Körpergewichtes fast 25% der Körperenergie [Anderson 1995], Dennoch ist das Gehirn mit einem Energieverbrauch von 10^{-16} Joule pro Rechenoperation im Vergleich zu den etwa 10^{-6} Joule eines Computers außerordentlich effizient. Zudem wiegt es nur ca. 1500 Gramm, benötigt keine spezifische Kühlung und erfüllt neben der Verarbeitung zudem die Aufgabe der Speicherung und des Abrufens der Informationen [Haykin 1999, S. 6].¹ Exakte numerische Berechnungen können dagegen durch Computer um ein Vielfaches effizienter, schneller und genauer durchgeführt werden als durch das menschliche Gehirn. Biologische Neuronen sind mit einer Schaltzeit im Millisekunden-Bereich (10^{-3} Sekunden) deutlich langsamer als silikonbasierte Mikrochips mit Schaltzeiten von Nanosekunden (10^{-9} Sekunden) [Haykin 1999, S. 6; Churchland et al. 1990, S. 47]. Daher sind die Fähigkeiten von NN in der Informationsverarbeitung auf die hohe Anzahl einfacher neuronaler Verarbeitungseinheiten und deren massive parallele Vernetzung zurückzuführen [Haykin 1999, S. 6; Strecker 1997, S. 7]. Die Simulation dieser biologisch inspirierten Strukturen auf Computern verspricht so Ansätze zur effizienten Lösung einzelner Probleme der Informationsverarbeitung.

Andererseits besitzen natürliche Nervensysteme auch negative Eigenschaften: Biologische NN sind langsam, ungenau, machen fehlerhafte Generalisierungen, sind voreingenommen bzw. befangen und üblicherweise unfähig, ihre Handlungen explizit zu erklären [Anderson 1995, S. vii]. “We should ... expect, that brains (or neural networks) cannot compute everything, but only arrive at reasonable solutions for a small but useful set of problems” [Anderson 1995, S. 4].² NN können vor allem dort angewendet werden, wo ein genau definierter Kompromiss zwischen effektiver, verlässlicher Problemlösung und effizienter Performanz angestrebt wird [Anderson 1995, S. x]. Die Problemstellungen vieler betriebswirtschaftlicher Entscheidungen ähneln denen, die von den biologischen NN von Tieren oder Menschen effizient gelöst werden: Assoziation, Prognose, Kategorisierung, Generalisierung sowie schnelle Reaktion und Inferenz aus verrauschten, unvollständigen Daten [Anderson 1995, S. 3]. Die Prognose unbekannter Realisationen in der Zukunft auf der Basis komplexer, verrauschter Muster in historischen Beobachtungen aus Quellen unterschiedlicher Wichtigkeit und Zuverlässigkeit repräsentiert eine potenzielle Anwendungsdomäne von NN, in der die numerische Genauigkeit durch die Anforderungen der stochastischen Struktur der Entscheidungssituation überlagert wird. Die numerische Effizienz stellt vielfach keinen Engpass der Problemlösung dar, wie durch den Einsatz menschlicher Experten in subjektiven Prognosemodellen belegt wird.

Die im Folgenden zur Prognose eingesetzten NN haben nur noch wenig mit dem biologischen Vorbild der NN gemein. Nach ZELL sind jedoch „nur diese stark vereinfachten Modelle derzeit einer genauen theoretischen Analyse zugänglich“ [Zell 2000, S. 35]. Der hohe Abstraktionsgrad der künstlichen Modelle erlaubt folglich Untersuchungen zur Ähnlichkeit der NN mit statistischen Verfahren, nicht aber Rückschlüsse auf die Funktionsweise biologischer NN. Entsprechend beschränkt sich auch die nachfolgende Diskussion der historischen Entwicklung und der biologischen Grundlagen von NN auf vereinfachte Modelle.

4.1.3 Historische Entwicklung

Die interdisziplinäre Erforschung der NN reicht bis zu den Anfängen der computergestützten Informationsverarbeitung zurück. Die Biophysiker MCCULLOCH und PITTS entwickelten bereits 1943 ein erstes formales, logikbasiertes Modell der Informationsverarbeitung in

¹ Dies repräsentiert gleichzeitig ein weiteres Vorbild für alternative Hardwarelösungen der Informatik.

² NN können und sollen also nicht alle Formen der Informationsverarbeitung ablösen, sondern lediglich in jenen Problemstellungen auf ihre Eignung geprüft werden, in denen die Dominanz menschlicher Experten und die Prinzipien neuronaler Informationsverarbeitung ihren Einsatz rechtfertigen, z. B. in Management-Support- oder in Forecasting-Support-Systemen.

biologischen neuronalen Netzen [McCulloch und Pitts 1943]. Sie zeigten, dass ein Netzwerk von Neuronen aus einfachen Verarbeitungseinheiten mit einer Schwellwertlogik und binären Ausgaben von $\{0,1\}$ jede finite logische Aussage realisieren kann und somit immense Rechenkapazität besitzt [Anderson 1995, S. 51]. Die Mächtigkeit dieser strukturellen Organisation inspirierte später Forscher unterschiedlicher Disziplinen: VON NEUMANN [1943, S. 82 ff.; 1958] zum Entwurf des allgemeingültigen digitalen Computermodells, WIENER [1948; 1965] zur Analogie von Computern und Gehirnen in „Cybernetics“, MINSKY [1968; 1973] zur Entwicklung des Forschungsfeldes makroskopischer und symbolischer künstlicher Intelligenz sowie ROSENBLATT [1958] zur Entwicklung eines ersten künstlichen NN.¹

Bei dem Versuch, neurophysiologische und psychologische Erkenntnisse zu verbinden, postulierte HEBB 1949 ein universelles Lernkonzept: Lernen begriff er als strukturelle Veränderung biologischer Neuronenverbindungen im Langzeitgedächtnis [Levine 2000, S. 17]. Diese Hypothese einer formalen Lernregel liefert bis heute die Grundlage vieler Lernverfahren für künstliche NN [Zell 2000, S. 28]. Eine erste computergestützte Software-Simulation, und somit die erste veröffentlichte Anwendung künstlicher NN, gelang 1956 den Wissenschaftlern der IBM Research Laboratories ROCHESTER, HOLLAND, HAIBT und DUDA, als sie die HEBB'schen Theorien einer detaillierten Evaluation unterzogen. [Hecht-Nielsen 1990, S. 15].² Der Psychologe ROSENBLATT entwickelte 1958 das Perceptron als Klasse lernender Systeme zur optischen Mustererkennung an einer künstlichen Retina, u.a. zur Erkennung handgeschriebener Zeichen [Rosenblatt 1958]. Es erlaubte eine erfolgreiche Mustererkennung und zeigte ein komplexes adaptives Verhalten, Ansätze zu einem assoziativen Lernen, Fehlerrobustheit bei Beschädigung einzelner Verarbeitungselemente und Plastizität nach Beschädigung oder Zerstörung durch erneutes Trainieren. Zudem konnte das Perceptron Muster trotz Zufallsschwankungen in den Eingabedaten klassifizieren und eingeschränkte Generalisierungen durchführen [Levine 2000, S. 17-23; Neuralware 1991, S. 36]. Diese Eigenschaften führten zu einer Forschungseuphorie für NN, trotz des Problems der manuellen bzw. zufälligen Parameterfindung. „Work .. progressed using the three techniques used for research on neural networks today: mathematical analysis, computer simulation, and special purpose hardware“ [Anderson und Rosenfeld 1988, S. 135], vor allem in den anwendungsorientierten Forschungszweigen der Ingenieurwissenschaften und der Physik. 1960 stellten WIDROW und HOFF [Widrow und Hoff 1960] ein dem Perceptron verwandtes Modell von adaptiven linearen Neuronen (ADALINE)³ sowie eine schnell konvergierende Lernregel vor, die als mathematische Grundlage zur Fehlerberechnung mittels einer quadratischen Zielfunktion noch immer in zahlreichen Verfahren des überwachten Lernens eingesetzt wird (Widrow und Hoff, 1960). Ihre Lernregel war verfahrenstechnisch motiviert und nicht durch biologische Prozesse erklärbar. Somit fand eine Abkehr von einer biologisch plausiblen Modellierung hin zu einer anwendungsorientierten Modellbildung statt (vgl. ausführlich 4.3.2.1). Ihre Ausführungen fanden insbesondere im Bereich der Elektrotechnik, der Signalverarbeitung und der Telekommunikation erste kommerzielle Anwendungen und Erfolge [Levine 2000, S. 24].

Die Forschungseuphorie, gestützt von dem Optimismus, die grundlegenden Prinzipien der selbstlernenden, „intelligenten“ Systeme entdeckt zu haben, wurde zunächst durch intensive

¹ Die grundlegenden Begriffe „artificial intelligence“ und „cognitive science“ wurden ebenso wie der Begriff „artificial neural networks“ 1956 auf der Dartmouth-Konferenz erstmalig verwendet.

² Sie präsentierten ihre Arbeit im „Dartmouth Summer Research Project on Artificial Intelligence“, der ersten Konferenz, die den offiziellen Beginn der Forschungsfelder Künstliche Intelligenz und NN markierte. ROSENBLATT'S Forschergruppe am MIT, die Erfinder des Perceptrons, baute erst 1958 den ersten erfolgreichen Neurocomputer (Mark 1 Perceptron) für Mustererkennungsprobleme. MINSKY hatte bereits 1951 mit dem „Snark“ einen Neurocomputer entwickelt, jedoch nicht praktisch eingesetzt [Hecht-Nielsen 1990, S. 15].

³ Die Bezeichnung als „ADaptiven LInearen Neuronen“ des ADALINE wurde im Zuge der wachsenden Kritik an NN in „ADaptive LINear Element“ umbenannt.

Forschungsförderung unterstützt [Zell 2000, S. 29-30]. In Ermangelung praktischer Erfolge der Modelle sank die Popularität der NN jedoch zum Ende der 60er Jahre. Die Forschungsanstrengungen wurden 1969 mit einer mathematischen Analyse der theoretischen und praktischen Grenzen des Perceptrons durch MINSKY und PAPERT [1969] beendet.¹ Sie zeigten u. a., dass Perceptrons selbst einfache nichtlineare Probleme wie z. B. das logische „exklusive oder“ (XOR) nicht repräsentieren können [Minsky und Papert 1969, S. 157-168]. Dies stellte den theoretischen Nutzen und die praktische Anwendbarkeit von NN grundsätzlich in Frage. MINSKY und PAPERT [1969, S. 232] erweiterten ihre pessimistische Einschätzung auch auf zukünftige Entwicklungen mehrlagiger Perceptrons, die die angesprochenen Probleme theoretisch hätten lösen können, für die zum damaligen Zeitpunkt aber noch kein Lernverfahren existierte. „It was the suggestion that learning procedures for complex networks could never be developed that contributed to the loss of interest“ [Zipser und Rumelhart 1990]. Dies führte in den folgenden Jahren nahezu zum Stillstand der Forschung an NN.²

Erst 1986 gelang es einer Forschergruppe um RUMELHART und MCCLELLAND, die Kritik von MINSKY und PAPERT [1969] zu widerlegen. Sie präsentierten eine Erweiterung des WIDROW-HOFF-Lernverfahrens namens „error back-propagation“, durch das Perceptrons mit mehreren versteckten Schichten, so genannte Multilayer Perceptrons, robust lernen konnten [Rumelhart et al. 1994]. Seit diesem Durchbruch erhält das Forschungsfeld der NN wieder wachsende Aufmerksamkeit in Wissenschaft und Praxis. Gestützt durch Fortschritte in der Computertechnik und in der Theorie von NN ist dies vor allem auf die wachsende Komplexität technischer Problemstellungen in industriellen Anwendungsbereichen, insbesondere „in applied areas such as knowledge processing, robotic control, pattern classification, speech synthesis and recognition, and machine vision“ [Levine 2000, S. 3] und die Weiterentwicklung herkömmlicher Computersysteme und traditioneller Ansätze der künstlichen Intelligenz (KI) zur Lösung dieser Problemstellungen zurückzuführen. Seitdem finden NN zunehmend Anwendung auch bei betriebswirtschaftlichen Problemstellungen, initiiert durch Anwendung in finanzwissenschaftlichen Bereichen, die im Anschluss an die Darstellung der verfahrenstechnischen Grundlagen näher vorgestellt werden.³

4.2 Aufbau und Funktionsweise von Neuronalen Netzen

4.2.1 Biologisches Vorbild

Die Modelle der künstlichen NN wurden in Analogie zu Neuronenverbänden des menschlichen Zentralen Nervensystems (ZNS) entworfen. Trotz der fortschreitenden Distanz zu neurophysiologisch korrekten Modellen der biologischen Realität bestehen Analogien zwischen biologischen und künstlichen NN, die durch die Struktur der Modelle und die verwendete Terminologie die Basis für das Verständnis computergestützter Modelle legen. Nachfolgend wird somit die biologische Morphologie von NN vereinfacht skizziert und anschließend die technische Realisation der biologischen Grundlagen in NN dargestellt.⁴

Das ZNS des Menschen enthält ca. 10 bis 100 Milliarden Nervenzellen, sog. Neuronen, unterschiedlicher Form, Größe und Funktion in unterschiedlichen Regionen des Gehirns

¹ Beide waren prominente Vertreter der konkurrierenden symbolischen künstlichen Intelligenz.

² Diese Forschung legte jedoch die theoretischen Grundlagen für viele gegenwärtige Forschungsgebiete zu NN, die im Bereich der vorliegenden Arbeit allerdings nicht relevant erscheinen. Zu nennen sind insbesondere KOHONENS Arbeiten zu selbstorganisierenden Karten, GROSSBERGS neuronale Modelle der adaptiven Resonanztheorie und HOPFIELDS sog. Hopfield-Netze. Für einen Überblick vgl. [Zell 2000].

³ Die Darstellung der theoretischen Grundlagen beschränkt sich dabei auf jene Aspekte, die zur Anwendung von NN bei der Prognose im Sinne der Punktprognose relevant sind. Insbesondere werden jene Grundlagen betrachtet, die für die Ableitung und Legitimation von asymmetrischen Zielfunktionen notwendig sind.

⁴ Eine Diskussion der intra- und interzellulärer, biochemischer und biophysischer Vorgänge von NN sowie eine detaillierte Diskussion neurobiologischer Grundlagen geben [Anderson 1995; Zell 2000].

[Haykin 1999, S. 7; Anderson 1995].¹ Ein Neuron besteht aus einem Zellkörper, dem Soma, baumartigen Eingangsstrukturen, den Dendriten, und einer langen Nervenfasern zur Reizweiterleitung, dem Axon, an dessen Ende sich verzweigte Verdickungen zur Signalübertragung, die Synapsen, befinden. Jedes der 10^{10} bis 10^{11} Neuronen ist über seine ca. 10^3 bis 10^4 Synapsen mit den Dendriten nachfolgender Neuronen verbunden, wodurch ein teilverbundenes Netzwerk von ca. 10^{15} Verbindungen im Gehirn entsteht [Anderson 1995, S. 10; Haykin 1999, S. 7; Müller et al. 1995]. Eine schematische Darstellung ist in Abbildung 4.1 dargestellt.

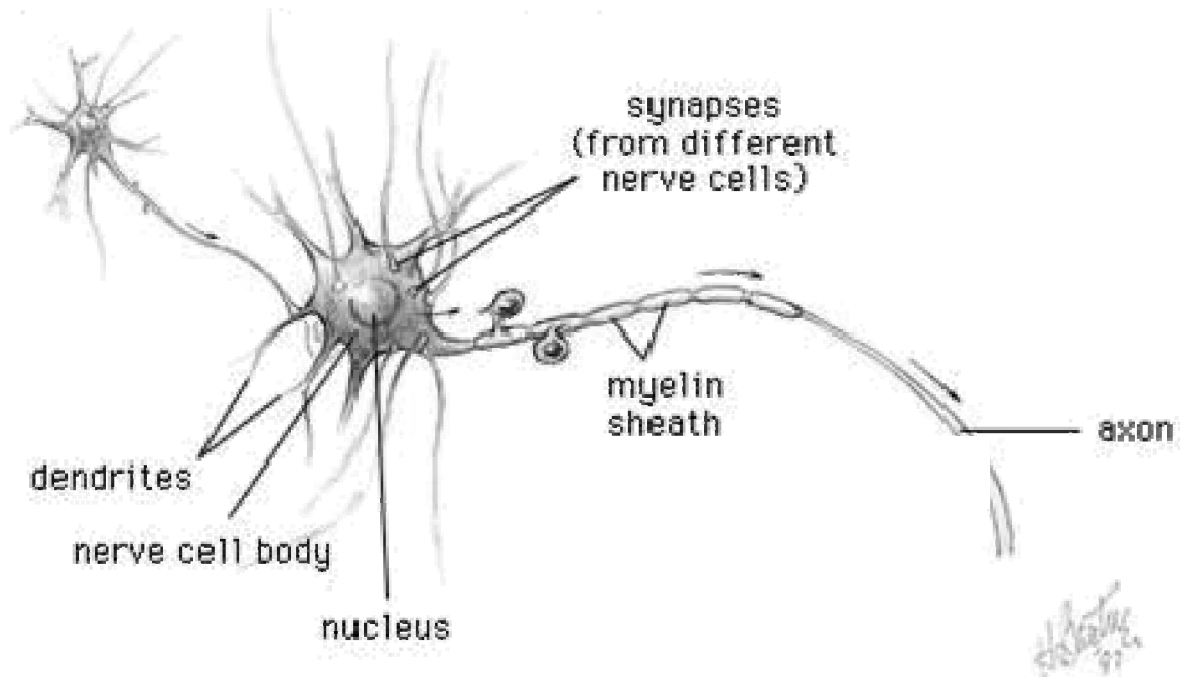


Abbildung 4.1: Biologisches Neuron im Neuronenverbund

Biologische Neuronen und daraus kombinierte Netzwerke bilden komplexe Strukturen, deren Informationsverarbeitung, in Abbildung 4.2 dargestellt, in beliebigen Abstraktionsgraden bis hin zur Interaktion einzelner Moleküle modelliert werden kann [Haykin 1999, S. 9].

Die anwendungsbezogene Modellierung künstlicher Neuronen und lokaler Neuroneneinheiten erfolgt auf einem der Problemstellung angemessenen Abstraktionsniveau einfacher, neuronaler Teilnetze, um inhärente Funktionen der Informationsverarbeitung als Systemeigenschaften abzubilden und die Komplexität zu verringern [Anderson 1995, S. 48]. Das ZNS führt die Informationsverarbeitung in einem dreigeteilten System eines Stimulus-Organismus-Response-Modells durch, in dem Neuronen nach ihrer Funktion und Position unterschieden werden [Haykin 1999, S. 6]. Sensorische Neuronen nehmen Reize der Umwelt als Stimuli bzw. Input in Rezeptoren auf, kodieren sie zur Informationsverarbeitung in elektrische Signale und leiten sie an Zwischenneuronen des ZNS weiter, die ausschließlich die Nervensignale verarbeiten und an andere Neuronen weiterleiten. Die verarbeiteten Nervensignale werden abschließend an motorische Neuronen, sog. Effektoren, weitergeleitet, die mit Muskelzellen oder Organen verbunden sind und somit aus den Aktivierungen des ZNS eine Handlung als Output bzw. Systemreaktion weiterleiten. In der Verarbeitung der elektrischen

¹ Obwohl die Funktionsweise des Neokortex nicht vollständig wissenschaftlich geklärt ist, wird angenommen, dass Neuronen die elementaren informationsverarbeitenden Einheiten in ZNS und Gehirn darstellen und dass ihre Zusammenarbeit die Informationsverarbeitung des Gehirns bestimmt [Anderson 1995, S. 2; Haykin 1999, S. 10]. Anderen Zellklassen des Gehirns, z. B. den Glia-Zellen, wird allerdings vereinzelt ebenfalls eine Rolle in der Informationsverarbeitung zugesprochen, vgl. [Anderson 1995].

Nervensignale erhält jedes Zwischenneuron Eingangssignale von bis zu 10^3 Vorgängerzellen und leitet Signale an ebenso viele Nachfolgerzellen weiter, was zu einer komplexen, massiv parallelen Informationsverarbeitung führt.

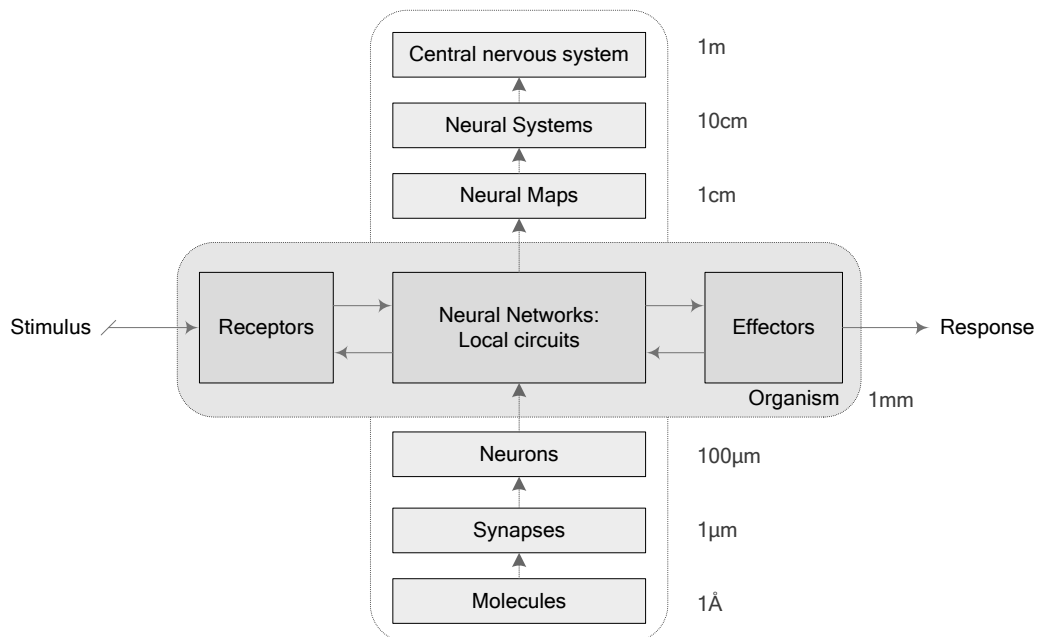


Abbildung 4.2: Zentrales Nervensystem und die Ebenen struktureller Organisation¹

Die Impuls- bzw. Informationsverarbeitung im Neuron erfolgt unidirektional durch elektrochemische Vorgänge. Die Dendriten nehmen die eingehenden Nervenimpulse vorgelagerter Neuronen in Form positiver oder negativer Veränderungen des elektrischen Membranpotenzials der Nervenzelle auf. Die Eingangssignale werden im Soma zu einem Aktionspotenzial kumuliert, das den Grad der Aktivierung des Neurons angibt. Übersteigt die kombinierte Veränderung des Membranpotenzials einen bestimmten Schwellwert mittels einer anhaltenden, kombinierten Reizung des Neurons durch seine Vorgängerneuronen, entsteht im Soma des Neurons ein elektrischer Impuls: Das Neuron wird aktiv, es „feuert“. Dieses Aktionspotenzial wandert als Fluss geladener Ionen entlang des Axons in die Synapsen, wo es die Freisetzung von chemischen Neurotransmittern bewirkt.² Diese übertragen den Nervenimpuls auf die Dendriten nachfolgender Neuronen und bewirken somit wiederum eine Veränderung des elektrischen Potentials in allen verbundenen Nachfolgerzellen [Anderson 1995, S. 7-35; Zell 2000, S. 35-68]. Je nach Stärke und Art der Synapse bewirkt der Nervenimpuls entweder eine Potenzialerhöhung oder eine Potenzialerniedrigung in der empfangenden Nachfolgerzelle, im Sinne einer verstärkenden („exzitatorischen“) oder einer hemmenden („inhibitorischen“) Wirkung auf die Reizweiterleitung der Nachfolgerzelle [Zell 2000, S. 45].

Ausgehend von der umstrittenen Überzeugung, dass beim Menschen nach Abschluss der Embryonalphase Neuronen nur absterben, nicht aber neu gebildet werden können, wurden alle Vorgänge des Lernens, der Speicherung von Erfahrungen und der Informationsverarbeitung auf Veränderungen der Struktur und Stärke der synaptischen Verbindungen zwischen Neuronen, der sog. synaptischen Plastizität, zurückgeführt [Anderson 1995, S. 37-61; Gupta und Rao 1994a, S. 6; Zell 2000, S. 45]. Da die Synapsen die elementaren funktionalen

¹ Erweiterte Darstellung in Anlehnung an [Churchland et al. 1990, S. 47-49; Haykin 1999, S. 6]

² Obwohl die Auslösung eines Reizes als einer „Alles-oder-nichts“-Entscheidung mit gleichmäßig hohem elektrischen Potenzial erfolgt, vernachlässigt diese Betrachtung die Signalkodierung unterschiedlich starker Erregungszustände durch die zeitliche Abfolge von Aktivierungen, sog. Spikes [Haykin 1999, S. 7]. Auf eine Untersuchung der pulskodierten Informationsübertragung wird hier verzichtet; eine Übersicht findet sich z. B. bei [Maass und Bishop 1999].

Strukturelemente der Informationsvermittlung und Informationsspeicherung darstellen, bestimmt ihre Anzahl die Komplexität und Verarbeitungsmächtigkeit des Systems, während das Zusammenwirken der erregenden und der hemmenden synaptischen Verbindungen in der Bildung von Aktionspotenzialen das Systemverhalten determiniert. Diese grundlegende Struktur der Informationsverarbeitung wird bei der künstlichen Modellierung von Neuronalen Netzen nachempfunden.¹

4.2.2 Technische Realisierung

Die modelltechnische Realisierung und mathematische Formulierung der Signalverarbeitung in künstlichen Neuronen und künstlichen NN erfolgt durch abstrakte Modelle der Informationsverarbeitung in Analogie zum elektrochemischen, neurobiologischen Prozess. Der Grundbaustein künstlicher NN bildet das stark idealisierte, künstliche Neuron.² Jedes künstliche Neuron u_i ist eine abstrakte Einheit zur Informationsverarbeitung, die über n Eingabeschnittstellen numerische Daten aus dem System oder der Systemumwelt empfängt und über m Ausgangsschnittstellen verarbeitete Informationen an verbundene Systemelemente u_j oder die Systemumgebung weiterleitet. In Abbildung 4.3 ist ein künstliches Neuron in Übereinstimmung mit der biologischen Darstellung in Abbildung 4.1 schematisch dargestellt.

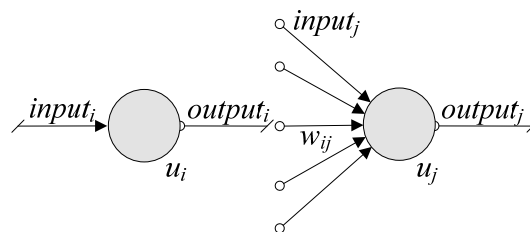


Abbildung 4.3: Künstliches Neuron im Neuronenverbund

In Analogie zur biologischen Verarbeitung erfolgt die numerische Berechnung je Neuron in drei disjunkten mathematischen Berechnungsvorschriften: der Eingabe-, der Aktivierungs- und der Ausgabefunktion. Die Informationsverarbeitung erfolgt sodann durch die Berechnung eines neuen Aktivierungszustandes aus den vorausgehenden Einheiten und dem aktuellen Zustand und dessen Weitergabe an verbundene Verarbeitungseinheiten [Strecker 1997, S. 13; Hecht-Nielsen 1990]. Das Neuron empfängt einen oder mehrere Datenwerte o_i über eine Eingangsschnittstelle und kombiniert diese durch einfache Berechnungen der Eingabefunktion net_j in einen numerischen Wert, in Anlehnung an die Funktion der biologischen Dendriten [Zell 2000, S. 83].³ Die Eingangswerte werden durch die Aktivierungsfunktion $f_{act}(net_j)$ ähnlich einem biologischen Aktivierungspotenzial transformiert, der transformierte Datenwert wird als Ausgabewert kodiert $f_{out}(f_{act}(net_j))$ und über eine Ausgabeschnittstelle ausgegeben [Stahlbock 2002, S. 33]. Die Ausgabe erfolgt in Anlehnung an die Alles-oder-nichts-Aktivierung biologischer Neuronen vielfach als binäre, nominal skalierte Ausgabe $\{0; 1\}$ oder alternativ als numerische Ausgabe $[0, 1]$. THIESING [1998, S. 6] interpretiert das Neuron selbst als Entscheidungsmodell, das durch die Informationsfindung, Informationsüberlagerung und Übergang eine dichotome „Ja/Nein“-Aussage im Sinne einer Entscheidung findet.⁴ Aufgrund

¹ Die Vorstellung dieser einfachen Arbeitsweise des Gehirns – und somit der Ähnlichkeit von digitalen Computern und dem Gehirn – ist nach dem gegenwärtigen Stand der Forschung überholt.

² Sie werden in Abgrenzung zum biologischen Vorbild synonym auch als Verarbeitungseinheit (engl. *processing unit*, *processing element* oder nur *unit*) [Zell 2000, S. 71] oder Knoten (engl. *node*) [Strecker 1997, S. 12] bezeichnet. Nachfolgend bezieht sich der Begriff ‚Neuron‘ ebenfalls auf künstliche Verarbeitungseinheiten.

³ Aus systemtheoretischer Betrachtungsweise kann ein Neuron auch als Multiple-Input-Single-Output-System interpretiert werden [Gupta und Rao 1994a, S. 7].

⁴ Dies erlaubt die weiterführende Interpretation von NN als Systeme von sich gegenseitig beeinflussenden dichotomen Entscheidungen, z. B. im Sinne von interagierenden Agenten auf einem Markt unter Kombinati-

der systembildenden Bedeutung und des hohen Freiheitsgrads der Modellierung dieser Funktionen werden sie in Abschnitt 4.2.3 detailliert dargestellt [Gupta und Rao 1994a, S. 7].

Die Stärke der synaptischen Verbindung zweier Neuronen wird als numerischer Wert des Verbindungsgewichtes w_{ij} zwischen dem sendenden Neuron u_i und dem empfangenden Neuron u_j abgebildet [Zell 2000, S. 70]. Die Verbindungen zwischen Neuronen werden als gerichtete, gewichtete Verbindung dargestellt, unter Vernachlässigung weiterer biologischer Schnittstellen. Die Menge aller Verbindungsgewichte kann in einer Gewichtsmatrix W abgebildet werden, in der ein Eintrag w_{ij} die Art und Stärke der künstlichen synaptischen Verbindung von Neuron u_i zu Neuron u_j bestimmt. Der absolute Betrag des Verbindungsgewichtes $|w_{ij}|$ spezifiziert die Stärke der übertragenen Signale, das Vorzeichen determiniert seine Wirkung. Bei positivem Vorzeichen wird eine exzitatorische Synapse mit verstärkender Wirkung auf die Aktivierung modelliert, bei negativem Vorzeichen die hemmende Wirkung einer inhibitorischen Synapse [Zell 2000, S. 46]. Der Wert $w_{ij} = 0$ bezeichnet das Fehlen einer Verbindung [Rumelhart et al. 1986, S. 49].

Der Ausgabewert o_j eines Neurons u_j wird über eine gewichtete Verbindung w_{ij} an nachfolgende Neuronen oder die Systemumgebung weitergeleitet, wodurch ein Netzwerk künstlicher Neuronen, ein künstliches NN, entsteht. Das künstliche NN ist aufgrund der Veränderbarkeit der numerischen Verbindungsgewichte zwischen den künstlichen Neuronen lernfähig, in Anlehnung an die synaptische Plastizität biologischer NN [Zell 2000, S. 46]. In NN wird dafür ein Algorithmus verwendet (als Lernregel bezeichnet), anhand dessen das NN durch wiederholte Präsentation von Trainingsmustern und entsprechende Modifikation der Verbindungsstärke den Fehler zwischen erwarteter und richtiger Ausgabe des NN für alle Trainingsmuster minimiert und so ihre korrekte Abbildung lernt.

In Analogie zum biologischen Vorbild werden Neuronen nach ihrer Funktion im NN differenziert und topologisch in Schichten geordnet. Eingabeneuronen (engl. *input units*), als Äquivalent sensorischer biologischer Neuronen, leiten Eingabemuster aus der Systemumwelt des Netzwerkes ohne eigenständige Berechnungen an interne Verarbeitungseinheiten weiter.¹ In den vor der Systemumwelt versteckten und von außen nicht manipulierbaren Neuronen (engl. *hidden units*) findet die originäre Informationsverarbeitung und Speicherung statt, deren Systemergebnis an die Ausgabeneuronen (engl. *output units*) zur Zwischenspeicherung und Aufbereitung einer sinnvollen und interpretierbaren Netzausgabe an die Systemumgebung weitergeleitet wird. Die Systemumgebung bezeichnet dabei die netzwerkexterne Umwelt, die die Eingabedaten an das NN anlegt und die Netzausgabe entgegennimmt [Strecker 1997, S. 13]. Nachfolgend bezeichnet der Begriff NN grundsätzlich die formalen Konstrukte künstlicher NN, unter Vernachlässigung der Abgrenzung zu biologischen Ausprägungen.

Die Modellbildung von NN erfordert eine eindeutige Spezifizierung aller Freiheitsgrade der Modellierung. Sie kann nach ALEX und STAHLBOCK [1998] formal durch die Ausprägungen des Tupels $U = [V, T, W, K]$ beschrieben werden, dargestellt durch die Signalverarbeitung im Neuron V , die Kombination von Neuronen durch ihre Verbindungs- oder Gewichtsmatrix W , ihre dynamische Aktivierungsstrategie K erweitert um die Netzwerktopologie T [Alex 1998, S. 85].² Die Signalverarbeitung V jedes Neurons wird durch die Verarbeitungsfunktionen der Netzwerk-Eingabe durch die Eingabefunktion F_I , die Berechnung des Aktivierungszustandes

on von Entscheidungen [Zimmermann et al. 2002]. Diese Sichtweise wird nachfolgend nicht weiter untersucht.

¹ Da in ihnen keine eigentliche Informationskodierung oder Verarbeitung stattfindet, werden sie nur der Vollständigkeit halber als Zwischenspeicher dargestellt.

² ALEX und STAHLBOCK [1996] benutzen eine reduzierte, aber erweiterungsfähige Notation, weshalb hier abweichende Symbole verwendet werden.

des Neurons durch die Aktivierungsfunktion F_A und die Berechnung der Ausgabe des Neurons durch die Ausgabefunktion F_O zum Tupel $V=[F_I, F_H, F_O]$ spezifiziert [Alex 1998, S. 86]. Die Notation nach wird nachfolgend um Komponenten der Lernparameter zur Parametrisierung und der Netzwerkarchitektur erweitert, um eine eindeutigen Spezifikation in der Modellierung von NN zur Prognose zu ermöglichen. Die Festlegung aller Parameter erlaubt anschließend eine eindeutige formale Identifikation sowie Replikation eines spezifischen NN und seiner Verhaltensweise. Nachfolgend werden die Freiheitsgrade der Modellierung im Rahmen dieser Systematik diskutiert und eine Auswahl von Teilfunktionen, Topologien und Aktivierungsstrategien zur Prognose mit NN vorgestellt.

4.2.3 Signalverarbeitung in künstlichen Neuronen

4.2.3.1 Eingabefunktionen

Die Eingabefunktion F_I^1 spezifiziert die Berechnungsvorschrift der Eingabesignale in ein Neuron u_j [Zell 2000, S. 83]. Die mathematische Berechnungsvorschrift der Summenbildung $F_I = net_j$ aggregiert die Ausgabesignale o_i ($i = 1, 2, \dots, m$) aller m verbundenen Vorgängerzellen in den Eingabewert net_j , unter Berücksichtigung der variablen Verbindungsgewichte w_{ij} , die die Intensität der Beziehung von Neuron u_i zu Neuron u_j angeben [Hecht-Nielsen 1990, S. 55-56; Stahlbock 2002, S. 36]. Die Netzeingabe net_j eines Neurons u_j berechnet sich üblicherweise als gewichtete Summe (Sum) [Zell 2000, S. 72-73] in der Form²

$$net_j = \sum_{i=1}^m o_i w_{ij} \quad . \quad (4.1)$$

Diese Linearkombination der Ausgaben o_i multipliziert mit den Verbindungsgewichten orientiert sich grob an der vereinfachten Verarbeitungsweise biologischer Neuronen [Zimmerer 1997, S. 20]. So verstärkt ein positives Verbindungsgewicht $w_{ij} > 0$ die Aktivierung des Neurons u_j oder hemmt sie bei negativem Verbindungsgewicht $w_{ij} < 0$. Durch die Abkehr von der biologischen Interpretation ist jedoch keine Beschränkung auf diese Eingabefunktion erforderlich. In Tabelle 4.1 ist eine Auswahl alternativer Eingabefunktionen mit spezifischen Eigenschaften für die Informationsverarbeitung in unterschiedlichen Netzwerkarchitekturen abgebildet, die selektiv zur Anwendung kommen.

¹ Sie wird auch als Nettoeingabefunktion (engl. *input function*) oder in Anlehnung an die engl. Bezeichnung „*propagation function*“ als Propagierungsfunktion bzw. Propagierungsregel bezeichnet [Rumelhart et al. 1986; Zell 2000; Haykin 1999, S. 10].

² Auf eine explizite Unterteilung der dynamischen Verarbeitungsschritte t in Eingabefunktion und nachfolgende Verarbeitungsfunktionen wie Aktivierungsfunktion etc. wird aus Gründen der Übersichtlichkeit verzichtet. Im Fall zeitlicher Interdependenzen durch rekursive Verbindungen anderer Neuronen oder vorheriger Aktivierungszustände desselben Neurons $a_j(t-1)$ werden diese als Verbindungsgewicht explizit modelliert und durch die Eingabefunktion integriert. Zu einer erweiterten, dynamischen Notation sowie einer alternativen Notation in Matrixschreibweise siehe [Zell 2000].

Tabelle 4.1: Eingabefunktionen in Neuronen

Bezeichnung	Formel	
Summe	$net_j = \sum_i o_i w_{ij}$	(4.2)
Dynamisch kumulierte Summe	$net_j^{t+1} = net_j^t + \sum_i o_i w_{ij}$	(4.3)
Maximum	$net_j = \text{Max}_i \{o_i w_{ij}\}$	(4.4)
Minimum	$net_j = \text{Min}_i \{o_i w_{ij}\}$	(4.5)
Majorität	$net_j = \sum_i \text{sgn}(o_i w_{ij})$ mit $\text{sgn}_i = \begin{cases} +1 & \forall w_{ij} o_i > 0 \\ -1 & \forall w_{ij} o_i \leq 0 \end{cases}$	(4.6)
Produkt (Sigma-Pi-Zellen)	$net_j = \prod_i o_i w_{ij}$	(4.7)
Euklidische Distanz	$net_j = \sqrt{\sum_i (o_i - w_{ij})^2}$	(4.8)
City-Block-Distanz	$net_j = \sum_i o_i - w_{ij} $	(4.9)

Die Verwendung alternativer Eingabefunktionen führt zur Abbildung eines veränderten Systemverhaltens. In Anwendungen zur Prognose kommen insbesondere Netzarchitekturen unter Verwendung statischer Summen (4.2) und dynamisch kumulierter Summenfunktionen (4.3) zum Einsatz. Zur Abbildung spezifischer Sachverhalte, z. B. der kausalanalytischen Modellierung mit Dummy-Variablen, können jedoch auch sog. Sigma-Pi-Netzwerkarchitekturen unter Multiplikation der Eingabewerte gebildet werden [Rumelhart et al. 1986].

4.2.3.2 Aktivierungsfunktionen

Anschließend an die Eingabe der Werte in ein Neuron berechnet die Aktivierungsfunktion f_{act} den Aktivierungszustand a_j eines Neurons u_j aus den Parametern der Netzeingabe net_j und dem Schwellwert des Neurons θ_j [Zell 2000, S. 72]¹, durch

$$a_j = f_{act}(net_j, \theta_j) \quad . \quad (4.10)$$

Der Aktivierungszustand a_j beschreibt den Grad der Aktivierung bzw. Erregung oder Ausgabe einer Zelle u_j im Zeitpunkt t als reelle Zahl [Zell 2000, S. 72]; der Vektor aller Aktivierungszustände a_j repräsentiert den Systemzustand des NN im Zeitpunkt t in Abhängigkeit von der dynamischen Aktivierung aller verbundenen Neuronen im Zeitablauf [Rumelhart et al. 1986, S. 48]. Somit kommt der Form der Aktivierungsfunktion, die im Rahmen der verwendeten Notation zur Informationsverarbeitung auch als F_A bezeichnet wird, eine besondere Bedeutung zu.

Der Schwellwert θ_j (engl. *threshold* oder *bias*) bezeichnet dabei jenen Wert, ab dem ein Neuron aus der kumulierten Netzeingabe aktiviert wird und eine Ausgabe an nachfolgende Neuronen oder die Systemumwelt vollzieht. „Mathematisch gibt er die Stelle der größten Steigung monoton wachsender Aktivierungsfunktionen an. Im biologischen Äquivalent

¹ Die Aktivierungsfunktion (engl. *activation function*) wird auch als Schwellwertfunktion bezeichnet. Der resultierende Zustand wird in Anlehnung an seine biologische Herkunft auch als Aktivierung (engl. *activation*) bezeichnet.

entspricht er der Reizschwelle, die benötigt wird, bis das Neuron feuert“ [Zell 2000, S. 81]. In der Regel wird der Schwellwert als additive Komponente modelliert, mit

$$a_j = f_{act}(net_j - \theta_j) \quad . \quad (4.11)$$

Alle mathematischen Aktivierungsfunktionen sind um den Wert Null zentriert, so dass für $\theta_j = 0$ ein Neuron umso stärker aktiviert bzw. eingeschaltet ist, je stärker seine Netzeingabe positiv ist, mit $net_j > 0$, und umso stärker inaktiv bzw. ausgeschaltet, je stärker diese negativ ist, mit $net_j < 0$. Der Schwellwert θ_j , ab dem ein Neuron j aktiviert wird, unterscheidet sich je Neuron und Aufgabenstellung des NN. Er stellt mit den Verbindungsgewichten w_{ij} einen freien Parameter der Modellbildung dar und wird im Rahmen der Parametrisierung auf den Daten bestimmt, um die Aktivierungsfunktion je Neuron auf einen spezifischen Wert zu zentrieren. Aufgrund ihrer zentralen Bedeutung für die Informationsverarbeitung wurde eine Vielzahl möglicher Aktivierungsfunktionen mit individuellem Systemverhalten und spezifischen Aktivierungszuständen entwickelt, die im Kontext der Prognose Verwendung finden können. In Abbildung 4.4 ist in Anlehnung an ZELL [2000, S. 76] eine Auswahl verbreiteter Aktivierungsfunktionen, ihrer Ableitungen und Aktivierungszustände dargestellt.

Die Aktivierungsfunktionen bilden eine nichtlineare Schwellwertlogik um einen variablen Schwellwert ab. Die Darstellung der Aktivierung a_j erfolgt hier in Abhängigkeit von der Netzeingabe net_j bei einem Schwellwert $\theta_j=0$. Eine Veränderung von θ_j durch Training etc. entspricht einer Verlagerung des Schwellwertes als Bereich des stärksten Anstiegs der Aktivierung entlang der Abszisse. Aktivierungsfunktionen werden in diskrete, binäre Aktivierungsfunktionen und stetige Funktionen metrischer Aktivierungen unterschieden.

Die binäre Schwellwertfunktion¹ in Abbildung 4.4 b.) bildet die ursprüngliche, binäre Schwellwertlogik der Signalverarbeitung des MCCULLOCH-PITTS-Neurons als logische „Wahr-oder-falsch“-Aussage ab, weshalb ihr eine besondere theoretische Bedeutung zukommt [Anderson 1995, S. 49-51]. Sie ist an der Sprungstelle nicht differenzierbar, und ihre Ableitung beträgt sonst Null [Zell 2000, S. 90]. Lernverfahren für mehrlagige Netzwerkarchitekturen, z. B. der Backpropagation-Algorithmus, benötigen als Verfahren des Gradientenabstiegs eine vollständig differenzierbare erste Ableitung der Aktivierungsfunktion zur Konvergenz. Neben Anwendungen in historischen MCCULLOCH-PITTS-, WIDROW-HOFF- und HOPFIELD-Modellen werden binäre Aktivierungsfunktionen aufgrund ihrer Integrationsfähigkeit im Bereich digitaler Elektronik, ihrer Nähe zur Modellierung formaler Logik und ihrer einfachen Anwendung in ausgewählten Problemstellungen noch vereinzelt in modernen Netzarchitekturen verwendet [Anderson 1995, S. 58 ff.]. Die lineare Schwellwertfunktion mit Sättigungsgrenzen in Abbildung 4.4 c.) wurde als Weiterentwicklung der binären Aktivierungsfunktion für spezielle Netzwerkarchitekturen konzipiert. Sie steigt im Intervall]-1, 1[monoton linear mit einer konstanten Ableitung $f'_{act}(net) = 1$. Da sie außerhalb des Intervalls und an den Sprungstellen -1 und 1 nicht definiert ist, kann sie ebenfalls nicht ohne Weiteres in Lernverfahren des Gradientenabstiegs eingesetzt werden [Anderson 1995, S. 58]. Da binäre Schwellwertfunktionen nicht differenzierbar sind, fand eine Abkehr von diskreten, binären oder bipolaren Aktivierungsfunktionen zugunsten semilinear, monoton steigender und vollständig differenzierbarer Aktivierungsfunktionen statt [Anderson 1995, S. 55].

¹ Sie wird auch als Schrittfunktion (engl. *step function*), *squashing function* oder *limiting function* bezeichnet.

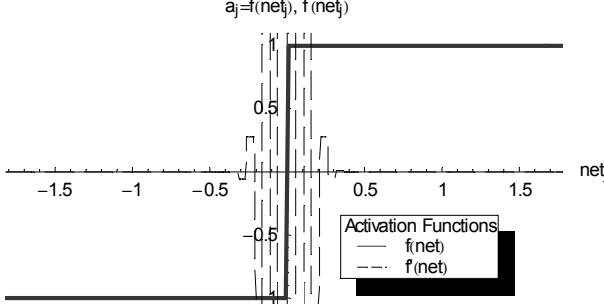
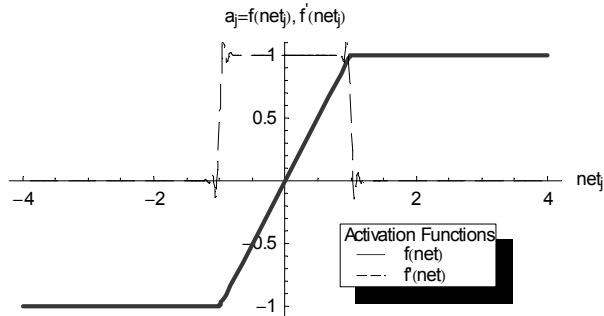
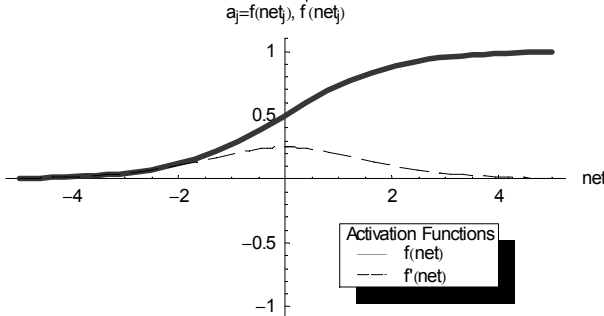
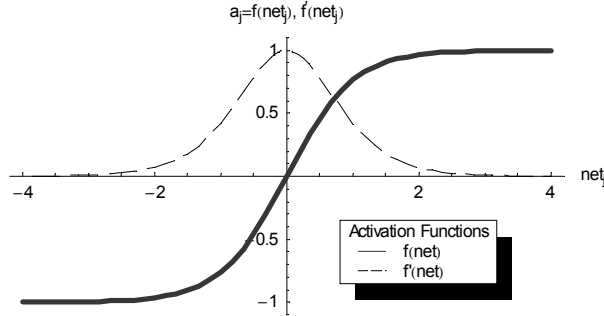
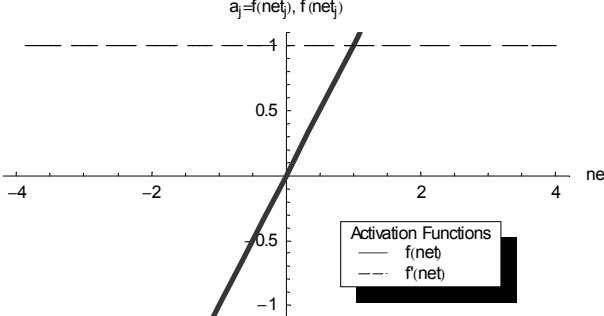
Aktivierungsfunktion $a_j = f(\text{net}_j)$	Aktivierungszustand a_j
<p>a.) $a_j = f(\text{net}_j), \hat{f}(\text{net}_j)$</p>  <p>Activation Functions — f(net) - - - f̂(net)</p>	<p>Bezeichnung: Binäre Schwellwertfunktion Zustand: Binär $\{0;1\}$ oder $\{-1;1\}$</p> $f_{act}(\text{net}_j, \theta_j) = \begin{cases} 1 & \forall \text{net}_j \geq \theta_j \\ -1 & \forall \text{net}_j \leq -1 + \theta_j \end{cases}$ $= \text{sgn}(\text{net}_j)$
<p>b.) $a_j = f(\text{net}_j), \hat{f}(\text{net}_j)$</p>  <p>Activation Functions — f(net) - - - f̂(net)</p>	<p>Bezeichnung: Linear bis Sättigung Zustand: Intervall $[-1, 1]$</p> $f_{act}(\text{net}_j, \theta_j) = \begin{cases} -1 & \forall \text{net}_j \leq -1 + \theta_j \\ \text{net}_j & \forall -1 + \theta_j \leq \text{net}_j \leq 1 + \theta_j \\ +1 & \forall \text{net}_j \geq 1 + \theta_j \end{cases}$
<p>c.) $a_j = f(\text{net}_j), \hat{f}(\text{net}_j)$</p>  <p>Activation Functions — f(net) - - - f̂(net)</p>	<p>Bezeichnung: Logistische Funktion Zustand: Intervall $]0; 1[$</p> $f_{act}(\text{net}_j, \theta_j) = \frac{1}{1 + e^{-\text{net}_j - \theta_j / T}}$ $= \log(\text{net}_j - \theta_j)$
<p>d.) $a_j = f(\text{net}_j), \hat{f}(\text{net}_j)$</p>  <p>Activation Functions — f(net) - - - f̂(net)</p>	<p>Bezeichnung: Tangens Hyperbolicus Zustand: Intervall $] -1; 1[$</p> $f_{act}(\text{net}_j, \theta_j) = \frac{e^{\text{net}_j - \theta_j} - e^{-\text{net}_j - \theta_j}}{e^{\text{net}_j - \theta_j} + e^{-\text{net}_j - \theta_j / T}}$ $= \tanh(\text{net}_j - \theta_j)$
<p>e.) $a_j = f(\text{net}_j), \hat{f}(\text{net}_j)$</p>  <p>Activation Functions — f(net) - - - f̂(net)</p>	<p>Bezeichnung: Identität Zustand: Unbeschränkt in \mathbb{R}</p> $f_{act}(\text{net}_j, \theta_j) = a_j = \text{net}_j - \theta_j$

Abbildung 4.4: Aktivierungsfunktionen und Aktivierungszustände

Die Anwendung stetiger, nichtlinearer Aktivierungsfunktionen mit sigmoidem Verlauf geht über RUMELHART et al. [1986] auf GROSSBERG [1976a; 1976b] und WILSON und COWAN [1972] zurück. Sie bilden die binäre Schwellwertlogik angenähert als kontinuierliche Funkti-

on ab.¹ MÜLLER [1995] interpretiert ihre Anwendung als Übergang von einer deterministischen Aktivierungsfunktion mit zwei binären Zuständen zu einer stochastischen Modellierung, bei der anstelle der Aktivierungswerte die Wahrscheinlichkeit modelliert wird, dass die Aktivierung a_j den Wert 0 oder 1 annehmen wird, im Sinne von

$$p(a_j) = f_{act}(net_j - \theta_j) \quad (4.12)$$

Dabei ist die Aktivierungsfunktion $f_{act}(\bullet)$ im Intervall $[0, 1]$ beschränkt

$$f(net \rightarrow -\infty) = 0, f(net \rightarrow +\infty) = 1 \quad (4.13)$$

und dazwischen monoton steigend und vollständig differenzierbar [Müller et al. 1995, S. 38]. Vielfach wird hierzu die Logistische Funktion (Log) verwendet², mit

$$f_{act}(net_j, \theta_j) = \left(1 + e^{\frac{-net_j - \theta_j}{T}}\right)^{-1} \quad (4.14)$$

Der Skalierungsparameter T beträgt standardmäßig $T=1$ und bestimmt die Steigung der logistischen Aktivierungsfunktion [Zell 2000, S. 91]. Für Grenzwerte des Skalierungsparameters $T \rightarrow 0$ wird die Steigung der Logistischen Funktion verringert, und sie nähert sich der linearen Funktion an [Zell 2000, S. 92]; für $T \rightarrow \infty$ wird die Steigung erhöht, die Funktion nähert sich einer binären Schwellwertfunktion an, und das stochastische Netzwerk geht in eine deterministische Modellierung über [Müller et al. 1995, S. 38]. Die Ableitung der Logistischen Aktivierungsfunktion ist zudem leicht aus sich selbst zu berechnen, was implementierungstechnische Vorteile hat [Zell 2000, S. 92].

Im Gegensatz zu binären oder linear gesättigten Aktivierungsfunktionen sind sigmoide Funktionen im offenen Wertebereich $]0, 1[$ stetig, überall differenzierbar und punktsymmetrisch um den Schwellwert. Somit ist ihre Ableitung an jeder Stelle definiert [Zell 2000, S. 90]. Die Steigung der sigmoiden Aktivierungsfunktion spezifiziert die Stärke des nichtlinearen Übergangs aus einem Aktivierungszustand in einen anderen, im Sinne ihrer Sensibilität hinsichtlich Veränderungen der Netzwerkeingabe. Der Bereich der steilsten Steigung aller sigmoiden Funktionen liegt um den Schwellwert θ_j , in der Abbildung 4.4 gilt $\theta_j = 0$ ³, während für extrem hohe oder niedrige Eingabewerte in den Sättigungsgrenzen nur geringe Veränderungen der Ausgabe aus veränderten Netzwerkeingaben resultieren. Somit lösen sigmoide Funktionen das Problem, „dass ein Netzwerk häufig sowohl auf Signale kleiner als auch sehr großer Amplitude reagieren muss, wobei es bei kleinen Amplituden wesentlich sensibler sein muss“ [Zell 2000, S. 90]. Der beschränkte Reaktionsbereich kann theoretisch durch die Verschiebung des Parameters des Schwellwertes θ_j je Neuron an die entsprechenden Eingabewerte angepasst werden, erfordert jedoch vielfach die Skalierung der Eingabewerte auf das Intervall der jeweiligen Aktivierungsfunktion.

Die ebenfalls populäre sigmoide Funktion des Tangens Hyperbolicus (TanH) gibt für Eingabewerte im Intervall $[-\infty; \infty]$ Ausgaben der Aktivierung zwischen $] -1; 1[$ an [Anderson

¹ Detaillierte Erläuterungen zur biologischen Motivation und Adäquanz sigmoider Aktivierungsfunktionen und negativer Aktivierungsbereiche findet sich bei [Anderson 1995, S. 57-60; Hopfield 1984].

² Sie wird in Anlehnung an die thermale Energieverteilung in einem System identischer „fermions“ in der Physik auch als Fermi-Funktion bezeichnet. Der Skalierungsparameter der Temperatur T modelliert die Annäherung der Fermi-Funktion an die binäre Schwellwertfunktion für $T \rightarrow \infty$, bzw. unter alternativer Notation von β als inverse mit $\beta^{-1} = T$ für $\beta \rightarrow 0$. Für eine physikalische Interpretation der Aktivierungsfunktionen siehe [Müller et al. 1995, S. 38 ff.].

³ ZELL [2000, S. 90] bezeichnet die lokale Gegend um den Schwellwert auch als Arbeitspunkt.

1995]. Bei Verwendung des Lernverfahrens Backpropagation wird durch den TanH das Lernen beschleunigt, da bei Nichtaktivierung des Vorgängerneurons u_i mit $net_i = -1$ das Verbindungsgewicht w_{ij} aufgrund der Lernregel reduziert wird, während bei Verwendung der logistischen Aktivierungsfunktion die Aktivierung $net_i = 0$ als Produkt in der Lernregel keine Gewichtsänderung bewirkt [Zell 2000, S. 92].¹ Die Ableitung des TanH ist ebenso leicht zu berechnen wie die Ableitung der logistischen Aktivierungsfunktion und bringt daher keine verfahrenstechnischen Vorteile [Zell 2000, S. 92].

Darüber hinaus existiert eine Vielzahl weiterer semilinear (monoton und vollständig differenzierbarer), trigonometrischer und hyperbolischer Aktivierungsfunktionen, die jedoch vergleichsweise selten Anwendung bei der Prognose finden. Trotz der theoretischen Bedeutung nichtlinearer Funktionen des Sinus (\sin), mit $\sin(net_j)$ im Bereich $[-1, 1]$, und GAUSS'scher Glockenfunktionen in der Klasse sog. Radialer Basisfunktionsnetze (RBF NN) werden sie von den sigmoiden Funktionstypen in Theorie und Praxis dominiert [Zimmerer 1997, S. 21].

Die lineare Identitätsfunktion (Id) in Abbildung 4.4 a.) ist im Intervall $[-\infty; \infty]$ definiert und stellt keine eigentliche Aktivierungsfunktion für versteckte Neuronen dar. Vielmehr wird sie als Ausgabe- und Aktivierungsfunktion in Eingabe- und Ausgabeneuronen verwendet und hier nur der Vollständigkeit halber abgebildet.

Aktivierungsfunktionen f_{act} beschränken den auszugebenden Funktionswert auf einen Bereich, in Anlehnung an die biologisch motivierte binäre „Alles-oder-nichts“-Ausgabe [Anderson 1995, S. 57].² RUMELHART et al. [1986] bezeichnen sie entsprechend als „squashing function“, die Werte aus $[-\infty, +\infty]$ in das Intervall des Aktivierungszustands $[l, u]$ abbildet [Anderson 1995, S. 57]. ZELL [2000] unterscheidet (quasi-)kontinuierliche Wertebereiche, unter unbeschränkter Abbildung aller reellen Zahlen oder unter Beschränkung auf ein Intervall $[l, u]$ aus einer unteren Grenze l und einer oberen Grenze u , wobei vielfach die zulässigen Wertebereiche $l = 0, u = 1$ durch die populäre logistische Aktivierungsfunktion oder $l = -1, u = +1$ durch die Aktivierungsfunktion des Hyperbolischen Tangens (TanH) abgebildet werden. Diese unterscheiden sich von den diskreten Wertebereichen der binären Aktivierungszustände $\{0, 1\}$ und den bipolaren Werten $\{-1, 1\}$ bzw. $\{-, +\}$. In Abbildung 4.5 gibt es eine Übersicht alternativer Aktivierungszustände dargestellt.

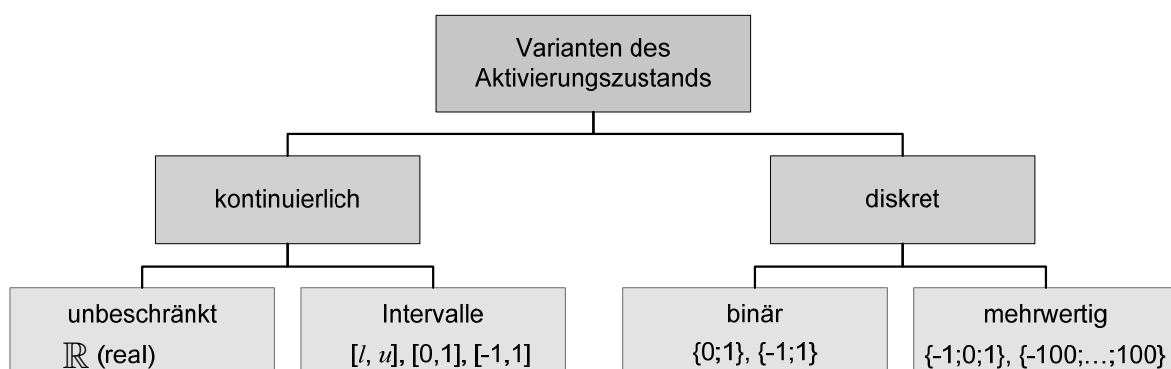


Abbildung 4.5: Aktivierungszustände von Aktivierungsfunktionen

Die unterschiedlichen Varianten sind unterschiedlich gut für die Anwendung zur Punktprognose im Sinne einer Regression geeignet. Diskrete Aktivierungszustände von Ausgabeneuronen eignen sich zur Abbildung von Klassenzugehörigkeiten, nicht aber zur

¹ Somit sollten Lernverfahren theoretisch durch die Verwendung des TanH schneller konvergieren, was jedoch keine Aussagen über die Fehlerrobustheit ihrer Verwendung zulässt.

² Dies wird auch durch die beschränkte Anzahl von dynamischen Aktionspotenzialen biologisch begründet.

Ausgabe metrisch skaliertes abhängiger Variablen im Sinne der Punktprognose.¹ Kontinuierliche Aktivierungszustände erlauben dagegen die Abbildung metrischer abhängiger Variablen durch unbeschränkte Ausgabewerte. Für Aktivierungen in Intervallen $[l, u]$ ist dagegen eine Reskalierung der Ausgabe jedes Neurons erforderlich, um eine Interpretation in der ursprünglichen Domäne zuzulassen.

4.2.3.3 Ausgabefunktionen

Die Ausgabefunktion F_O berechnet die Ausgabeinformation o_j eines Neurons u_j aus seiner Aktivierung a_j [Zell 2000, S. 76] in der Form

$$o_j = f_{out}(a_j) \quad . \quad (4.15)$$

Sie erlaubt die Kodierung der Ausgabe durch Skalierung der Aktivierungswerte in ein abweichendes Intervall sowie die Modellierung z. B. statistisch motivierter Verhaltensweisen spezifischer NN-Architekturen. Dabei kommen spezifische Funktionen zur Koordination der Ausgabe bei Klassifikationsaufgaben zum Einsatz, die in den Anwendungen der Regression nicht relevant sind.² Im Fall der nachfolgend analysierten Netzarchitekturen zur Prognose mit nichtlinearen Aktivierungsfunktionen kann die Ausgabefunktion daher vernachlässigt bzw. als Identitätsfunktion (siehe Abbildung 4.4) modelliert werden.³ Die Ausgabe o_j des Neurons u_j ist dann gleich seiner Aktivierung a_j , also

$$o_j = f_{ident}(a_j) = a_j \quad . \quad (4.16)$$

Die Ausgabefunktion kann somit in den nachfolgenden grafischen Darstellungen zur Punktprognose zur Vereinfachung der Formeln und der Notation vernachlässigt werden. Die Verwendung der Identitätsfunktion als Ausgabefunktion sollte jedoch zur vollständigen Spezifikation in der Notation der Informationsverarbeitung $V=[F^I, F^A, F^O]$ durch $F^O = \text{Id}$ angegeben werden.

4.2.3.4 Kombination der Verarbeitungsfunktionen

Die Informationsverarbeitung im Neuron erfolgt durch die sequenzielle Berechnung der Verarbeitungsfunktionen der Eingabe, der Aktivierung und der Ausgabe des Neurons, mit

$$o_j = f_{out}(f_{act}(net_j - \theta_j)) \quad . \quad (4.17)$$

Die Kombination der unterschiedlichen Eingabefunktionen F_I , der Aktivierungsfunktionen F_H und der Ausgabefunktionen F_O determiniert die Signalverarbeitung eines Neurons $V=[F_I, F_H, F_O]$ und somit sein Verhalten und das der resultierenden Netzwerkarchitektur.

Verarbeitungsfunktionen werden in der Literatur uneinheitlich bezeichnet und kombiniert, z. B. durch Zusammenfassung der Aktivierungs- und der Ausgabefunktion, der Eingabe- und der Aktivierungsfunktion oder gar der Kombination aller Verarbeitungsfunktionen zu einer

¹ Auf eine Untersuchung diskreter Aktivierungszustände in allen oder ausgewählten versteckten Neuronen und linearen Aktivierungsfunktionen in Ausgabeneuronen für metrische Prognoseaussagen wird hier verzichtet. Sie könnten zur Abbildung dichotomer Aussagen und Regimewechsel beitragen, erfordern jedoch eine eigenständige Untersuchung.

² Sie modellieren beispielsweise bei wettbewerbsorientiertem Lernen in Mehrklassenproblemen die Auswahl des Ausgabeneurons mit der höchsten Aktivierung sowie die normalisierte Ausgabe einer Wahrscheinlichkeit der Klassenzugehörigkeit durch eine Softmax-Ausgabefunktion [Neuralware 1991].

³ Einige Autoren postulieren die Verwendung nichtlinearer „Ausgabefunktionen“ oder „Transferfunktionen“ anstelle der Aktivierungsfunktion [Zell 2000, S. 76]. Dies ist auf eine irreführende Bezeichnung bei der Kombination von Verarbeitungsfunktionen zurückzuführen, wie im nachfolgenden Abschnitt gezeigt wird.

allgemeinen Transferfunktion [Zell 2000, S. 89; Rumelhart et al. 1986; Alex 1998, S. 88]. Grundsätzlich erfolgt eine Verarbeitung jedoch in den elementaren Schritten Eingabe, Berechnung der Aktivierung und Kodierung der Ausgabe, weshalb die einzelnen Berechnungen als disjunkte Verarbeitungsfunktionen abgebildet werden. Zur vereinfachten Darstellung der Informationsverarbeitung in künstlichen Neuronen wird anstelle der mathematischen Darstellung in Formel (4.17) vielfach auf eine grafische Repräsentation der Informationsverarbeitung als gerichteter, gewichteter Graph zurückgegriffen, die in Abbildung 4.6 dargestellt ist.

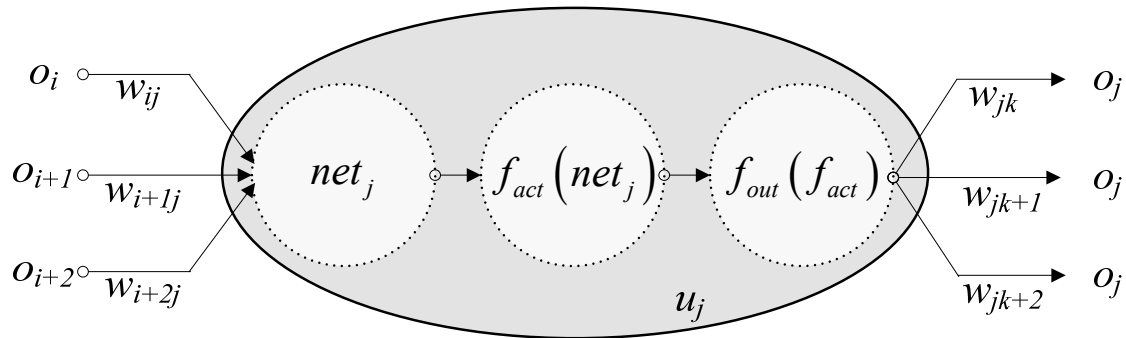


Abbildung 4.6: Signalverarbeitung im künstlichen Neuron

Nach MCCULLOCH und PITTS [1943] werden drei Gruppen von Neuronen je nach ihrer Funktion und ihrer topologischen Ordnung im NN unterschieden: Eingabe-Neuronen (engl. *input units*), versteckte Neuronen (engl. *hidden units*¹) und Ausgabe-Neuronen (engl. *output units*). „The input units react to particular data features from the environment. ... The output units generate particular organismic responses. ... The hidden units ... are neither input nor output unit themselves, but, via network connections, influence output units to respond to prescribed patterns of input unit firings or activities” [Levine 2000, S. 15]. Diese Unterscheidung erfolgt in Übereinstimmung zur biologischen Unterscheidung von Neuronen in sensorische, verarbeitende und motorische Neuronen (vgl. Abschnitt 4.2.1) [Levine 2000, S. 15]. Ihre unterschiedliche Funktionsweise in NN wird allein durch die Kombination von Eingabe-, Aktivierungs- und Ausgabefunktionen bestimmt.

Eingabeneuronen leiten Eingabemuster aus der Systemumwelt an interne Verarbeitungseinheiten weiter. In ihnen erfolgt grundsätzlich keine lineare oder nichtlineare Informationsverarbeitung durch Berechnung von Aktivierungspotenzialen, sondern lediglich eine Weiterleitung von Werten an interne Verarbeitungseinheiten. Dabei nimmt ein Eingabeneuron u_j grundsätzlich genau ein Eingabesignal x_j auf, mit

$$o_j = x_j \quad (4.18)$$

Unter Verwendung einer einheitlichen Notation entspricht dies der Berechnung der Eingabefunktion net_j als Summe über ein einzelnes Signal $m = 1$ und der Identitätsfunktion als Aktivierungs- und Ausgabefunktion, was einem einfachen Anlegen des Eingabewertes der Systemumgebung an die Ausgabe des Neurons entspricht, mit $F_I = \text{Sum}$, $F_H = \text{Id}$ und $F_O = \text{Id}$ bzw. dem Tupel $V = [\text{Sum}, \text{Id}, \text{Id}]$.

In den vor der Systemumwelt versteckten und nicht unmittelbar manipulierbaren Neuronen findet die originäre Informationsverarbeitung des NN statt [Zell 2000, S. 74]. Sie verwenden überwiegend nichtlineare, sigmoide Funktionen $V = [\text{Sum}, \text{TanH}, \text{Id}]$ oder $V = [\text{Sum}, \text{Log}, \text{Id}]$.

¹ Die englische Bezeichnung geht auf RUMELHART und MCCLELLAND [1986] sowie die Mitglieder ihrer Forschungsgruppe des „Parallel Distributed Processing“ (PDP) an der UCSD zurück.

Die Ausgabe eines versteckten Neurons u_j mit nichtlinearer Signalverarbeitung berechnet sich unter Verwendung der Summe als Eingabefunktion, des TanH als Aktivierungsfunktion und der Identitätsfunktion als Ausgabefunktion, mit $V=[\text{Sum}, \text{TanH}, \text{Id}]$, durch

$$o_j = \tanh\left(\left(\sum_{i=1}^N o_i w_{ij}\right) - \theta_j\right) = \frac{\left(e^{\left(\sum_{i=1}^N o_i w_{ij}\right) - \theta_j} - e^{-\left(\sum_{i=1}^N o_i w_{ij}\right) - \theta_j}\right)^2}{\left(e^{\left(\sum_{i=1}^N o_i w_{ij}\right) - \theta_j} + e^{-\left(\sum_{i=1}^N o_i w_{ij}\right) - \theta_j}\right)^2} . \quad (4.19)$$

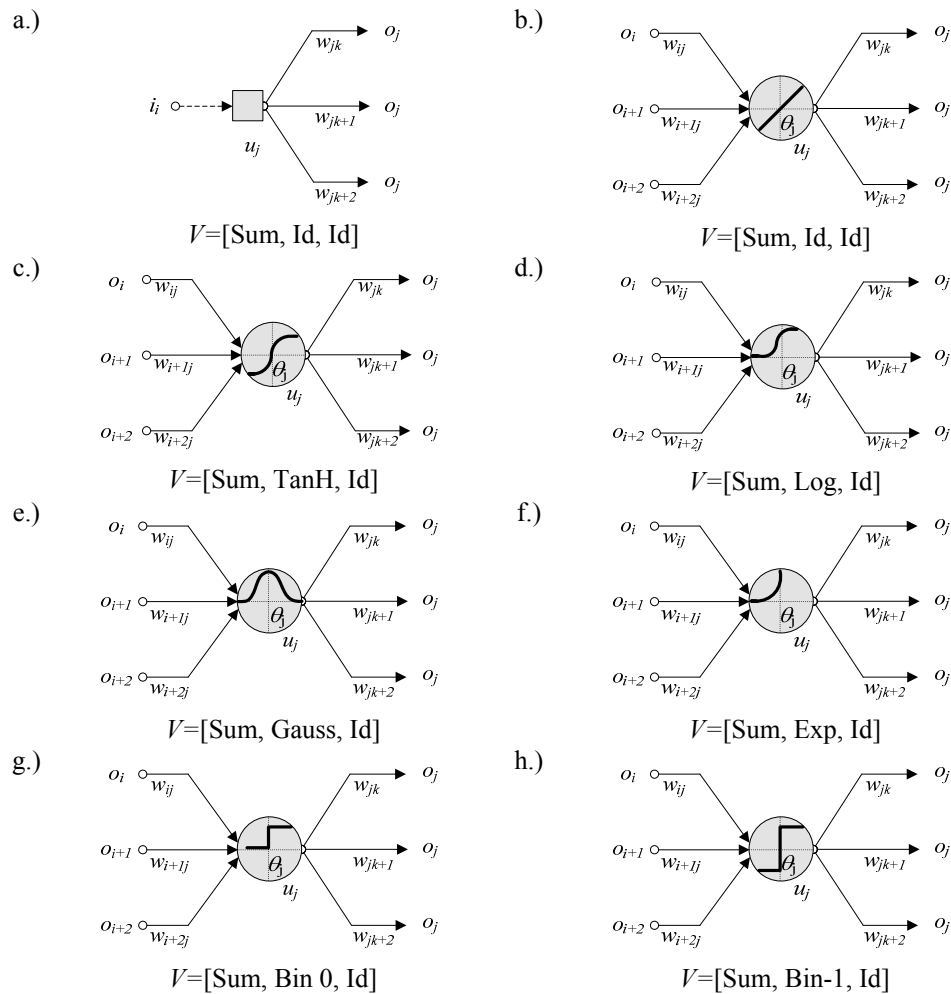
Die Verwendung von linearen Aktivierungsfunktionen der Identität in versteckten Neuronen $V=[\text{Sum}, \text{Id}, \text{Id}]$ erscheint wenig sinnvoll, da die Berechnung zweier nachgeschalteter Neuronen $V=[\text{Sum}, \text{TanH}, \text{Id}]$ und $V=[\text{Sum}, \text{TanH}, \text{Id}]$ mit beliebigen Verbindungsgewichten durch ein einzelnes Neuron $V=[\text{Sum}, \text{TanH}, \text{Id}]$ mit kombiniertem Verbindungsgewicht effizienter modelliert werden kann [Zell 2000].¹

In Ausgabeneuronen wird die empfangene Information in eine sinnvolle und interpretierbare Netzausgabe aufbereitet. Die Verarbeitung erfolgt je nach Modellierungsziel und -aufgabe entweder als nichtlineare Informationsverarbeitung analog zu den versteckten Neuronen $V=[\text{Sum}, \text{TanH}, \text{Id}]$ oder $V=[\text{Sum}, \text{Log}, \text{Id}]$ ², zur Abbildung eines spezifischen Systemverhaltens durch spezielle Ausgabefunktionen $V=[\text{Sum}, \text{Id}, \text{Softmax}]$ oder Eingabefunktionen $V=[\text{Max}, \text{Id}, \text{Id}]$ etc. oder vielfach einfach als lineare Ausgabe der nichtlinearen Informationsverarbeitung der versteckten Neuronen durch $V=[\text{Sum}, \text{Id}, \text{Id}]$. Die Verwendung einer linearen Aktivierungsfunktion $V=[\text{Sum}, \text{Id}, \text{Id}]$ in Ausgabeneuronen mit Eingängen von n Vorgängerneuronen berechnet eine einfache Linearkombination, und erlaubt somit auch eine Interpretation der Informationsverarbeitung als Gewichtung unterschiedlicher Teilausgaben vorgelagerter Neuronen im Sinne einer Gremien- oder Ensemblebildung.

Da die Kombination der einzelnen Verarbeitungsfunktionen das Verhalten des Neurons und das Netzverhalten bestimmt sowie die Vielfalt möglicher Modellierungsansätze NN zur Prognose begründet, ist in Abbildung 4.7 eine exemplarische Übersicht häufig verwendeter Neuronentypen dargestellt.

¹ ZELL [2000, S. 89f.] beweist für vorwärtsgerichtete Netze, dass ein NN aus n Neuronenschichten mit linearen Aktivierungsfunktionen durch ein Netzwerk mit $n-1$ versteckten Schichten bei identischer Informationsverarbeitung mit weniger freien Parametern und somit effizienter realisiert werden kann.

² Von einer Modellierung als $V=[\text{Sum}, \text{TanH}, \text{Log}]$ oder $V=[\text{Sum}, \text{Log}, \text{TanH}]$ ist abzusehen, da die unterschiedlichen Aktivierungs- und Ausgabebereiche $[0, 1]$ und $[-1, 1]$ zu Problemen in der Modellierung führen können.

Abbildung 4.7: Grafische Notation der Informationsverarbeitung im Neuron¹

Die vereinfachte Notation in Abbildung 4.7 nach $V=[\text{Sum}, f_{act}, \text{Id}]$ vernachlässigt die vielfach unveränderten Eingabe- und Ausgabefunktionen. Sie spezifiziert die Informationsverarbeitung im Neuron als Kreis unter Abbildung der Form der linearen bzw. nichtlinearen Aktivierungsfunktionen. Eingabeneuronen ohne eigentliche Informationsverarbeitung, aber mit linearer Weitergabe eines einzelnen Beobachtungswertes an das Netz, werden nach Abbildung 4.7 a.) abweichend als Rechtecke dargestellt. Trainierbare Verbindungen mit veränderbaren Gewichten w_{ij} werden als gerichteter Pfeil dargestellt, nicht trainierbare Verbindungen als unterbrochene Linie visuell abgegrenzt.

In Abbildung 4.7 werden Neuronen mit linearer Aktivierungsfunktion (b.), wie sie vielfach als Ausgabeneuron verwendet wird, mit einer nichtlinearen Aktivierungsfunktion des TanH (c.), der logistischen Funktion (d.), einer GAUSS'schen Funktion zur Modellierung radialer Basisfunktionen (e.), der Exponentialfunktion (f.), sowie der Binären Schwellwertfunktion im Intervall $[0; 1]$ (g.) sowie im Intervall $[-1; 1]$ (h.) abgebildet. Diese Neuronen werden anschließend unter grafischer Notation zu Netzwerken kombiniert, in denen die Kombination der einzelnen Neuronen das Systemverhalten bestimmt.

¹ Die Darstellung zeigt Neuronen in der grafischen Notation mit unterschiedlichen Aktivierungsfunktionen, bei konstanter Summe als Eingangsfunktion und der Identitätsfunktion als Ausgabefunktion mit zu- und abgehenden Verbindungen.

4.2.4 Informationsverarbeitung in Netzwerken von Neuronen

4.2.4.1 Netzwerktopologie

Die Informationsverarbeitung im künstlichen Neuron erfolgt in Analogie zum biologischen Vorbild sequenziell und unidirektional. Durch gewichtete, gerichtete Verbindungen eines Neurons u_i mit anderen Neuronen u_j , $j = 1, \dots, m$, entsteht ein Netzwerk von Neuronen, welches entsprechend als künstliches Neuronales Netz bezeichnet wird.¹ Das Netzwerk entspricht einer sequenziellen Folge funktionaler Berechnungen in Neuronen, im Sinne von

$$o_k = \tanh \left(\sum_k w_{kj} \tanh \left(\sum_i w_{ij} \tanh \left(\sum_j w_{ji} o_j - \theta_j \right) - \theta_i \right) - \theta_k \right) \Rightarrow \text{Min!} \quad (4.20)$$

NN können somit als Netz von Funktionen aufgefasst werden [Rojas 1993, S. 74].² Aus Gründen der Übersichtlichkeit wird anstelle der mathematischen Darstellung nachfolgend eine Visualisierung als gerichteter Graph verwendet, bei dem die Neuronen als Knoten und die Verbindungen zwischen den Neuronen als gerichtete Kanten realisiert sind [Zell 2000].³ Die grafische Darstellung der Neuronen im Netzwerk beschreibt die Netzwerktopologie T als Struktur der Verarbeitungseinheiten und ihrer Verbindungen, die Rückschlüsse auf den Netztyp und die Fähigkeit zur Informationsverarbeitung erlaubt. Die Darstellung einer vorwärts gerichteten Netzwerktopologie findet sich in Abbildung 4.8.

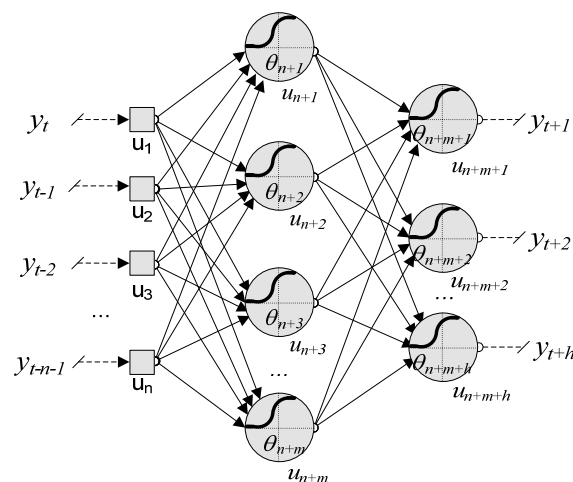


Abbildung 4.8: Mehrschichtiges Neuronales Netz

Entsprechend ihrer struktureller Anordnung und Funktion im Netzwerk können die einzelnen Neuronen topologisch in Schichten identischer Funktion geordnet werden. So unterscheidet ROSENBLATT (1.) die Eingabeschicht, deren Zellen aus Eingabeneuronen die Eingangssignale der Systemumwelt ohne Informationstransformation in das Netz weiterleiten und die keine vorgelagerten Neuronen besitzen, (2.) die Ausgabeschicht als Menge eines oder mehrerer Ausgabeneuronen, die die Netzausgabe an die Systemumgebung weiterleiten, und (3.) eine oder mehrere versteckte Schichten aus versteckten Neuronen zur eigentlichen Informations-

¹ Es ist anzuführen, dass die deutsche Bezeichnung „Netz“ im Gegensatz zu „Netzwerk“ nicht treffend ist, aus Gründen des allgemeinen Gebrauchs jedoch auch nachfolgend verwendet wird.

² Diese Interpretation erscheint für die nachfolgende Betrachtung und Analyse von NN zur Prognose im Vergleich zur funktionalen Abbildung von Prognoseverfahren und Prognosemodellen hilfreich.

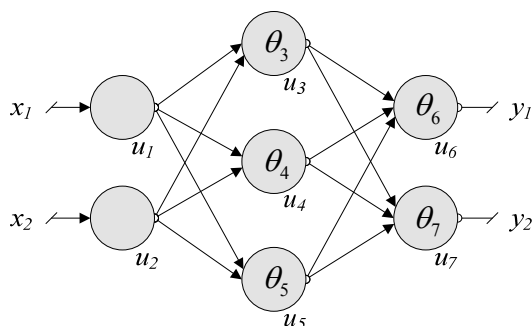
³ Nach [Alex 1998, S. 104-106] finden unterschiedliche Darstellungsformen Verwendung. Aus Gründen der Übersichtlichkeit und der Vollständigkeit der Abbildung wird eine erweiterte Notation in Anlehnung an [Haykin 1999; Reed und Marks 1999; Rojas 1993; Zell 2000] entwickelt.

verarbeitung [Rosenblatt 1958].¹ Die Systemumgebung bezeichnet dabei die netzwerkexterne Umwelt, die die Eingabedaten an das NN anlegt und die Netzausgabe entgegennimmt [Strecker 1997, S. 13].

Die Anzahl der Schichten eines NN wird uneinheitlich spezifiziert. Nach ZELL [2000, S. 73] werden die Schichten trainierbarer Gewichte gezählt. Dies erlaubt jedoch keine eindeutige Spezifikation von Topologien mit nicht-trainierbaren Gewichten, Verbindungen zwischen Neuronen über Schichten hinweg oder nicht vollständig verbundenen, modularisierten Netzwerken. Andere Konventionen zählen die Neuronenschichten von Input- bis Outputschicht, oder aber zählen die Input-Schicht nicht mit, da in ihr keine Informationsverarbeitung im eigentlichen Sinne erfolgt. Aufgrund der uneinheitlichen Bezeichnungen auf der einen Seite und der Bedeutung der Topologie auf der anderen erscheint es notwendig, letztere vollständig zu beschreiben, um Missverständnissen vorzubeugen und eine nachträgliche Simulation bzw. Replikation von Experimenten durch Dritte zu ermöglichen. Sie wird daher wie folgt spezifiziert: Die Netzwerktopologie wird durch die Anzahl von Neuronen N in den einzelnen Schichten festgelegt. Die Anzahl der Eingabeneuronen N_I entspricht der Anzahl simultan verarbeiteter Eingabewerte und determiniert die Länge des Inputvektors. Die Anzahl der versteckten Schichten legt die Komplexität abbildbarer Systemzustände fest, während die Anzahl der Neuronen N_H je Schicht $N_{H1}, N_{H2}, \dots, N_{HN}$ die Komplexität der Verarbeitung festlegt.² Die Anzahl von Neuronen in der Ausgangsschicht N_O spezifiziert die Anzahl simultan ausgegebener Netzwerkausgaben und somit die Länge des Ausgabevektors. Zusammenfassend wird die Topologie T eines NN durch ein Tupel der Neuronenanzahl in den einzelnen Schichten, mit $T = [N_I - N_{H1} - N_{H2} - \dots - N_O]$, ausgehend von der Eingabeschicht hinreichend beschrieben.

Abweichend von der oben modellierten Netzwerkarchitektur kann die Funktionalität des Schwellwertes θ_j aller Neuronen $u_j, j = 1, 2, \dots, m$, der verdeckten Schichten und der Ausgangsschicht alternativ als ein Schwellwert-Neuron (engl. *on-* oder *bias-node*) u_0 modelliert werden, dessen Ausgabe konstant eins ist, $o_0 = 1$, und das mit allen Neuronen u_j mit Schwellwerten über trainierbare Verbindungsgewichte verbunden ist [Zell 2000, S. 82].³

a.)



b.)

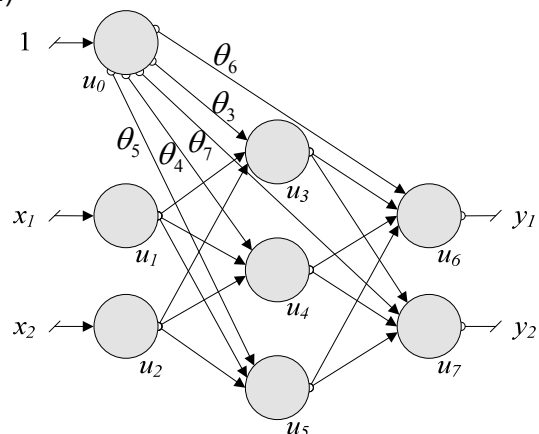


Abbildung 4.9: Modellierung der Schwellwerte der Aktivierung durch ein Bias-Neuron

Dies erlaubt sowohl eine vereinfachte mathematische Darstellung als auch eine vereinfachte Parametrisierung des NN ohne eine separate Berechnung der trainierbaren Schwellwerte in

¹ ROSENBLATT [1958] bezeichnete die Schichten ursprünglich als „sensory layer“, „associative layer“ und „response layer“.

² Es werden nur selten mehr als drei Schichten versteckter Neuronen benötigt, selbst zur Modellierung komplexer, nichtlinearer Interaktionen. Dies ist jedoch für den Anwendungsfall der Prognose zu bestätigen.

³ Es ist zu beachten, dass diese Formulierung ausschließlich für den Sonderfall einer additiven Eingabefunktion Gültigkeit besitzt.

der Lernphase [Anderson 1995, S. 56]. Auf eine explizite Erweiterung der Notation in der topologischen Darstellung als Graph oder in der Verbindungsmatrix wird dabei meist verzichtet.

4.2.4.2 Struktur der Verbindungsgewichte

Die Abbildung des Verbindungsnetzwerkes der Zellen erfolgt anhand der Topologie des gewichteten Graphen oder in einer Gewichtsmatrix aller Verbindungsgewichte W^1 , wobei der Eintrag w_{ij} die Art und Stärke der künstlichen synaptischen Verbindung zwischen dem sendenden Neuron u_i zum empfangenden Neuron u_j bestimmt. Der absolute Betrag des Verbindungsgewichtes $|w_{ij}|$ spezifiziert die Stärke der übertragenen Signale, das Vorzeichen des Verbindungsgewichtes seine Wirkung durch Verstärkung der Aktivierung bei positivem Vorzeichen (als biologisches Äquivalent einer exzitatorischen Synapse) oder durch Hemmung der Aktivierung bei negativem Vorzeichen (im Sinne einer inhibitorischen Synapse) [Zell 2000]. Alle Verbindungen bezeichnen gerichtete, unidirektionale Verbindungen², so dass für einen Eintrag in der Matrix der Verbindungsgewichte W gilt:

$$W = [w_{ij}]$$

mit

$$\begin{aligned} w_{ij} = 0 & \quad \text{keine Verbindung zwischen Neuron } i \text{ und } j \\ w_{ij} < 0 & \quad \text{Neuron } i \text{ hemmt seinen Nachfolger } j \text{ durch ein Gewicht der Stärke } |w_{ij}|. \\ w_{ij} > 0 & \quad \text{Neuron } i \text{ erregt seinen Nachfolger } j \text{ durch ein Gewicht der Stärke } |w_{ij}|. \end{aligned} \quad (4.21)$$

Verbindungsgewichte werden als reelle numerische Werte, $w_{ij} \in \mathbb{R}$, realisiert.³ Die Verbindungsmatrix spezifiziert neben der grundlegenden Topologie des NN auch die Struktur des Informationsflusses im Netzwerk und somit die Art der Informationsverarbeitung. Anhand der Struktur der Verbindungsmatrix unterscheidet man Netze ohne Rückkopplungen der Verbindungsstruktur, auch als vorwärtsgerichtete neuronale Netze bezeichnet, und Netze mit Rückkopplungen, die auch als rekurrente Netze bezeichnet werden [Zell 2000, S. 78]. Beide Netztypen können gleichermaßen in unterschiedlichen Anwendungen einschließlich der Prognose eingesetzt werden, bilden jedoch ein signifikant abweichendes Systemverhalten ab und erfordern unterschiedliche Ansätze bei der Parametrisierung.

In NN ohne Rückkopplungen existiert kein Pfad, der von einem Neuron u_i direkt oder indirekt über zwischengeschaltete Neuronen wieder zum Neuron u_i zurückführt. In Anlehnung an die unidirektionale Informationsverarbeitung von der Eingabeschicht zur Ausgabeschicht werden sie auch als vorwärtsgerichtete NN (engl. *feed forward neural networks*) bezeichnet und in der Notation als $W = \text{FF}$ spezifiziert. Mathematisch entspricht dies einem azyklischen Graph; in der Matrixschreibweise ist lediglich die obere Dreiecksmatrix mit Werten $w_{ij} \neq 0$ besetzt. Innerhalb der Klasse vorwärtsgerichteter NN unterscheidet ZELL [2000, S. 78] ebenelementar verbundene NN (mit Verbindungen von jeweils einer Schicht zur unmittelbar nachfolgenden) und generalisierte oder auch allgemein vorwärtsgerichtete Netze mit schichtübergreifenden

¹ Darüber hinaus werden anregende und hemmende Gewichte zur Effizienzsteigerung in Software- oder Hardwaresimulatoren häufig durch separate Matrizen getrennt gespeichert [Zell 2000, S. 76], zur Verringerung des Speicherbedarfs durch eine Zeigerstruktur aus rückwärtsgerichteten Adjazenzlisten abgebildet [Zell 2000, S. 77] oder in separaten Matrizen für Verbindungsstruktur und -gewichte modelliert [Alex 1998]. Die Matrixschreibweise dient nachfolgend der vereinfachenden Abbildung.

² In Abweichung dazu stellen einige Autoren die Verbindung als eine bidirektionale dar, siehe z. B. [Alex 1998, S. 83]. Der korrespondierende Fluss der Lerninformationen eines NN verläuft zwar entgegengerichtet zu der Verbindungsrichtung der Kanten, hat jedoch auf die Verarbeitung der Information im NN keine Auswirkungen und wird somit nicht im Graph abgebildet [Alex 1998, S. 83].

³ Das komplexe biologische, dynamische Verhalten mit nichtlinearen Interaktionen innerhalb der Synapsen sowie die frequenzkodierte Übertragung von Aktionspotenzialen der Neuronen [Anderson 1995, S. 95] wird bei der Modellierung künstlicher NN zur Anwendung außerhalb des biologischen Kontextes vernachlässigt.

Verbindungen (engl. *shortcut connections*), die einzelne Schichten überspringen können, um zusätzliches Systemverhalten zu modellieren. In Abbildung 4.10. sind beide Netzwerktopologien sowohl als gerichteter Graph als auch als Gewichtsmatrix dargestellt.¹

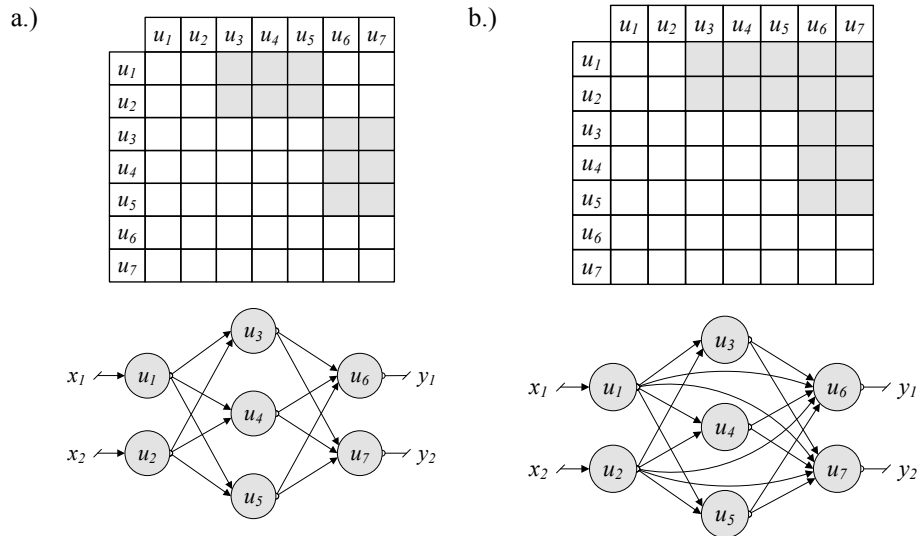


Abbildung 4.10: Topologien und Verbindungsmatrizen für vorwärtsgerichtete NN ohne (a.) und mit schichtübergreifenden Verbindungen (b.)²

Die Modellierung schichtübergreifender Verbindungen $W = SC$ in Abbildung 4.10 b.) erfolgt zur simultanen Überlagerung linearer und nichtlinearer Komponenten des Eingavektors auf den Ausgabevektor zur Abbildung eines komplexen Systemverhaltens [Zell 2000].³

Rückkopplungen innerhalb der Netzwerkstruktur werden in direkte Rückkopplungen, indirekte Rückkopplungen zwischen Schichten und laterale Rückkopplungen innerhalb einer Schicht unterteilt [Zell 2000, S. 78]. Topologien mit direkten Rückkopplungen $W = DF$ (engl. *direct feedback*) erlauben über die direkte Verbindung der Ausgabe eines Neurons zu seiner Eingabe die Verstärkung oder Abschwächung der eigenen Aktivierung. Die Fähigkeit, durch direkte Eigenhemmung oder Selbstverstärkung die Grenzzustände der Aktivierung anzunehmen, wird durch eine besetzte Hauptdiagonale der Verbindungsmatrix in Abbildung 4.11 a.) abgebildet. NN mit indirekten Rückkopplungen $W = IF$ (engl. *indirect feedback*) besitzen Rückkopplungen von Neuronen höherer Schichten zu Neuronen niedriger Schichten, um z. B. im Sinne einer selektiven Mustererkennung die Aufmerksamkeitssteuerung auf bestimmte Bereiche von Eingabeneuronen oder auf bestimmte Eingabemerkmale durch das Netz zu lenken [Zell 2000, S. 78], siehe Abbildung 4.11 b.). Als Sonderform indirekter, rekurrenter Verbindungen in vorwärtsgerichteten Netzen gelten auch Kontextzellen, also separate verdeckte Zellen zur temporären Speicherung von Aktivierungen, die Eingaben von versteckten Neuronen und Ausgabeneuronen erhalten und diese verarbeitet an das Netzwerk zurückleiten [Zell 2000, S. 138]. Netztopologien mit lateralen Rückkopplungen innerhalb einer Schicht $W = LF$ (engl. *lateral feedback*) in Abbildung 4.11 c.) [Lohrbach 1994, S. 15] erlauben durch die Verwendung hemmender Verbindungen die Modellierung selbstregelnder Prozesse, in denen nur ein einziges Neuron einer Schicht aktiviert wird und alle anderen

¹ Während die Notation der Verbindungsmatrix eine übersichtliche Abbildung insbesondere rekursiver Verbindungen erlaubt, können in gerichteten Graphen zusätzlich Informationen der Verarbeitungsfunktionen etc. angegeben werden. Aus diesem Grund werden NN nachfolgend als gerichtete Graphen abgebildet.

² Für zwei verbundene Neuronen u_i, u_j wird der Ursprung der Verbindung u_i in der Spalte und der Empfänger der Verbindung u_j in der Zeile angegeben.

³ Die Überlagerung linearer und nichtlinearer Komponenten erfolgt durch die Verwendung einer linearen Aktivierungsfunktion in den Neuronen der Ausgabeschicht und von nichtlinearen Aktivierungsfunktionen in den versteckten Neuronen, so dass die schichtenweise Berechnung nichtlineare, die der direkten Verbindungen lineare Zusammenhänge abbilden kann.

inaktiv werden, im Sinne biologisch motivierter ‘‘Winner-takes-all‘‘-Netzwerke [Anderson 1995, S. 85-127; Zell 2000, S. 78]. Eine Übersicht gibt Abbildung 4.11.

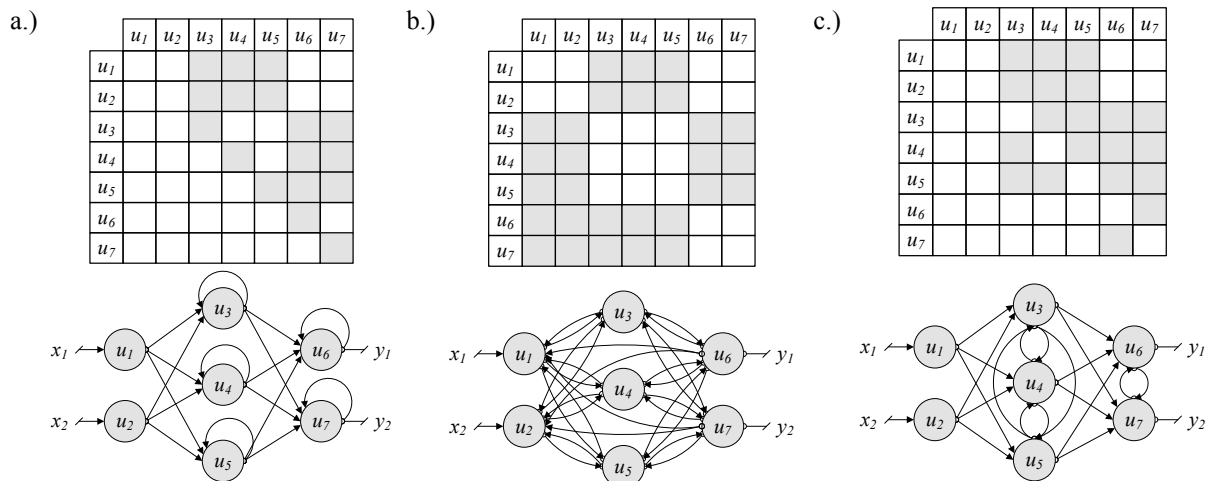


Abbildung 4.11: Topologien und Verbindungsmatrizen für rückgekoppelte NN mit direkten (a.), indirekten (b.) und lateralen Rückkopplungen (c.).

Durch Kombination der Rückkopplungsarten zwischen allen Neuronen ergeben sich vollständig verbundene rekurrente NN. Sie besitzen eine vollständig ausgefüllte Verbindungsmatrix, mit Ausnahme der Eingabeschicht, die lediglich Eingangssignale an das NN anlegt und typischerweise weder laterale noch direkte Rückkopplungen besitzt [Zell 2000, S. 79]. Eine eindeutige Zuordnung einzelner Neuronen zu versteckten Schichten ist bei Auftreten von Rückkopplungen nicht länger möglich, was die Komplexität der Informationsberechnung und die Beschreibung der Verbindungsstruktur erheblich erhöht. Rückkopplungen erfordern somit die Berücksichtigung einer Reihenfolge zur Berechnung in Form einer Aktivierungsstrategie.

4.2.4.3 Aktivierungsstrategien der Informationsverarbeitung

Die Aktivierungsstrategie K spezifiziert die Reihenfolge der einzelnen Berechnungsschritte zur Informationsverarbeitung in NN. Im Gegensatz zu biologischen neuronalen Netzen mit simultaner, synchroner und massiv paralleler Berechnung von Aktivierungspotenzialen erfolgt die Berechnung in künstlichen NN in diskreten Zeitschritten in Hardware und Software [Zell 2000].

Grundsätzlich wird nach Anlegen eines Eingabemusters als n -dimensionaler Eingabevektor (x_1, x_2, \dots, x_n) an die Neuronen der Eingabeschicht die Information unter Berechnung von Aktionspotenzialen über die Neuronen der versteckten Schichten bis an die Neuronen der Ausgabeschicht zur Ausgabe eines m -dimensionalen Ausgabevektors (y_1, y_2, \dots, y_m) weitergeleitet [Rojas 1993, S. 29], ein Vorgang, der auch als Vorwärts-Propagierung (engl. *forward propagation* oder *forward pass*) bezeichnet wird. Dabei erhält die Reihenfolge, in der Neuronen ihre Netzwerkeingabe erhalten und in der sie ihre Ausgaben berechnen und weiterleiten, eine entscheidende Bedeutung [Zell 2000, S. 87]. Da vorwärtsgerichtete Netzwerkarchitekturen durch das Fehlen direkter, indirekter oder lateraler Rückkopplungen keine zeitliche Entwicklung früherer Aktivierungszustände vor- oder nachgelagerter Neuronen berücksichtigen, ist die Aktivierung ausschließlich von den momentanen Aktivitätsniveaus der vorgeschalteten Neuronen abhängig. Da jedes Neuron eine funktionale Transformation seiner Eingabewerte x in Aktivierungszustände $g(x)$ berechnet, kann ein stark vereinfachtes NN mit sequenzieller Aktivierungsstrategie als schrittweise Berechnung von Funktionen $f(g(x))$ betrachtet werden, wie in Abbildung 4.12 dargestellt. Ein NN

bildet entsprechend beliebige Funktionen des Inputvektors durch sequenzielle Transformation auf den Ausgabevektor ab, mit $F : R^n \rightarrow R^m$ [Rojas 1993, S. 29].

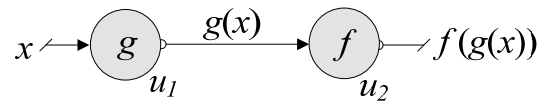


Abbildung 4.12: Funktionsverarbeitung in vorwärtsgerichteten Neuronalen Netzen¹

Aufgrund der topologischen Ordnung aller Neuronen in Schichten existiert eine eindeutig determinierte, schrittweise Berechnungsvorschrift der schichtweisen Aktivierungsreihenfolge von der Eingabeschicht zur Ausgabeschicht. Aus einem anliegenden Eingabevektor kann jederzeit ein eindeutiger Aktivierungszustand des gesamten NN berechnet werden. Diese unidirektionale Verarbeitungsrichtung des Informationsflusses in vorwärtsgerichteten NN begünstigt eine besonders effiziente Berechnung in Software-Simulatoren [Zell 2000, S. 88].

Dies gilt nicht für rekurrente Netzwerke², in denen der rückgekoppelte Informationsfluss topologisch nicht eindeutig festgelegt ist und spezifische Aktivierungsstrategien die Aktivierungsreihenfolge bestimmen müssen, um die Entstehung stabiler Systemzustände zu erlauben [Zell 2000]. Der Zustand eines einzelnen Neurons $g(x)$ hängt somit von seinem vorherigen Zustand $g(x_t, g(x_{t-1}))$ und aller vorangegangenen Zustände ab. Der Zustand eines Neurons $f(g(x))$ kann sich so mit jedem Berechnungsschritt einer rekursiven Schleife in Abhängigkeit seiner eigenen Aktivierungsreihenfolge $f(g(x_t, g(x_{t-1}, \dots)))$ und der Aktivierungsreihenfolge anderer Neuronen $g(x_t, g(x_{t-1}))$ verändern, und so das Netz in unterschiedliche stabile Zustände konvergieren.

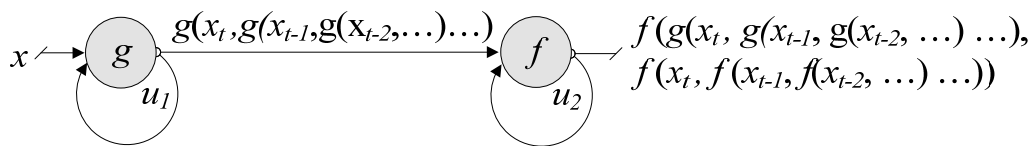


Abbildung 4.13: Funktionsverarbeitung in rekurrenten Neuronalen Netzen³

Die asynchrone Berechnung der Aktivierungszustände kann hingegen nach einer festen Ordnung, einer zufälligen Ordnung oder einer zufälligen Permutation spezifiziert werden und wird als Propagierungsschema oder Aktivierungsstrategie bezeichnet. Eine Diskussion dynamischer, synchroner und asynchroner Aktivierungsstrategien findet sich bei [Zell 2000, S. 87-89].

4.2.5 Ausgewählte Architekturen

Durch Kombination der untersuchten Modellierungsparameter der Verarbeitungsfunktionen zur Signalverarbeitung im Neuron V mit den Parametern ihrer Anordnung in der Topologie T zu einem Netzwerk sowie der Verbindungsstruktur der Gewichtsmatrix W und der Aktivierungsreihenfolge K wird die Architektur eines NN spezifiziert. Unter der Architektur eines NN wird „das Minimum der im Rahmen eines Gestaltungs- oder Bewertungsprozesses zu erfassenden Merkmale verstanden, die ein konkretes NN von anderen Netzen unterscheidet und seine im Hinblick auf die Problemstellung notwendigen Eigenschaften beschreibt“ [Alex 1998, S. 103]. Sie spezifiziert entsprechend alle Parameter der Modellierung.

¹ Eigene Darstellung nach ROJAS [1993]

² ZELL [2000, S. 83] spricht im Fall zeitlichen Verknüpfung und expliziten Speicherung vorangegangener Aktivierungszustände auch von einer „Gedächtnisfunktion“ der Aktivierungszustände

³ Eigene Darstellung in Anlehnung an ROJAS [1993].

Eine spezifische Netzwerkarchitektur A wird hinreichend durch die Freiheitsgrade der Modellierung V , T , W und K spezifiziert und kann in Erweiterung der Notation von ALEX und STAHLBOCK [Alex 1998; Alex und Stahlbock 1996] als Tupel $A=[V, T, W, K]$ abgebildet werden. Die Architektur wird durch die Teilentscheidungen der Verarbeitungsfunktionen $V=[F_I, F_H, F_O]$ je Neuron sowie die Topologie $T=[N_I-N_{H1}-N_{H2}-\dots-N_O]$ der einzelnen Schichten genauer beschrieben. Anstelle der Spezifikation einer Verarbeitungsfunktion individuell für jedes Neuron wird nachfolgend eine identische Signalverarbeitung aller Neuronen einer Schicht unterstellt. Da in Neuronen der Eingabeschicht einheitlich keine originäre Informationsverarbeitung stattfindet, wird sie vielfach vernachlässigt. Jedoch ist die Architektur hinreichend durch das Tupel $A=[[N_I-N_{H1}(F_I, F_A, F_O)-N_{H2}(F_I, F_A, F_O)-\dots-N_O(F_I, F_A, F_O)], W, K]$ spezifiziert.

Durch Kombination der Gestaltungsmöglichkeiten haben sich spezifische Netzarchitekturen in Wissenschaft und Praxis etabliert, die unterschiedliche Anwendungspotenziale für die Prognose besitzen. Spezifische Netzwerkarchitekturen werden synonym als Modellklassen oder Netzwerkparadigmen bezeichnet. Unter Berücksichtigung der vorangehend entwickelten als grafischen Notation werden in Abbildung 4.14 zwei historische Architekturen als gerichtete Graphen dargestellt.

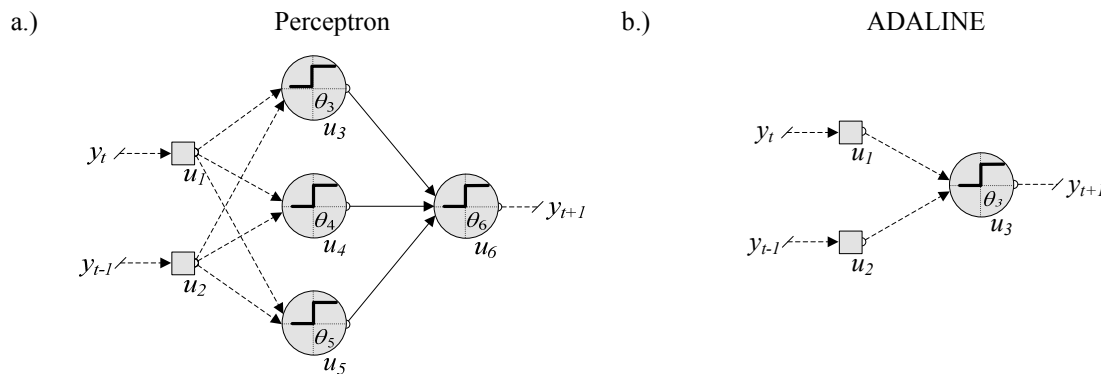


Abbildung 4.14: Netzwerkarchitektur des Perceptron (a.) und ADALINE (b.)

Die historische Modellklasse der Perceptrons in Abbildung 4.15 a.) stellt eine einfache Feed-forward-Netzwerkarchitektur mit einer Schicht trainierbarer Verbindungsgewichte und mit binären Aktivierungsfunktionen dar [Rosenblatt 1958; Minsky und Papert 1988, S. 97-103; Anderson 1995, S. 209-238; Zell 2000].¹ Das ADALINE in Abbildung 4.15 b.) repräsentiert eine frühe NN-Architektur nach WIDROW [1990], die erhebliche Aufmerksamkeit erhielt und zu deren Parametrisierung der Delta-Lernalgorithmus als Vorläufer heutiger Ansätze zur Parametrisierung von NN entwickelt wurde (siehe Abschnitt 4.3.2.1).

Darauf aufbauend wurden aus den gleichen Grundbausteinen jene mehrlagigen Architekturen nichtlinearer NN entwickelt, die heute überwiegend angewendet werden. In Abbildung 4.15 wird eine Übersicht verbreiteter Netzwerkarchitekturen des Multilayer Perceptron, Generalised Feed Forward NN, radiales Basisfunktionsnetz, Hyper-Basisfunktionsnetz, Probabilistisches NN und eines rekurrenten Jordan-NN als gerichtete Graphen präsentiert.

¹ Beide historischen Netzwerkarchitekturen entziehen sich der Notation aufgrund der historischen Verwendung nicht-trainierbarer Verbindungsgewichte. Näherungsweise kann das Perceptron durch $A=[[2-3(\text{Sum}, \text{Bin}, \text{Id})-1(\text{Sum}, \text{Bin}, \text{Id})], \text{FF}, \text{F}]$ und das ADALINE durch $A=[[2-1(\text{Sum}, \text{Bin}, \text{Id})], \text{FF}, \text{F}]$ spezifiziert werden.

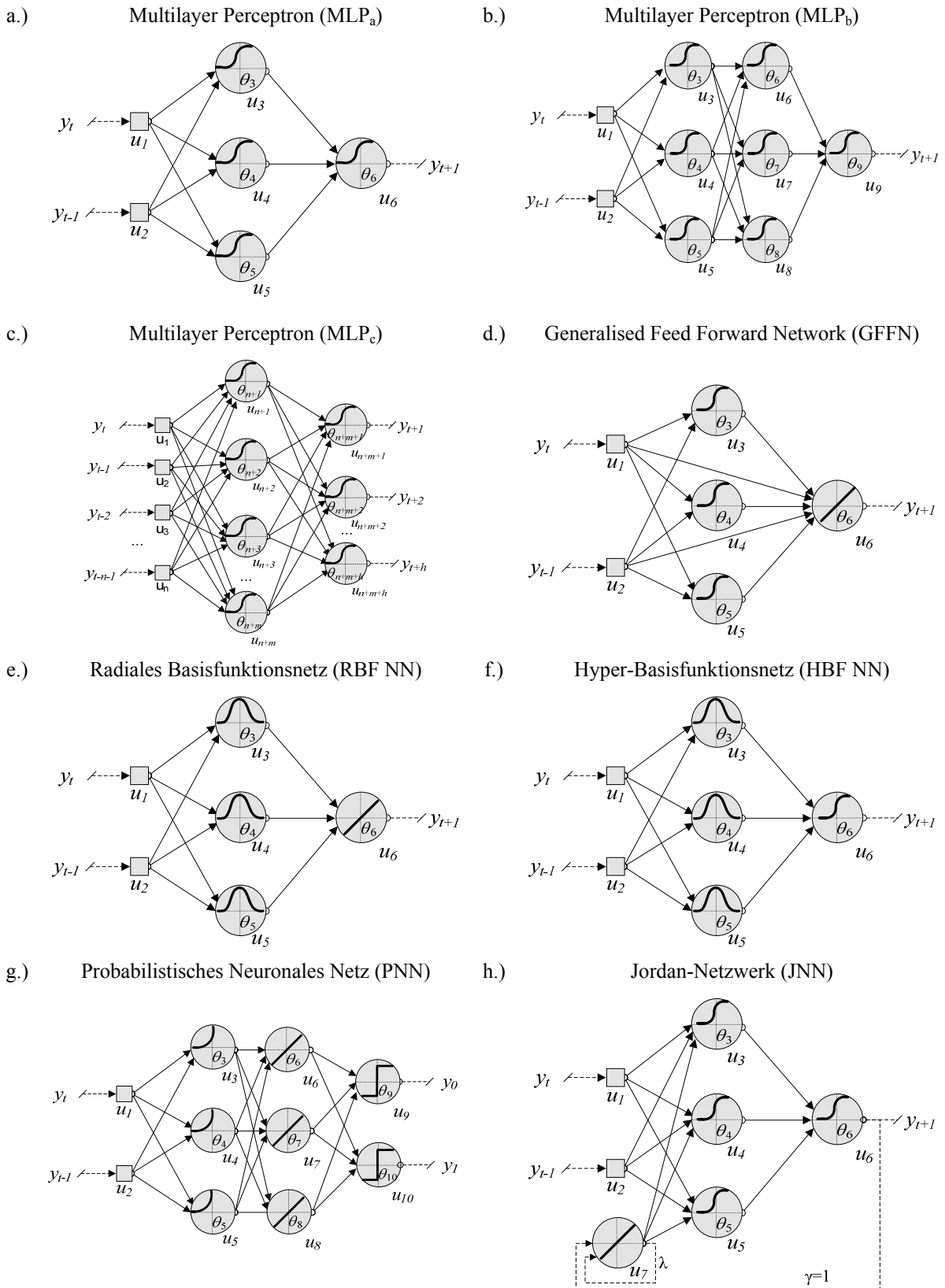


Abbildung 4.15: Prototypische Netzwerkarchitekturen des MLP (b.), (c.) und (d.), GFFN (d.), RBF NN (e.), HBF NN (f.), PNN (g.) und JNN (h.)

Ausgehend von dem einlagigen Perceptron wurde von RUMELHART et al. [1986] die Architektur auf mehrlagige vorwärtsgerichtete NN erweitert, sog. Multilayer Perceptrons (MLP), siehe

Abbildung 4.15 a.) bis c.). Das MLP¹ repräsentiert eine mächtige Klasse vorwärtsgerichteter NN mit diversen Freiheitsgraden der Modellierung durch beliebig viele Schichten von Neuronen, einer beliebigen Anzahl von Neuronen je Schicht, einer großen Auswahl nichtlinearer, sigmoider Aktivierungsfunktionen $F_A = \{\text{TanH}, \text{Log}, \text{Id}\}$ und diversen theoretischen Weiterentwicklungen und Erweiterungen der Netzwerkarchitektur.² MLPs lassen sich hinreichend durch die folgende Notation beschreiben:

$$A = [[N_{I-N_{HI}}(F_I, F_A, F_O) - N_{H2}(F_I, F_A, F_O) - \dots - N_O(F_I, F_A, F_O)], W, K] \quad (4.22)$$

Beispielsweise wird das MLP in Abbildung 4.15 a.) durch

$$A = [[2-3(\text{Sum}, \text{Log}, \text{Id})-1(\text{Sum}, \text{Id}, \text{Id})], \text{FF}, \text{F}] \quad (4.23)$$

und das MLP in Abbildung 4.15b.) durch

$$A = [[2-3(\text{Sum}, \text{Log}, \text{Id})-3(\text{Sum}, \text{Log}, \text{Id})-1(\text{Sum}, \text{Id}, \text{Id})], \text{FF}, \text{F}] \quad (4.24)$$

hinreichend spezifiziert. Seit der Einführung des MLP konzentriert sich ein Großteil der wissenschaftlichen Forschung und der praktischen Anwendungen auf diese Netzwerkarchitektur beschränkt, auch im Bereich der Prognose (vgl. Abschnitt 4.4.1.2).

Generalisierte vorwärtsgerichtete NN (engl. *Generalised Feed Forward Neural Network*, GFNN) in Abbildung 4.15 d.) erweitern die Architektur der MLP durch ebenenübergreifende Verbindungen, welche eine Überlagerung nichtlinearer und linearer Interaktionen zwischen Eingabe- und Ausgabevektor in der Modellierung berücksichtigt. Die Notation ihrer Architektur A unterscheidet sich lediglich durch die Spezifikation ihrer Verbindungsmatrix $W = \text{FF} + \text{S}$ bei identischer Bezeichnung von T und S wie für das MLP.³

Die Modellklassen der Radialen Basisfunktionsnetze (RBF NN) in Abbildung 4.15c. und ihre Erweiterung zu Hyper-Basisfunktionsnetzen (HBF NN) in Abbildung 4.15d. repräsentieren spezielle Klassen vorwärtsgerichteter NN mit einer Schicht versteckter Neuronen, die aus der mathematischen Approximationstheorie von Funktionen entwickelt wurde [Zell 2000, S. 226]. Sie kombinieren radialsymmetrische Basisfunktionen in den versteckten Neuronen zur Interpolation und Approximation einer mehrdimensionalen Funktion anhand ihrer Stützstellen und erlauben so die Extrapolation im Sinne der Zeitreihenprognose. Die wissenschaftliche Untersuchung und praktische Anwendung von RBF-Netzen zur Prognose lieferte bislang nur inkonsistente Erkenntnisse [Sarle 1994]. Die erweiterte Modellierung kausaler Interdependenzen durch Interventionsvariablen exogener Sondereinflüsse oder dynamischer Regressionsmodelle sind nicht ohne Weiteres abzubilden und nicht untersucht. Daher werden RBF-Netze nachfolgend zugunsten populärer und mächtigerer Modellklassen in der Prognose vernachlässigt. Eine Übersicht über die Theorie, die Anwendung und Erweiterungen verallgemeinerter RBF-Netze geben [Poggio und Girosi 1990; Zell 2000, S. 225-240].⁴

Vollständig und partiell rekurrente Netzwerkarchitekturen mit lateralen sowie direkten Rückkopplungen werden in Architekturen zur multikriteriellen Klassifikation eingesetzt,

¹ MLPs werden von einigen Autoren fälschlicherweise nach dem zu ihrer Parametrisierung entwickelten, jedoch auch für andere Netzwerkarchitekturen anwendbaren Lernverfahren auch als Backpropagation-Netze [Rojas 1993] oder allgemein als Feed-Forward Neural Networks (FFNN) bezeichnet.

² Aufgrund ihrer flexiblen Modellierung können MLPs sowohl zur zeitreihenanalytischen als auch zur kausalanalytischen Modellierung in der Prognose angewendet werden, wie nachfolgend noch detailliert gezeigt wird.

³ Sie sind nicht mit den Architekturen sog. Generalised Regression Neural Networks (GRNN) zu verwechseln; GRNN ist die ursprüngliche Bezeichnung von SPECHT für die nachfolgend dargestellten Radialen Basisfunktionsnetze [Zell 2000].

⁴ Eine Untersuchung von RBF-Netzen in der Prognose erfolgt u.a. bei [Chiu 1994; Duncan 1996; Wu 1998].

weniger zur Prognose. Dagegen sind rekurrente NN mit indirekten Rückkopplungen zur Prognose von Zeitreihen geeignet [Mozer 1994; Pearlmutter 1990] und werden vielfach gegenüber vorwärtsgerichteten Architekturen präferiert, so z. B. durch ZIMMERMANN et al. [2000a] oder auch WERBOS [1990; 1996; 1988b].

Partiell rekurrente Netzwerkarchitekturen wie Jordan-Netzwerke in Abbildung 4.15 h.) oder hierarchische Elman-Netzwerke haben aufgrund von Problemen in ihrer Spezifikation jedoch nur eine geringere empirische Aufmerksamkeit erfahren. Insbesondere werden Probleme in der Vermeidung von Oszillationen in rekursiven Schleifen, der hohen Komplexität der Architekturentscheidungen zur Bestimmung geeigneter Topologien und der Konfiguration und Kontrolle der Lernverfahren angeführt. Rekurrente NN erfordern eine detaillierte Betrachtung von Architektur und Lernverfahren, die eine Darstellung vorwärtsgerichteter NN übersteigt. Da der Schwerpunkt dieser Untersuchung auf einer Erweiterung der Modellierung von NN auf asymmetrische Zielfunktionen liegt, unterbleibt hier die ansonsten durchaus sinnvolle Betrachtung dieser wichtigen Modellklasse.¹ Entsprechend beschränken sich die nachfolgenden Ausführungen auf die populärste Klasse vorwärtsgerichteter Netzwerkarchitekturen, das MLP. Eine detaillierte Diskussion der Grundlagen rekurrenter Netze findet sich in [Medsker und Jain 2001; Michel und Liu 2002] sowie im inhaltlichen Zusammenhang der Prognose in [Rangadass 1996].²

Die Architektur A eines NN bestimmt durch die Auswahl von $A=[V, T, W, K]$ maßgeblich seine Repräsentierfähigkeit³, also die „Fähigkeit eines Netzes, eine gegebene Funktion ... mit dem neuronalen Netz realisieren zu können“ [Zell 2000]. So sind z. B. NN-Architekturen mit nur einer Schicht trainierbarer Gewichte lediglich in der Lage, einfache nichtlineare Funktionen zu approximieren bzw. linear separierbare Probleme zu lösen [Minsky und Papert 1969]. Dagegen sind Architekturen mit mindestens einer Schicht verdeckter Neuronen bzw. zwei oder mehr Lagen trainierbarer Gewichte auch in der Lage, komplizierte nichtlineare, diskontinuierliche und lineare Muster sowie konvexe und konkave Polygone korrekt abzubilden [Reed und Marks 1999; Zell 2000, S. 99-102]. Im Gegensatz zu seiner Repräsentierfähigkeit bezeichnet die Lernfähigkeit eines NN hingegen die Fähigkeit eine repräsentierbare Funktion aus den Daten zu erlernen, indem die Gewichte und Schwellwerte bei gegebener Architektur durch einen Lernalgorithmus korrekt bestimmt werden, im Sinne der Parametrisierung eines Verfahrens.⁴ Netzwerkarchitektur und Lernalgorithmus müssen aufeinander abgestimmt sein, um ein gegebenes Anwendungsproblem repräsentieren und aus den Daten lernen zu können.⁵ Der Ablauf und die relevanten Verfahren zur Parametrisierung von MLPs für Prognosen werden daher nachfolgend untersucht.

¹ Die Analyse und die Erkenntnisse zum Training von NN mit asymmetrischen Kostenfunktionen werden jedoch allgemein gehalten und lassen sich so auf rekurrente Architekturen übertragen. Sie sind dort jedoch erneut theoretisch sowie in einem empirischen Verfahrenvergleich zu überprüfen.

² Modellklassen vorwärtsgerichteter Netze mit binären Ausgabewerten umfassen z. B. Probabilistische Neuronale Netze (PNN) (siehe Abbildung 4.15e), Varianten der lernenden Vektorquatisierung (LVQ) Netze [Stahlbock 2002; Kohonen 1989; Zell 2000; Somervuo und Kohonen 1999], selbstorganisierende Karten [Kohonen 1982; Kohonen 1985; Kohonen 2001] oder Counterpropagation-Netze [Hecht-Nielsen 1988; Hecht-Nielsen und McKenna 2003]. Diese Netzwerkarchitekturen wurden speziell zur Klassifikation bzw. Vorhersage von nominal skalierten abhängigen Variablen entwickelt und sind nicht zur Zeitreihenprognose mit metrischen Prognosewerten nicht anwendbar. Sie werden somit nicht näher untersucht.

³ Sie wird von ZELL [2000, S. 99] abweichend als „Repräsentierbarkeit“ bezeichnet.

⁴ Dies wird nach ROSENBLATT auch als „perceptron convergence theorem“ bezeichnet [Zell 2000, S. 99].

⁵ Diese Darstellung vernachlässigt allerdings die Bedeutung des Lernzieles, welche durch die vielfach im Lernalgorithmus implizit spezifizierte Zielfunktion repräsentiert wird. Diese ist im Anschluss an die Darstellung von NN zur Prognose zu diskutieren.

4.3 Parametrisierung von Neuronalen Netzen

4.3.1 Grundlagen des Lernens

Nachdem ein NN durch eine Architektur aus den Verarbeitungsfunktionen der Neuronen, ihre Topologie, die Verbindungsstruktur und die Aktivierungsstrategie spezifiziert ist, wird das NN zur Anwendung in einer Aufgabenstellung, z. B. für eine Zeitreihenprognose auf einem spezifischen Datensatz parametrisiert. In Anlehnung an die Vorgänge des Lernens zur Speicherung von Erfahrungen in biologischen NN durch die Veränderung von Struktur und Stärke der synaptischen Verbindungen wird die Parametrisierung von NN in Form einer Anpassung der Verbindungsgewichte durchgeführt, ein Vorgang, der auch als Lernen oder Training bezeichnet wird.

Die Parametrisierung von NN erfolgt mit Hilfe architektur- und anwendungsspezifischer Lernalgorithmen. "The procedure used to perform the learning process is called learning algorithm, the function of which is to modify the synaptic weights of the network in an orderly fashion to attain a desired objective" [Haykin 1999, S. 2]. Der Lernalgorithmus, auch als Lernverfahren oder Lernregel bezeichnet [Zell 2000, S. 83], bestimmt die freien Parameter eines NN direkt aus den präsentierten Daten. Die Findung der Parameter erfolgt in einer der Anwendung vorausgehenden Lernphase durch Anwendung eines iterativen Lernalgorithmus [Rojas 1993], äquivalent zur Parametrisierung herkömmlicher mathematisch-statistischer Verfahren, z. B. der Regression, zur Anpassung an die Datenbasis. Ausgehend von einer vorher festgelegten NN-Architektur erfolgt die Parametrisierung durch die Bestimmung der freien Parameter der Verbindungsgewichte w_{ij} und der Schwellwerte θ_j . Lernen bezeichnet somit die Modifikation der Stärke der Verbindungen Δw_{ij} und des Schwellwertes des Neurons $\Delta \theta_j$ zur Abbildung unterschiedlicher Aufgaben bzw. funktionaler Zusammenhänge, ohne Veränderungen der Netzwerkarchitektur \bar{A} abzubilden.¹

Es existiert eine Vielzahl konkurrierender Ziele, Paradigmen und Algorithmen zur Durchführung des Lernprozesses; die in Abbildung 4.16. einander zugeordnet werden.

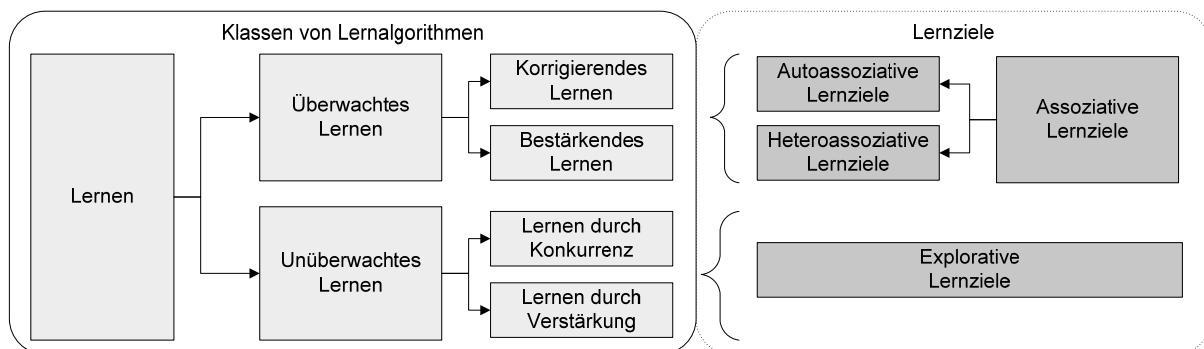


Abbildung 4.16: Zusammenhang von Lernzielen und Lernalgorithmen² [erweiterte Darstellung nach Rojas 1993]

¹ Darüber hinaus spezifiziert ZELL [2000, S. 84] architekturverändernde Lernarten: die Veränderungen der Verbindungsmatrix ΔW durch die Entwicklung neuer Verbindungen $w_{ij} \neq 0$ oder das Absterben bestehender Verbindungen $w_{ij} = 0$, außerdem die Veränderung der Netzwerktopologie im Lernprozess ΔT durch die Entwicklung neuer oder das Löschen bestehender Neuronen, und die Modifikation der Verarbeitungsfunktionen in den Neuronen ΔV . Diese Lernarten repräsentieren jedoch eher Ansätze einer dynamischen Optimierung der Netzwerkarchitektur durch Ausdünnung von Gewichten oder Neuronen (engl. *pruning*) nach Abschluss des Trainings oder auch evolutionäre Ansätze zur Findung geeigneter Netzarchitekturen im Sinne einer Methodologie zur Modellierung.¹ Das Ziel der gleichzeitigen Bestimmung einer optimalen Netzwerkarchitektur und einer Parametrisierung derselben [Zell 2000, S. 84] wird nachfolgend in den disjunkten Entscheidungen hinsichtlich der Bestimmung und der Optimierung der Netzwerkarchitektur im Modellierungsprozess diskutiert.

² Eigene, erweiterte Darstellung in Anlehnung an ROJAS [1993]

Nach KOHONEN [1989, S. 160] werden zunächst die unterschiedlichen Lernziele des assoziativen Lernens vom explorativen Lernen unterschieden, die durch die Aufgabenstellung der Anwendungsdomäne bestimmt werden und die Auswahl der Lernalgorithmen bestimmen [Levine 2000, S. 68]. Die unterschiedlichen Lernziele und Lernalgorithmen sind in Abbildung 4.16 einander zugeordnet. Im Fall des explorativen Lernens liegen dem Netz keine Informationen über das gewünschte Ausgabemuster vor [Levine 2000, S. 68]. Das Netz sucht einen stabilen Zustand aus den Eingabedaten, etwa in den Anwendungsfällen der Clusteranalyse oder der Optimierung, die hier aufgrund der abweichenden Problemstellung und wegen des Erfordernisses andersartiger Netzwerkarchitekturen nicht weiter betrachtet werden.

Durch ein assoziatives Lernen von Mustern verändert das NN die freien Parameter derart, dass ein NN nach wiederholter Präsentation von Paaren aus Eingabe- und Ausgabemustern die korrekte Assoziation für bekannte Muster selbstständig vornehmen kann und darüber hinaus unbekannte, ähnliche Eingabemuster im Sinne einer Generalisierung korrekt zuordnet [Zell 2000]. Ferner werden zwei Arten des assoziativen Lernens unterschieden: „the auto-associative recall, whereby an incomplete key pattern is replenished into a complete (stored) version, and the hetero-associative recall which selectively produces an output pattern \vec{y}_k in response to an input pattern \vec{x}_k ; in the latter case the paired associates \vec{x}_k and \vec{y}_k can be selected freely, independently of each other” [Kohonen 1989, S. 162] (siehe Abbildung 4.16).

In Anwendungen des heteroassoziativen Lernens sind die Vektoren der Eingabemuster und Ausgabemuster verschieden, $\vec{x}_k \neq \vec{y}_k$ [Anderson 1995, S. 170]. Dies entspricht der Modellierung und Parametrisierung konventioneller Prognosemodelle und -verfahren, da aus einer Menge verzögerter Realisationen abhängiger und unabhängiger Variablen der Eingabedaten eine oder mehrere zukünftige Realisationen der abhängigen Variablen vorhergesagt werden. Da konventionelle Anwendungen der zeitreihenanalytischen bzw. kausalanalytischen Punktprognose einem heteroassoziativen Lernziel entsprechen, werden sie nachfolgend intensiver untersucht. Netzwerktopologien zur Abbildung heteroassoziativer Lernaufgaben besitzen vielfach eine abweichende Anzahl von Eingabe- und Ausgabeneuronen $N^I \neq N^O$ zur Abbildung der heterogenen Datenvektoren, z. B. in den Anwendungen der einfachen oder der Mehrschritt-Prognose.

Anwendungen des autoassoziativen oder auch homoassoziativen Lernens mit $\vec{x}_k \rightarrow \vec{y}_k$ reproduzieren Eingabemuster auf identische Ausgabemuster. Homoassoziative Lernziele erfordern eine identische Anzahl von Neuronen in der Eingabe- und der Ausgabeschicht, $N^I = N^O$, mit dem Ziel der Erkennung, der Vervollständigung oder der Korrektur eines verrauschten Vektors der Eingabedaten in einen fehlerfreien Ausgabevektor [Alex 1998, S. 90]. Die Anwendung autoassoziativer Algorithmen zielen auf die Rekonstruktion verrauschter oder unvollständiger Muster aus Teilmustern, z. B. der Erkennung handgeschriebener Zeichen, der Restauration oder Rekonstruktion menschlicher Gesichter aus verschwommenen oder teilweise fehlenden Bildern etc. [Levine 2000, S. 68]. Die autoassoziative Modellierung von NN erlaubt keine unmittelbare Nutzung im Rahmen der Prognose zur Vorhersage zukünftiger, unbekannter Beobachtungswerte aus Realisationen unabhängiger und abhängiger Variablen mit unterschiedlichen Eingabe- und Ausgabevektoren $\vec{x}_k \neq \vec{y}_k$.¹

¹ Theoretisch sind auch für sie Ansätze der zeitreihenanalytischen Modellierung denkbar, in denen durch Bildung von Eingabe- und Ausgabevektoren identischer Länge $\vec{x}_k = \vec{y}_k$ eine Replikation gelernter Eingabemuster auf partiell vorhandene Ausgabemuster im Sinne teilweise vorhandener Zeitreihenelemente erfolgt. Ferner können autoassoziative Lernverfahren mit adäquaten Architekturen theoretisch in Problemstellungen der Zeitreihenanalyse und -prognose zur Glättung verrauschter Zeitreihen, zur Korrektur ungewöhnlicher Beobachtungswerte im Sinne der Ausreißerkorrektur oder zur Rekonstruktion fehlender Beobachtungswerte erforscht werden. Diese Ansätze sind bislang kaum untersucht, erfordern jedoch eine gesonderte Entwicklung und Untersuchung.

In Übereinstimmung mit den hetero- und den autoassoziativen Lernzielen können einzelne Klassen von überwachten und unüberwachten Lernalgorithmen unterschieden werden. Algorithmen des unüberwachten Lernens (engl. *unsupervised learning* oder *self-organised learning*) bilden Aufgaben des explorativen Lernens ab. Es existiert keine zugehörige Sollausgabe zu den Eingabemustern; der Lernalgorithmus versucht durch Selbstorganisation selbstständig, Gruppen bzw. Cluster statistisch ähnlicher Eingabevektoren zu identifizieren und diese auf Gruppen benachbarter Neuronen abzubilden. Unüberwachte Lernalgorithmen sind somit ebenfalls nicht für die Fragestellungen der Punktprognose geeignet.¹ Grundlagen und Beispiele des unüberwachten Lernens finden sich bei [Anderson 1995, S. 463-544].

Lernziele des assoziativen Lernens werden durch überwachte Lernalgorithmen abgebildet [Anderson 1995, S. 210]. Beim überwachten Lernen (engl. *supervised learning*) gibt ein externer „Lehrer“ zu jedem Eingabemuster der Trainingsmenge das korrekte bzw. beste Ausgabemuster an. Dabei können Ausgabemuster sowohl hetero- als auch homoassoziativer Art sein. Das bedeutet, dass dem neuronalen Netz immer gleichzeitig ein vollständig spezifiziertes Eingabemuster und das korrekte bzw. optimale, vollständig spezifizierte Ausgabemuster für die Eingabe vorliegen [Zell 2000, S. 93]. Aufgrund der Existenz eines externen Lehrers bzw. Trainers des NN wird auch von Training, Trainingsdaten und Trainingsprozessen etc. gesprochen. Die Ausgabe des Netzes wird beobachtet und im Fall einer falschen Ausgabe der Fehler berechnet. Man kann überwachte Lernalgorithmen weiter nach dem Zeitpunkt der Fehlerberechnung unterscheiden. Beim iterativen bzw. stochastischen Lernen (engl. *online-learning*) korrigiert der Lernalgorithmus die fehlerverursachenden Verbindungsgewichte nach jedem Anlegen eines Teilmusters aus der lokalen Differenz zwischen Ausgabe- und Sollwerten [Rojas 1993]. Im Gegensatz dazu aggregiert das kumulierte oder „stapelweise“ Lernen (engl. *batch-* oder *offline-learning*) alle einzelnen Fehler der Teilmuster eines Datensatzes und verwendet diesen Fehler zur Korrektur der Parameter.

Als Sonderfall des überwachten Lernens gilt das bestärkende Lernen (engl. *reinforcement learning*), auch als graduelles Lernen bezeichnet, das einige Autoren als eigenständige Lernart bezeichnen [Zell 2000; Hecht-Nielsen 1990]. Anstelle einer konkreten Zielausgabe für jedes Eingabemuster der Trainingsmenge wird anhand einer diskreten Kennziffer lediglich angegeben, ob die Netzausgabe aller Neuronen richtig oder falsch bzw. gut oder schlecht war [Hecht-Nielsen 1990, S. 49]. Ohne explizite Zielwerte für jedes einzelne Ausgabeneuron muss das NN die richtige Ausgabe selbst finden [Zell 2000, S. 95]. Beiden überwachten Lernarten ist gemein, dass die Soll-Ausgabe des NN für ein bestimmtes Muster bekannt ist und somit der Grad der Abweichung vom Sollwert angegeben werden kann. Dies entspricht der Nutzung von Rückkopplungen (engl. *feedback*) durch die Fehlerkorrektur (engl. *error correction*) im Training des NN zur Modifikation der Verbindungsgewichte, um im nächsten Zeitschritt besser auf das Muster reagieren zu können [Anderson 1995]. Eine Abbildung der beiden Lernkonzepte als Regelkreise von Assoziatoren mit und ohne Feedback gibt Abbildung 4.17.

¹ Auch wenn bei mikroskopischer Betrachtung unüberwachte Lernarten als biologisch am plausibelsten gelten [Zell 2000, S. 93], gleicht bei makroskopischer Betrachtung eher die assoziative Informationsverarbeitung von Mustern der eigentlichen Funktionsweise des Lernens, da die menschliche Informationsspeicherung und das Gedächtnis als inhärent assoziativ gelten und Lernen als Assoziation einer „response“ auf einen „stimulus“ verstanden wird [Anderson 1995, S. 143].

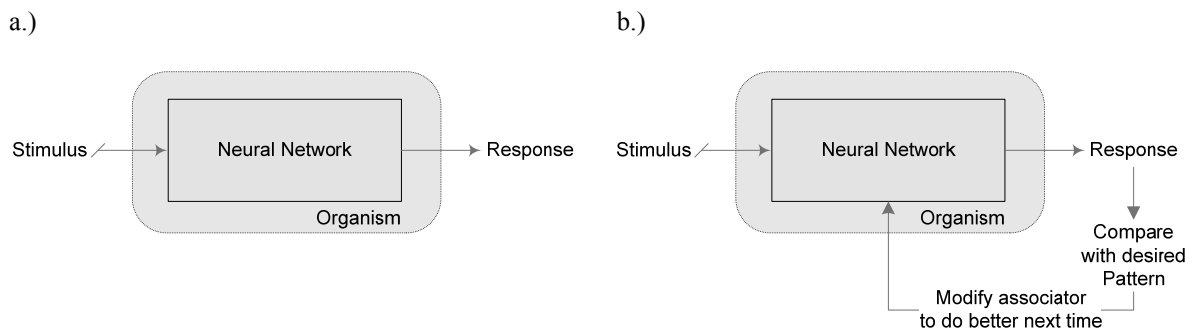
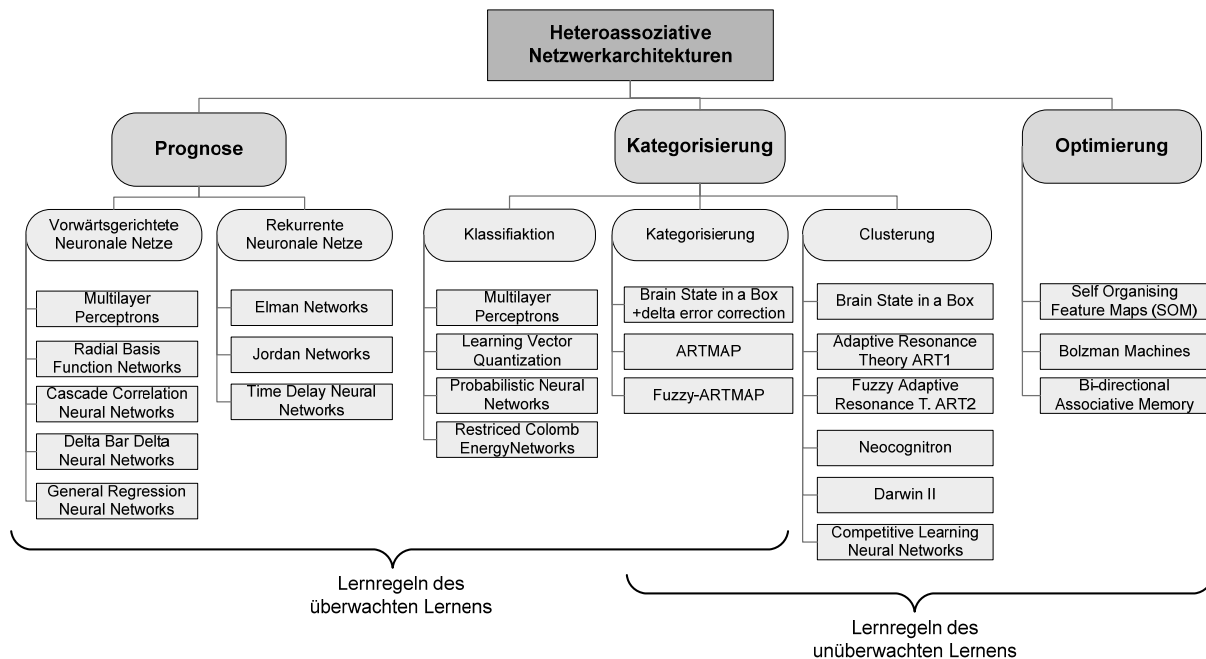


Abbildung 4.17: Lernen ohne Feedback (a.) und Lernen mit Feedback zur Fehlerkorrektur (b.)¹

Grundsätzlich können auch Ansätze des bestärkenden Lernens zum Training von NN zur Prognose eingesetzt werden. Es ist jedoch anzunehmen, dass „the more information we have about the training set and the network, the better we can learn“ [Anderson 1995, S. 210]. Während Netzarchitekturen beim bestärkenden Lernen nur im Fall eines Fehlers lernen, verbessert sich die Systemleistung beim überwachten Lernen auch dann, wenn eine richtige Ausgabe erfolgt [Anderson 1995, S. 214], da mehr Informationen zur korrekten Modifikation der Gewichte zur Verfügung stehen [Zell 2000, S. 93,95]. Entsprechend sind überwachte Lernalgorithmen für Problemstellungen der Prognose mit bekannten Eingabe- und Ausgabedaten die schnellste und effizienteste Methode der Parametrisierung [Zell 2000, S. 93]. Da im Sinne des heteroassoziativen Lernens bei der zeitreihen- oder kausalanalytischen Prognose sowohl Eingabe- als auch Ausgabevektoren bekannt sind, können überwachte Lernalgorithmen mit hoher Performanz zum Einsatz kommen. Zur Berücksichtigung architektonischer Besonderheiten, z. B. rekurrenter Verbindungen, existieren spezifische Lernalgorithmen für jede Klasse von Netzwerkarchitekturen [Zell 2000, S. 95]. Aufgrund der Beschränkung auf Feed-forward-Architekturen des MLP zur Prognose konzentrieren sich die folgenden Ausführungen auf überwachte Lernalgorithmen für MLPs. Heteroassoziatives Lernen in MLPs erfolgt durch die Anpassung der Verbindungsgewichte zwischen den Neuronen sowie ihrer Schwellwerte, um ein erwünschtes Systemverhalten zu trainieren. Dies entspricht dem Finden der freien Verfahrensparameter eines Prognosemodells, wobei die Summe der Gewichte und Schwellwerte der Anzahl freier Parameter entspricht.

In Abbildung 4.18 werden den Klassen von Anwendungen nach dem Skalenniveau und dem Zukunftsbezug der abhängigen Variablen gemäß Abschnitt 3.1.2.3 ausgewählte Netzwerkarchitekturen für heteroassoziative Lernziele zugeordnet. Diese Kategorisierung und Zuordnung wird in der Diskussion unterschiedlicher Anwendungen von NN in Abschnitt 4.4.1.2 erneut aufgegriffen.

¹ Eigene, erweiterte Darstellung in Anlehnung an ANDERSON [1995, S. 213].

Abbildung 4.18: Problemklassen der Modellierung und Architekturen Neuronaler Netze¹

Überwachte Lernverfahren erlauben Anwendungen der Prognose und Klassifikation. Dabei werden zur Prognose Architekturen der vorwärtsgerichteten MLP, RBF NN oder GRNN und einiger weitere Formen verwendet sowie rekurrente Elman- oder Jordan-Architekturen eingesetzt. Für Anwendungen der Klassifikation im Sinne einer überwachten Kategorisierung werden sowohl MLPs als auch eine Reihe spezieller Netzwerkarchitekturen der LVQ und PNN eingesetzt, was die universelle Anwendung und besondere Bedeutung der Architektur des MLP hervorhebt. Unüberwachte Lernregeln der Kategorisierung ohne Kenntnis der Klassenzugehörigkeit von Instanzen werden dagegen zum Clustering eingesetzt, wozu eine Reihe spezieller Netzwerkarchitekturen entwickelt wurde.² Da Architekturen und Lernverfahren der Kategorisierung in Form der überwachten Klassifikation bzw. der unüberwachten Clusteranalyse keinen im Hinblick auf Anwendungen der Punktprognose bieten, werden sie nachfolgend nicht weiter betrachtet.³

Aufgrund der vielseitigen Möglichkeiten der Informationsverarbeitung von NN kann die Netzwerkarchitektur eines MLP je nach dem jeweiligen Lernziel sowohl zur Realisierung homo- als auch heteroassoziativer Lernaufgaben verwendet werden. Dies erfordert eine Erweiterung der bislang architekturorientierten Notation $A=[V, T, W, K]$ um die Aspekte des Lernens L . Insbesondere sind neben dem Lernziel Z auch die nachfolgend erklärten Lernalgorithmen G und die verwendeten Lernparameter P zur Steuerung der Lernalgorithmen anzugeben, mit $L = [Z, G, P]$.⁴ Diese werden nachfolgend untersucht und spezifiziert.

¹ Eigene Darstellung in Anlehnung an LEVINE [2000, S. 200]

² Diese Kategorisierung ist vollständig korrekt, da auch unüberwachte Verfahren in Wahrheit durch interne Fehlersignale und Zielfunktionen gesteuert werden. Dennoch erscheint diese allgemein akzeptierte Bezeichnung sinnvoll [Levine 2000, S. 199].

³ Eine Übersicht über spezifische Netzwerkarchitekturen zur Kategorisierung und ihre methodischen Ursprünge findet sich in [Levine 2000, S. 198-279], eine Übersicht über Architekturen von NN zur kombinatorischen Optimierung in [Levine 2000, S. 280-364].

⁴ Anhand der strukturellen Übereinstimmung von Eingabe- und Ausgabeschicht können homoassoziative Netzwerktopologien von heteroassoziativen oft unterschieden werden. Dies muss jedoch nicht immer der Fall sein, da MLPs trotz eines heteroassoziativen Lernziels für simultane Mehrschrittprognosen eine identische Anzahl von Eingabe- und Ausgabeneuronen besitzen können. Dies erfordert eine explizite Kennzeichnung.

4.3.2 Algorithmen des überwachten Lernens

4.3.2.1 Klassische Lernalgorithmen

HEBB beschrieb 1949 als erster eine plausible neuro-biologische Lernregel, die auf der Basis der diskreten Informationsverarbeitung in binären MCCULLOCH-PITTS-Neuronen entwickelt wurde [Anderson 1995, S. 172]: „When an axion of cell u_i is near enough to excite a cell u_j and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that u_i 's efficiency, as one of the cells firing u_j , is increased” [Hebb 1949, S. 50].¹ Mathematisch wird die Veränderung des Gewichtes Δw_{ij} beschrieben durch

$$\Delta w_{ij} = \eta o_i o_j \quad , \quad (4.25)$$

mit einer konstanten Lernrate η mit $0 < \eta \leq 1$, der binären Ausgabe o_i von $\{0;1\}$, der Vorgängerzelle u_i und der binären Aktivierung o_j der Nachfolgerzelle u_j [Zell 2000, S. 85].² Entsprechend handelt es sich um eine Form des unüberwachten Lernens, bei der die Verbindungsgewichte mit der korrelierten Aktivität zweier verbundener Neuronen um den Betrag der Lernrate erhöht werden [Levine 2000, S. 68]. Die Gewichtsänderung erfolgt nur für einen Bruchteil der korrelierten Aktivierung und bildet so den schrittweisen und iterativen Charakter der Lernvorgänge ab. Bei binären Ausgaben $o_i, o_j \in \{0;1\}$ lässt die ursprüngliche Form der Lernregel nur eine Verstärkung der Gewichte zu, mit

$$\Delta w_{ij} = \begin{cases} \eta & \text{für } o_i = o_j = 1 \\ 0 & \text{für } o_i \neq o_j \end{cases} \quad . \quad (4.26)$$

Dagegen erlaubt eine einfache Erweiterung durch Skalierung der binären Ausgabewerte auf $[-1, 1]$ auch dann die Gewichtsverringerung, wenn nur eines der Neuronen aktiv ist. Die Idee der HEBB'schen Lernregel fand u. a. Anwendung im Training einlagiger Perceptrons und bildete die Grundlage nahezu aller nachfolgenden Lernregeln.

WIDROW und HOFF erweiterten 1960 die Idee der HEBB'schen Lernregel auf überwachte Lernvorgänge in NN für ADALINEs (siehe Abschnitt 4.2.5) mit einer Lage trainierbarer Verbindungen [Anderson 1995, S. 214]. Die WIDROW-HOFF-Lernregel wird auch als Delta-Lernregel (engl. *delta-rule*) oder als *Least-Mean-Squares*-Lernregel (LMS) bezeichnet. Sie berechnet eine Gewichtsänderung proportional zum Fehler der Netzwerkausgabe δ_j für ein Ausgabeneuron u_j . Dies beinhaltet eine Abkehr vom biologisch plausibleren korrelationsbasierten Lernen hin zu einer Form des fehlerbasierten Lernens unter Verwendung einer expliziten Zielfunktion und somit die Abkehr vom unüberwachten zum überwachten Lernen. Der Name der LMS-Regel hebt die besondere Bedeutung der impliziten quadratischen Zielfunktion hervor. Aufgrund der grundlegenden Bedeutung der LMS-Regel für den Lernalgorithmus der Backpropagation und nahezu alle nachfolgend entwickelten Verfahren des überwachten Lernens unter impliziter Verwendung dieser Zielfunktion mit quadratischen Fehlerkosten im Lernprozess von NN wird sie hier intensiver erläutert.³

¹ Die Originalnotation nach HEBB wurde im Zitat an die Notation der Arbeit angepasst, mit $A=u_i$ und $B=u_j$.

² Bei der üblichen Verwendung der Identitätsfunktion als Ausgabefunktion entspricht die Aktivierung a_j eines Neurons j seiner Ausgabe o_j , $o_j=a_j$. Aufgrund der vorliegenden synonymen Verwendung von Ausgabe und Aktivierung kann die ursprüngliche Formel nach HEBB entsprechend vereinfacht werden [Zell 2000, S. 85].

³ Alle Ausführungen lassen sich auch auf die auf der LMS-Regel aufbauenden, im nachfolgenden Abschnitt dargestellte Backpropagation-Lernregel anwenden. Aufgrund ihrer breiten Veröffentlichung und der erhöhten Komplexität der Herleitung der Backpropagation-Regel sowie der mangelnden Transparenz der Bedeutung der Zielfunktion in der Entwicklung der Lernregel wird hier die LMS-Regel dargestellt.

Das LMS-Lernverfahren zielt auf das Erlernen eines datengenerierenden Prozesses aus einzelnen Mustern p (engl. *pattern*) von bekannten Eingabe- und Ausgabedaten eines Datensatzes P , $p = 1, \dots, P$. Das NN hat die heteroassoziative Beziehung zwischen Eingabe- und Ausgabewerten erlernt, wenn die Verbindungsgewichte w_{ij} so bestimmt sind, dass der Fehler E als Differenz aus der tatsächlichen Netzwerkausgabe bzw. der Aktivierung o_{pj} und der erwarteten Soll-Ausgabe t_{pj} über alle Datenmuster p minimiert ist. Die Soll-Ausgabe t_{pj} wird entsprechend auch als Lerneingabe (engl. *teaching input*) bezeichnet. Ziel des Lernens ist somit die Minimierung des Ausgabefehlers E für alle Eingabe- und Ausgabemuster p in Abhängigkeit von den Netzwerkgewichten w_{ij} der Verbindungsmatrix W :

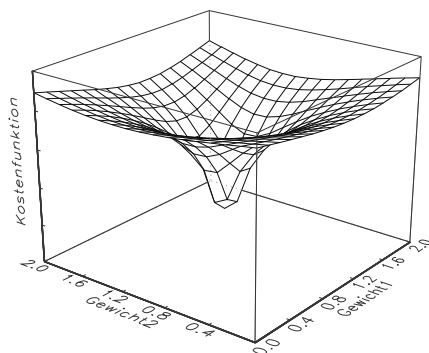
$$E(W) = \sum_{p=1}^m E_p(w_{p,1}, \dots, w_{p,n}) \rightarrow \text{MIN!} . \quad (4.27)$$

Das Problem der Minimierung des Netzwerkfehlers E für ein einzelnes Neuron j und ein einzelnes Musters p lässt sich für zwei Verbindungsgewichte w_1, w_2 und die Fehlerkosten, z. B. bewertet durch eine quadratische Kostenfunktion von

$$E_{pj} = (t_{pj} - o_{pj})^2 , \quad (4.28)$$

grafisch in Form eines Fehlergebirges in Abbildung 4.19a. veranschaulichen, das nur einem globales Minimum aufweist.

a.)



b.)

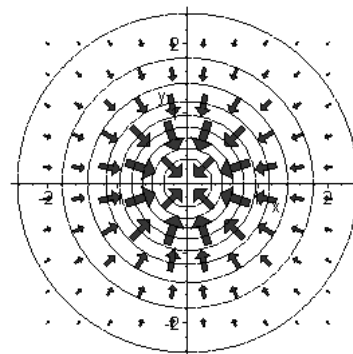


Abbildung 4.19: Netzwerkfehler als bivariate Fehlerfunktion mit einem globalen Minimum (a.) und als Gradienten der Fehlerfunktion (b.)¹

Ziel des Lernalgorithmus ist somit das möglichst schnelle und robuste Auffinden der Parameterkombination (w_1, w_2) , die den Netzwerkfehler E_j minimiert und somit das globale Minimum der Fehlerfläche findet [Zell 2000, S. 106]. Aufgrund der hohen Komplexität des Auffindens von globalen Minima im n -dimensionalen Raum über alle p Datenmuster der nichtlinearen Funktionsform kann i. d. R. keine analytische Lösung für ein globales Minimum der Fehlerfläche gefunden werden.² Deshalb wenden WIDROW und HOFF [1960] eine iterative, lokalen Suchheuristik des Gradientenverfahrens zur Suche von Minima im Lösungsraum an.³ Ausgehend von einem zufälligen Startpunkt wird lokal der Gradient der

¹ Die Gradienten werden als Pfeile dargestellt; die Höhenlinien zeigen die Steigung der Fehlerfläche.

² Minimierungsprobleme bei der Bestimmung aller freien Parameter eines NN haben eine vergleichsweise hohe Dimensionalität. Bereits eine einfache Netzwerktopologie (12-6-1) erfordert die simultane Bestimmung von 78 Parametern der Verbindungsgewichte w_{ij} und von 7 Parametern der Schwellwerte θ je Neuron und entspricht daher dem Finden eines globalen Minimums im 85-dimensionalen Raum.

³ Das Gradientenverfahren kann auch für andere unrestringierte, nichtlineare Optimierungsprobleme verwendet werden. Für eine Diskussion sei auf die Grundlagenliteratur der NN verwiesen [Bishop 1995; Haykin 1999].

Fehlerfunktion $-\nabla E_p(W)$ als Richtung des steilsten Abstiegs auf der Fehlerfläche durch partielle Ableitung ermittelt. Der resultierende Vektor spezifiziert für den berechneten Punkt die Stärke der Steigung als Betrag und zeigt in die Richtung der größten Funktionsabnahme. Abbildung 4.19b. veranschaulicht die Gradienten für unterschiedliche Stellen des Fehlergebirges in Abbildung 4.19a., die auf ein einzelnes globales Minimum zulaufen.¹ Die Richtung der Gradientenpfeile symbolisiert die Richtung des Abstiegs, die Länge der Pfeile sowie die Distanz der Höhenlinien symbolisieren den Betrag des Gradienten als Stärke der Steigung. Anschließend werden die Gewichte w_{ij} um einen Bruchteil η des negativen Gradienten der Fehlerfunktion verändert, mit

$$\Delta W = -\eta \nabla E(W), \quad (4.29)$$

wobei in Anlehnung an HEBB η ebenfalls als Lernrate bezeichnet wird und den iterativen Charakter des Algorithmus verdeutlicht. Somit wird ein neuer Berechnungspunkt erreicht, der entsprechend näher am Minimum der Fehlerfunktion liegt und so zu einer verringerten Differenz zwischen Ist-Ausgabe und Soll-Ausgabe, bewertet anhand der quadratischen Fehlerkosten, führt, im Sinne eines verbesserten bzw. erlernten Verhaltens. Der Algorithmus wird anschließend iterativ fortgesetzt und der Fehler bis zur Konvergenz in ein Minimum weiter verringert², dies entspricht einer Approximation der Lösung des Least-Mean-Squares- bzw. des Kleinste-Quadrate-Problems.

Die Beurteilung der Fehlerkosten als Differenz von Soll- und Istausgabe durch die Zielfunktion erfordert die formale Herleitung der LMS-Lernregel. Für ein einzelnes Gewicht w_{ij} und ein einzelnes Muster p gilt

$$\Delta_p w_{ij} = -\eta \frac{\partial E_p}{\partial w_{ij}}. \quad (4.30)$$

Unter Nutzung der Kettenregel folgt daraus

$$\frac{\partial E_p}{\partial w_{ij}} = \frac{\partial E_p}{\partial o_{pj}} \frac{\partial o_{pj}}{\partial w_{ij}}. \quad (4.31)$$

Das LMS-Verfahren beruht auf der Idee eines iterativen Kleinste-Quadrate-Schätzers nach WIENER [WIDROW 2003] durch eine Suche des lokalen Minimums einer Fehlerfläche mittels Gradientenabstiegs. Entsprechend bewerten WIDROW und HOFF [1960] die Fehlerabweichung durch die Summe der quadratischen Fehler (engl. *sum of squared errors*, *SSE*) als Zielfunktion zur Spezifikation von E_p :

$$E_p = \sum_j (t_{pj} - o_{pj})^2. \quad (4.32)$$

Aus der spezifizierten quadratischen Fehlerfunktion in (4.32) folgt die Ableitung des ersten Terms der Gleichung (4.31) zu

¹ Bei komplexen Zielfunktionen mit Nichtlinearitäten konvergiert das Gradientenverfahren nicht immer in ein globales Minimum. Entsprechende Erweiterungen gradientenbasierter Lernverfahren werden in Abschnitt 4.3.2.2 dargestellt.

² Das LMS-Verfahren ist ein iteratives Online-Lernverfahren auf einzelnen Datenmustern. Alternativ kann die LMS-Lernregel durch Summierung der Fehler über alle p in ein kumuliertes Trainingsverfahren erweitert werden, indem zunächst alle Muster des Datensatzes dem System präsentiert werden, bevor die Gewichte in einem weiteren Schritt modifiziert werden.

$$\begin{aligned}\frac{\partial E_p}{\partial o_{pj}} &= -(t_{pj} - o_{pj})^2, \\ &= -\delta_{pj}\end{aligned}\quad (4.33)$$

wobei der Fehler der Netzausgabe¹ für jedes Neurons j und ein Datenmusters p als δ_{pj} bezeichnet wird. Unter Verwendung von linearen Aktivierungsfunktionen im einlagigen ADALINE kann der zweite Term aus Gleichung (4.31) somit spezifiziert werden, zu

$$\frac{\partial o_{pj}}{\partial w_{ij}} = \frac{\partial \sum_i o_{pi} w_{ij}}{\partial w_{ij}} = o_{pi} \quad (4.34)$$

Aus den Umformungen (4.33) und (4.34) kann unter Anwendung von (4.32) folglich die Ursprungsgleichung (4.31) spezifiziert werden zu

$$\frac{\partial E_p}{\partial w_{ij}} = -o_{pi} \delta_{pj} \quad (4.35)$$

Die erste Ableitung der kontinuierlichen quadratischen Zielfunktion bzgl. der Gewichte ist entsprechend eine lineare Funktion der Gewichte, so dass ein Minimum der Fehlerfunktion exakt gefunden werden kann. Somit kann die allgemeingültige Modifikationsregel aus (4.30) durch Fehlerfunktion und Aktivierungsfunktionen abschließend spezifiziert werden zu

$$\begin{aligned}\Delta_p w_{ij} &= -\eta o_{pi} (t_{pj} - o_{pj}) \\ &= -\eta o_{pi} \delta_{pj}\end{aligned}\quad (4.36)$$

Die Gewichtsänderung Δw_{ij} erfolgt somit nicht länger proportional zur Aktivierung beider Neuronen, sondern proportional zum Fehler δ_{pj} der Netzausgabe, welcher über die Zielfunktion Richtung und Intensität der Gewichtsänderung definiert. Die Notwendigkeit einer exogenen Logik zur Fehlerberechnung durch die Annahme eines externen Lehrers oder Trainers, der zu jedem Zeitpunkt der Eingabe die gewünschte Ausgabe als Zielwert für jedes Neuron der Ausgabeschicht oder auch verdeckter Schichten kennt, stimmt nicht länger mit den Erkenntnissen biologischer Lernvorgänge überein. Somit repräsentiert die LMS-Regel einen entscheidenden Schritt zur Abkehr von der biologisch plausiblen Modellierung von NN hin zu einer anwendungsorientierten Entwicklung. Lernalgorithmen, die auf den Überlegungen der Delta-Lernregel aufbauen, wenden somit explizit eine quadratische Fehlerfunktion an, worauf in der Entwicklung nicht-quadratischer, asymmetrischer Zielfunktionen im folgenden Kapitel näher eingegangen wird.

4.3.2.2 Der Backpropagation-Lernalgorithmus

Die Berechnungsvorschrift der LMS-Lernregel zum Training von NN war auf eine einzelne Schicht trainierbarer Verbindungsgewichte und somit auf die Lösung linearer Problemklassen beschränkt [Anderson 1995, S. 255].² Algorithmen für ein überwachtes, hetero-assoziatives Training von NN zur Lösung nichtlinearer Modelle fehlten hingegen. „It was not until the development of the generalized error correcting rule now called backpropagation ... that it

¹ Der Fehler wird als Differenz von Ist-Ausgabe o_j und Soll-Ausgabe t_j ermittelt.

² Eine Diskussion ihrer Entwicklung und Anwendung im historischen Kontext findet sich bei [Anderson 1995, S. 239-254]

was possible to give efficient general learning rules for multilayer networks“ [Barto und Sutton, zitiert in Anderson 1995, S. 535].

Eine entsprechende Erweiterung des Algorithmus auf mehrere Schichten trainierbarer Verbindungsgewichte eines NN wurde bereits 1974 mathematisch von WERBOS formuliert [Werbos 1994] und unabhängig voneinander 1985 sowohl von LECUN als auch von PARKER modifiziert beschrieben [Levine 2000, S. 23; 57-60]. Jedoch erst die 1986 durch RUMELHART, HINTON und WILLIAMS [Rumelhart et al. 1986] formulierte Lernregel der *error backpropagation*, in Anlehnung an die Fehlerrückführung über mehrere Schichten, wurde weitläufig bekannt und verursachte eine Renaissance in der Anwendung und Erforschung von NN.¹ Die Backpropagation-Lernregel erlaubte erstmalig das Training mehrlagiger, vorwärtsgerichteter NN des Typs MLP zur Lösung nichtlinearer Problemklassen durch überwachte Lernverfahren [Zipser und Rumelhart 1990, S. 194]. Backpropagation ist eine Generalisierung der von WIDROW und HOFF entwickelten LMS-Regel auf mehrere Schichten von Neuronen mit semilinearen Aktivierungsfunktionen [Widrow und Lehr 1990]², weshalb sie auch als verallgemeinerte oder generalisierte Delta-Regel (engl. *generalised delta rule*) bezeichnet wird [Rumelhart und Zipser 1990, S. 194].

Aufbauend auf der Idee der LMS-Lernregel minimiert der Backpropagation-Algorithmus ebenfalls iterativ die quadratischen Fehler der *SSE* zwischen einer gewollten t_{pj} und der tatsächlichen Netzwerkausgabe o_{pj} für jedes präsentierte Muster p und jedes Neuron j mittels des Gradientenabstiegsverfahrens [Zell 2000, S. 85; Neuralware 1991, S. 46]. Die Berechnung der Gewichtsänderung durch die Backpropagation-Regel erfolgt analog zur LMS-Regel:

$$\begin{aligned}\Delta_p w_{ij} &= -\eta o_{pi} (t_{pj} - o_{pj}) \\ &= -\eta o_{pi} \delta_{pj}\end{aligned}\quad (4.37)$$

Dabei muss bei mehrlagigen NN zwischen der Berechnung des Fehlers δ_{pj} an Ausgabeneuronen von der Berechnung an versteckten Neuronen unterschieden werden [Zell 2000, S. 86], unter Annahme einer quadratischen Zielfunktion wie in Formel (4.28) und einer beliebigen nichtlinearen Aktivierungsfunktion $f_j(\text{net})$ des Neurons j mit

$$\delta_j = \begin{cases} f'_j(\text{net}_j)(t_j - o_j) & \forall \text{Ausgabezellen } j \\ f'_j(\text{net}_j) \sum_k (\delta_k w_{jk}) & \forall \text{verdeckte Zellen } j \end{cases} \quad (4.38)$$

Der Index k gibt für die Fehlerberechnung in Zellen der verdeckten Schichten die direkten Nachfolgezellen der Zelle j an und spezifiziert so den kumulierten Fehler aller nachgelagerten Zellen. Indem man konkrete Verarbeitungsfunktionen einsetzt, z. B. die gewichtete Summe als Eingabefunktion (vgl. Abschnitt 4.2.3.1)

¹ Nach eingehender Diskussionen um die Urheberschaft des Backpropagation-Algorithmus wurde WERBOS in 2005 auf der International Joint Conference on Neural Networks (IJCNN) offiziell als dessen Erfinder geehrt. Die grundlegende Idee zu diesem Lernverfahren der Fehler-Rückführung (engl. *back-propagation of errors*) und seine Namensgebung gehen jedoch bis auf ROSENBLATT zurück: „The procedure to be described here is called the ‚back-propagating error correction procedure‘ since it takes its cue from the error of the R -units, propagating corrections quickly at the response end“ [Rosenblatt 1958, S. 292; zitiert nach: Levine 2000, S. 23]. Die deutsche Übersetzung als „Fehlerrückführungs-Regel“ wird allgemein nicht verwendet.

² Es ist anzumerken, dass diese Aktivierungsfunktionen in der Grundform der Backpropagation monoton und vollständig differenzierbar sein müssen.

³ Dabei wird angenommen, dass die Schwellwerte als zusätzliches *Bias*-Neuron modelliert werden und so die Eingabefunktion net_j sowohl die Verbindungsgewichte w_{ij} als auch die Schwellwerte θ_j bezeichnet.

$$net_j = \sum_{i=1}^m w_{ij} o_i \quad , \quad (4.39)$$

und die logistische Aktivierungsfunktion

$$f(net_j) = \frac{1}{1 + e^{-\sum_i o_i(t) w_{ij}}} \rightarrow f'(net_j) = o_j(1 - o_j) \quad , \quad (4.40)$$

kann die Fehlerberechnung spezifiziert werden zu [Zell 2000, S. 85ff.]

$$\Delta w_{ij} = \eta o_i \delta_j \quad (4.41)$$

mit $\delta_j = \begin{cases} o_j(1 - o_j)(t_j - o_j) & \forall \text{ Ausgabezellen } j \\ o_j(1 - o_j) \sum_k (\delta_k w_{jk}) & \forall \text{ verdeckte Zellen } j \end{cases}$

Da die Fehler von der Ausgabeschicht über die verdeckten Schichten bis zur Eingabeschicht zurückgeführt werden, wird die Lernregel als Backpropagation-Algorithmus bezeichnet.

Ebenso wie die LMS-Regel ist die Backpropagation-Lernregel nicht länger biologisch motiviert, insbesondere wegen der Fehlerrückführung entlang der Verbindungen entgegen der Richtung ihres Informationsflusses und aufgrund der Veränderung der Vorzeichen von Verbindungsgewichten [Zipser und Rumelhart 1990, S. 198]. Die Anwendung des Backpropagation-Algorithmus zur Parameterbestimmung impliziert durch die Einbindung der quadratischen Zielfunktion in die mathematische Herleitung der Verfahrensvorschriften zur Fehlerberechnung grundsätzlich die Minimierung der spezifischen quadratischen Zielfunktion des SSE als Lernziel eines NN. Die genaue Verfahrensvorschrift des Backpropagation-Algorithmus wird aufgrund seiner Bedeutung in Theorie und Praxis und zur Offenlegung der implizit verwendeten Zielfunktion in Abschnitt 4.3.2.3 detailliert dargestellt und im Hinblick auf generelle Zielfunktionen und Kostenfunktionen generalisiert.

4.3.2.3 Erweiterungen des Backpropagation-Algorithmus

Das Auffinden des globalen Minimums der Fehler- bzw. Zielfunktion durch lokale Lernverfahren des Gradientenabstiegs wird durch die Erweiterung der Netzwerkarchitektur auf NN mit mehreren Schichten von Neuronen mit nichtlinearen Aktivierungsfunktionen erschwert, da sich Fehlerflächen mit multiplen lokalen Minima ausbilden. Dies wird in Abbildung 4.20 für zwei unterschiedliche Fehlerflächen, jeweils in Abhängigkeit von zwei Verbindungsgewichten w_1 und w_2 dargestellt.

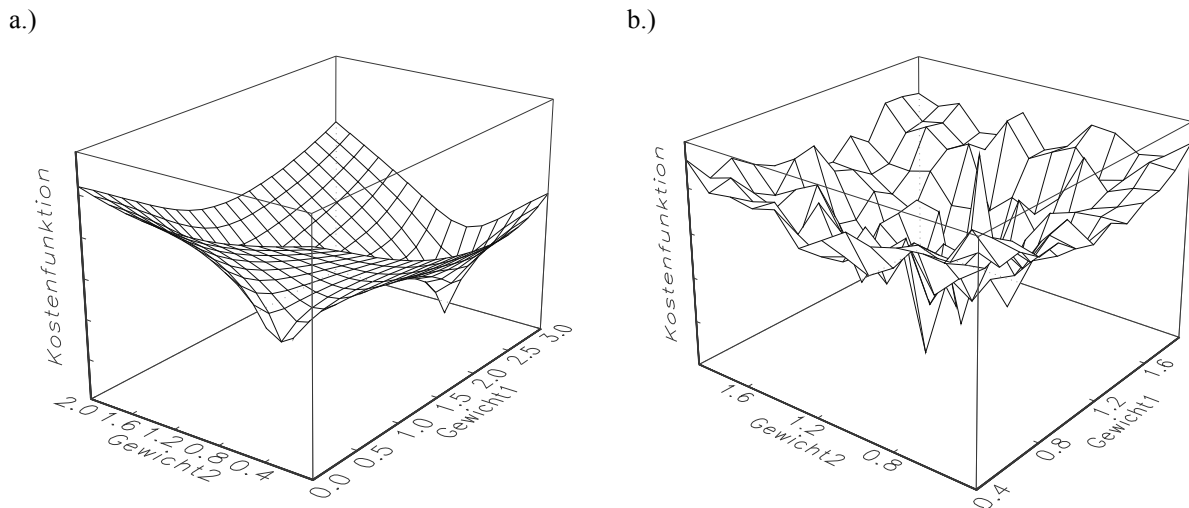


Abbildung 4.20: Bivariate Fehlerfunktionen mit je einem globalen und lokalen Minimum (a.) sowie mit multiplen lokalen Minima der Fehlerfläche (b.)

Aufgrund der lokalen Suche des Gradientenabstiegsverfahrens entsteht eine Vielzahl typischer Probleme, etwa die Konvergenz des Algorithmus in lokale Minima, die langsame Konvergenz des Trainings auf flachen Plateaus, die Oszillation in steilen Schluchten und das Verlassen vorzuziehender lokaler zugunsten schlechterer Minima [Zell 2000, S. 112ff.]. Darüber hinaus entstehen Zonen lokaler Minima mit vergleichbarer Güte, die jedoch keinen Aufschluss über die Robustheit der Lösung geben. Die Heterogenität nichtlinearer Fehlerflächen wird in Abbildung 4.21 anhand des Fehlergebirges in a.), der korrespondierenden Höhenlinien in b.) und der lokalen Gradienten der Fehlerfläche in c.) verdeutlicht.¹

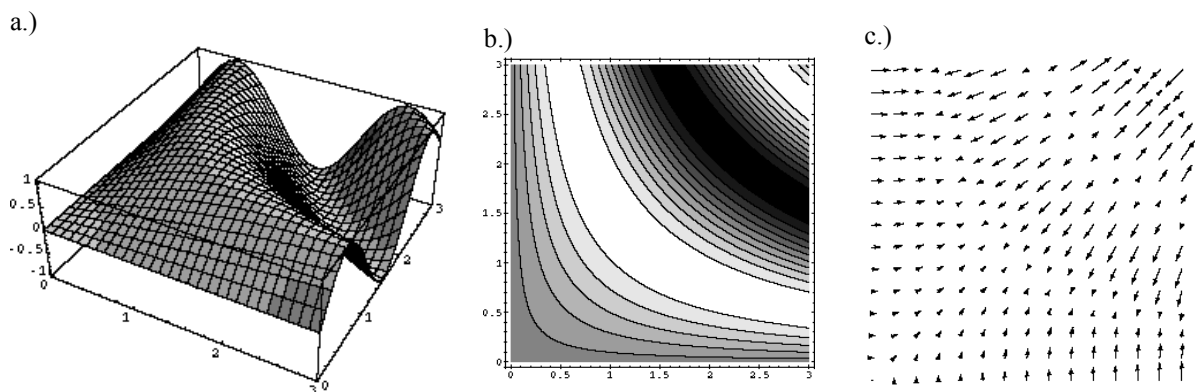


Abbildung 4.21: Bivariate Fehlerfunktion mit mehreren lokalen Minima (a.), Höhenlinien (b.) und Gradienten der Fehlerfunktion (c.).

Um diese Probleme zu umgehen, wurden seit 1986 vielfältige Erweiterungen des Backpropagation-Algorithmus sowie alternative Lernverfahren entwickelt [Silva und Almeida 1990]. Nachfolgend wird eine Auswahl verbreiteter Erweiterungen skizziert, um neben der Unabhängigkeit der Berechnungsvorschriften des Backpropagation-Algorithmus von der verwendeten Zielfunktion auch die Unabhängigkeit seiner Erweiterungen von der verwendeten Zielfunktion zu zeigen. Für eine detaillierte Diskussion der Erweiterungen sei auf [Zell 2000; Reed und Marks 1999] verwiesen.

Der konventionelle Backpropagation-Algorithmus erfordert zur Anpassung seines Lernverhaltens lediglich die Wahl der Lernrate η , welche die Schrittweite entlang des Gradienten auf der

¹ Die Richtung der Gradientenpfeile symbolisiert hier die Richtung des steilsten Anstiegs; die Länge der Pfeile symbolisiert den Betrag als Stärke der Steigung, was auch durch die verringerte Distanz der Höhenlinien symbolisiert wird.

Fehlerfläche bestimmt. Die Ex-ante-Bestimmung der Lernparameter erfolgt in Abhängigkeit von der Beschaffenheit der Fehlerfläche, die jedoch unbekannt ist.¹ Große Schrittweiten bewirken grundsätzlich einen schnelleren Abstieg entlang des Gradienten in das nächste Minimum, erlauben jedoch auch das Überspringen lokaler Minima oder das Verlassen guter Minima durch starke Bewegungen und Sprünge auf der Fehlerfläche. Bei Wahl einer geringen Lernrate kann hingegen der Lernprozess signifikant verlangsamt werden, da immer nur ein Bruchteil des Gradienten zur Gewichtsveränderung berücksichtigt wird. Die Bestimmung der Lernrate interagiert mit der Steigung der Fehlerfläche, da der Betrag des Gradienten in steilen Schluchten der Fehlerfläche stark ansteigt und so bei konstanter Lernrate die Stärke der Gewichtsanzpassung und die Schrittweite auf der Fehlerfläche erhöht. Dagegen nimmt die Ableitung der Fehlerfunktion auf flachen Plateaus der Fehlerfläche einen geringen Wert nahe 0 an, was die Schrittweite verringert und somit die Konvergenz des Algorithmus in ein Minimum verlängert [Zaghw und Dong 1994; Kamiyama et al. 1993]. Daher ist die Ex-ante-Wahl einer problemadäquaten Lernrate für die unbekannt Struktur der Fehlerfläche eines Datensatzes ein zentrales Problem der Parametrisierung von NN [Yu und Chen 1997].

Als mögliche Lösungen wurden die Erweiterung des Backpropagation-Algorithmus durch adaptive Lernraten in Abhängigkeit von der Beschaffenheit der lokalen Fehlerfläche [Rezaie et al. 2004; Yu und Liu 2002; Kamran et al. 1996; Nachtsheim 1994; Janakiraman und Honavar 1993] bzw. die schrittweise Verringerung der Lernrate von einem hohen Anfangswert für die Dauer des Lernprozesses vorgeschlagen. Dagegen postulieren u. a. NEUNEIER und ZIMMERMANN [1998] die Verwendung einer konstant hohen Schrittweite zur besseren Abdeckung des Suchraumes der Parameter.

Zur Behebung der langsamen Konvergenz auf flachen Plateaus und der Oszillation in steilen Schluchten wurde die Heuristik des sog. *momentum terms* zur Beschleunigung der Bewegung auf der Fehlerfläche eingeführt [Rumelhart et al. 1994; Rumelhart et al. 1986]. Die grundlegende Formel des Backpropagation-Algorithmus (4.37) wird dabei um den Term $\alpha \Delta_p w_{ij,t-1}$ des Momentum-Term α im Intervall $[0,2, 0,99]$ erweitert, der die vorangegangene Gewichtsänderung im Zeitpunkt $t-1$ für die aktuelle Gewichtsänderung in t berücksichtigt, mit

$$\Delta_p w_{ij,t} = \eta o_i \delta_j + \alpha \Delta_p w_{ij,t-1} . \quad (4.42)$$

Somit erfolgt eine Erhöhung der Gewichtsänderung Δw_{ij} zur Beschleunigung auf weiten Plateaus der Fehlerfläche sowie eine Verringerung der Schrittweite in steilen Schluchten [Qian 1999].² Das so erweiterte Lernverfahren wird auch als konjugierter Gradientenabstieg bezeichnet (engl. *conjugent gradient descent*). Es ermöglicht vielfach ein robusteres Lernverhalten als der Backpropagation-Algorithmus [Hagiwara und Sato 1995; Phansalkar und Sastry 1994], verursacht jedoch das zusätzliche Problem der Ex-ante-Bestimmung eines zweiten Parameters zur Kontrolle des Algorithmus.

Es existiert eine Vielzahl weiterer, potenziell sinnvoller Erweiterungen zur Optimierung von Lernraten im Backpropagation-Algorithmus, z. B. die Verringerung der Beträge von Verbindungsgewichten mit der Zeit durch *weight decay* nach WERBOS [Werbos 1988a; Mundie und Massengill 1991; Krogh und Hertz 1992; Gupta und Lam 1998]. Ferner wurden Ansätze zur

¹ Vielfach werden heuristische Regeln zur Bestimmung von Lernraten von $0 < \eta \leq 1$ postuliert, die jedoch nicht allgemein akzeptiert sind. HILL et al. schlagen eine anfängliche Schrittweite von $\eta=10$ für ihre Experimente mit MLPs zur Prognose vor, was einer Gewichtsänderung um das Zehnfache des Gradienten entspricht und somit mehr als fragwürdig ist [Hill et al. 1996].

² Es existieren ferner Erweiterungen des *momentum term*, z. B. durch einfache Exponentielle Glättung der vorhergehenden Gewichtsänderungen [Sejnowski und Rosenberg 1988], Windowing [Istook 2001], Generalisierungen [Zhao 1996] oder adaptive Momentum-Terme je Verbindungsgewicht etc. [Graepel und Schraudolph 2002; Yang et al. 2002; Drago et al. 1995; Qiu et al. 1992].

Verwendung einer spezifischen Lernrate η_{ij} für jedes Verbindungsgewicht w_{ij} entwickelt. Die Lernregel SuperSAB nach JACOBS [Jacobs 1988] erhöht die Lernraten je Verbindungsgewicht linear, wenn die partielle Ableitung über mehrere Zeitschritte das gleiche Vorzeichen besitzt bzw. verringert die Lernrate bei unterschiedlichen Vorzeichen [Zell 2000, S. 118; Tollenaere 1990]. Ähnlich verändert die Delta-Bar-Delta-Lernregel nach JACOBS und SUTTON die Verbindungsgewichte individuell durch lineares Anwachsen und exponentielles Abnehmen zur schnelleren Verringerung der Lernrate bei einem Richtungswechsel des Gradienten [Minai und Williams 1990; Sutton 1992]. Dabei erfolgt analog zur Erweiterung der Backpropagation-Lernregel eine Modifikation der Formel (4.37), nicht jedoch der Berechnungsvorschriften, die aus der Netztopologie oder der Zielfunktion abgeleitet werden.

Ein weiteres Problem in der Parametrisierung von NN entsteht durch die Aktivierung von Neuronen im Sättigungsbereich ihrer sigmoiden Aktivierungsfunktionen. Dort nimmt die Ableitung $f'_{act}(net_{pj})$ durch die geringe Steigung Werte nahe 0 an, sodass nur ein Bruchteil des Fehlersignals an vorgelagerte Neuronen propagiert und die Gewichtskorrektur verlangsamt wird. Zur Kompensation dieser sog. *flat-spots* der Ableitung schlägt FAHLMAN [1988] die Addition eines geringen, konstanten Betrags zur Ableitung einer beliebigen Aktivierungsfunktion $f'_{act}(net_j)$ in (4.38) bzw. (4.41) vor:

$$f'_{act}(net_j) + 0,1 \quad . \quad (4.43)$$

Anstelle einer derartig erweiterten Aktivierungsfunktion, auch als *modified sigmoid-prime function* bezeichnet [Zell 2000, S. 117; Balakrishnan und Honavar 1992], wird alternativ die Verwendung einer konstanten Ableitung von $f'_{act}(net_{pj}) = 0,5$ postuliert [Zell 2000, S. 116].

Neben diesen Erweiterungen des Backpropagation-Algorithmus wurden Lernverfahren zweiten Grades entwickelt, die neben der ersten Ableitung der Fehlerfläche ebenfalls höhere Momente berücksichtigen, um einen beschleunigten Abstieg auf der Fehlerfläche und somit ein schnelleres Lernen zu erreichen [Saito und Nakano 2000; Wang und Lin 1998; Osowski et al. 1996]. So wurden Lernverfahren höherer Ordnung unter Berechnung der lokalen oder inversen Hesse-Matrix entwickelt, die als Quasi-Newton- und Levenberg-Marquard-Verfahren breite Anwendung gefunden haben [Bishop 1995, S. 285ff.].¹ Verfahren höherer Ordnung zeigen im Vergleich zur Backpropagation i. d. R. eine höhere Lerngeschwindigkeit [Ramasubramanian und Karman 2004; Vrahatis et al. 2000], welche jedoch im Hinblick auf das Anwendungsgebiet vernachlässigbar sein kann² oder bei großen Netzwerktopologien durch den erhöhten Berechnungsaufwand einer großen Hesse-Matrix kompensiert wird. Zudem weisen mehrere Untersuchungen darauf hin, dass Lernverfahren höheren Grades „nicht immer die gleiche Generalisierungsleistung erzielen wie solche, die mit Backpropagation trainiert wurden“ [Zell 2000, S. 126]. Neben den etablierten Verfahren zur Parametrisierung von NN kann eine Vielzahl alternativer, lokaler Such-Heuristiken zur Parameterfindung angewendet werden, wie Tabu Search [Glover 1990; Voss 1996],

¹ Darüber hinaus existiert eine Vielzahl von Heuristiken, die aufgrund ihrer Ähnlichkeit zu bereits genannten Algorithmen hier nicht explizit besprochen werden. Beispielsweise repräsentiert der Quickprop-Algorithmus nach FAHLMAN [1988] eine Variante des sog. *line search* [Bishop 1995, S. 272], bei dem die Richtung des lokalen Gradienten verwendet wird, um in einem Schritt in ein angenommenes Minimum zu springen. Unter der Annahme eines näherungsweise quadratischen Verlaufs der lokalen Fehlerfläche kann der erwartete Scheitelpunkt der Parabel berechnet und somit die Schrittweite für eine direkte Konvergenz in das Minimum der Parabel bestimmt werden [Igel und Husken 2003].

² In Anwendungen zur Prognose liegen vergleichsweise wenige, niedrigdimensionale Datensätze vor, so dass die Lerngeschwindigkeit im Gegensatz zur Zeit der Modellbildung vernachlässigbar erscheint. Die durchschnittliche Trainingszeit auf einer Zeitreihe mit mehreren hundert Beobachtungswerten beträgt selbst auf handelsüblichen PCs unter einer Sekunde, wobei allerdings die Zeit zur Speicherung der Parametergewichte und der Fehlerverläufe eine längere Zeit in Anspruch nimmt. Dennoch ist auch bei Berechnung mehrfacher Zeitreihen kein signifikanter Zeitgewinn durch eine Beschleunigung des Lernalgorithmus zu erwarten.

natural analoge Ansätze des Threshold Accepting [Dueck und Scheuer 1990], Simulated Annealing [Kirkpatrick et al. 1983; Cerny 1985; Eglese 1990], evolutionäre und genetische Algorithmen [Fogel und Fogel 1994; Stahlbock 2002] oder Particle Swarm Optimisation etc. [Venter und Sobieszczanski-Sobieski 2003].

Im Hinblick auf vieler Anwendungen, darunter auch der Prognose, stellt sich jedoch die grundlegende Frage, inwieweit eine möglichst schnelle oder exakte Konvergenz in ein globales Minimum auf einer Stichprobe eines verrauschten Trainingsdatensatzes eine hohe Generalisierungsleistung auf ungesehenen Daten garantiert. Vielmehr ist in einem solchen Fall eine Überparametrisierung oder ein Overfitting auf den Trainingsdaten anzunehmen. Stochastische Lernalgorithmen einer iterativen Backpropagation auf einzelnen Datenmustern sind im Vergleich dagegen robust [Reed und Marks 1999], was ihre verbreitete Anwendung in Theorie und Praxis zum Teil erklärt.

Da aus der Betrachtung weiterer Lernalgorithmen kein direkter Wissenszuwachs für die Problemstellung der Bedarfsprognose und Disposition mit NN anzunehmen ist, werden Lernalgorithmen höherer Ordnung sowie Lernheuristiken in den Experimenten dieser Untersuchung nicht weiter betrachtet.¹ In Bezug auf die Darstellung unterschiedlicher Lernalgorithmen und ihrer Erweiterungen ist jedoch festzuhalten, dass sowohl die Berechnungsschritte eines Lernalgorithmus (zur Minimierung einer Zielfunktion) als auch Modifikationen der Lernalgorithmen unabhängig von der Festlegung der Zielfunktion – im Sinne der Lernaufgabe – erfolgen. Somit können die dargestellten Erweiterungen der Algorithmen auch bei einer Verwendung alternativer Zielfunktionen verwendet werden und im weiteren Zielfunktionen unabhängig von der anzuwendenden Lernregel aus der Anwendungsdomäne entwickelt und betrachtet werden.

4.3.3 Anwendung der Lernalgorithmen

Die Berechnung der Fehler des überwachten Lernens erfolgt in einem iterativen Prozess auf einer Menge von Eingabemustern und korrespondierenden Ausgabemustern. Der Prozess wird nachfolgend für iterativen Backpropagation-Algorithmus zum Training eines MLP dargestellt.

Zunächst werden alle Verbindungsgewichte w_{ij} mit zufälligen Werten in einem Intervall $[I_{min}, I_{max}]$ um 0 initialisiert.² Ein Eingabemuster p wird an die Input-Neuronen der Eingabeschicht angelegt und über die Verbindungsgewichte an die Neuronen der verdeckten Schichten vorwärts propagiert. In den versteckten Neuronen wird das Eingangssignal unter Berechnung der Eingabe-, der Aktivierungs- und der Ausgabefunktion verarbeitet und über nachfolgende Verbindungen und versteckte Schichten weitergeleitet, bis die Aktivierung die Ausgabeschicht erreicht, wo an den Ausgabeneuronen für die aktuelle Eingabe ein entsprechendes Ausgabemuster ausgegeben wird. Dies stellt die Phase der Vorwärtspropagierung (engl. *forward-pass*) der Informationsverarbeitung für ein einzelnes Datenmuster p dar.

Die berechnete Ist-Ausgabe des NN o_{pj} wird anschließend mit den gewünschten Soll-Ausgabewerten des Musters $t_{p,j}$ verglichen, im Sinne einer Lerneingabe (engl. *teaching input*) an das NN. Die berechnete Differenz zwischen Ist- und Sollausgabe an jedem Neuron j der Ausgabeschicht beschreibt somit den Fehler δ_j , den das NN in der Vorhersage des Musters p verursacht hat. Anschließend wird der Trainingsfehler eines jeden Neurons j durch partielle Ableitungen proportional zu ihrer Fehlerverursachung auf die eingehenden Verbindungsge-

¹ Grundsätzlich sind jedoch auch sie auf die Verwendung alternativer Zielfunktionen erweiterbar, so dass die Erkenntnisse aus den Experimenten in Kapitel 5 und 6 auch auf höherwertige Lernverfahren generalisierbar sind. Dies ist jedoch im Einzelfall durch Simulationsexperimente zu überprüfen.

² Dies ist erforderlich, um Probleme bei der Ausbildung identischer Gewichte mit symmetrischen Änderungen im Training zu vermeiden (engl. *symmetry-breaking*) [Zell 2000, S. 111].

wichte w_{ij} der vorangehenden Schicht umgerechnet. Die Verbindungsgewichte werden um die berechneten Anteile des Trainingsfehlers Δw_{ij} geringfügig korrigiert und so der Ausgabefehler δ_j für dieses Eingabemuster unter Nutzung eines Verfahrens des Gradientenabstiegs auf der Fehlerfläche verringert. Um die Ausgabefehler δ_j der Neuronen in versteckten Schichten zu berechnen, für die keine direkte Lerneingabe des Ausgabemusters verfügbar ist, wird der Fehler der nachfolgenden Schicht δ_j verwendet. Auf diese Weise werden die Fehler des Eingabemusters von der Ausgabeschicht über die verdeckten Schichten zur Verringerung des Fehlervektors rückwärts durch das neuronale Netzwerk bis hin zur Eingabeschicht propagiert, was die Phase der Rückwärtspropagierung beendet. Diese Berechnung in Vorwärts- und Rückwärtspropagierung wird für alle Paare von Eingabe- und dazugehörigem Ausgabemuster p im Datensatz $p = 1, 2, \dots, P$ so lange wiederholt, bis der Fehler auf den präsentierten Daten minimiert wurde [Zell 2000, S. 94]. Das Anlegen aller Datenmuster p eines Datensatzes mit P Musterpaaren wird auch als Epoche bezeichnet. Bei iterativen Lernalgorithmen erfolgt es i. d. R. zufällig für alle Muster einer Epoche, um wiederkehrende Oszillationen auf der Fehlerfläche durch die Reihenfolge der Datenmuster zu vermeiden.

Ziel des Trainings eines NN ist das Erlernen des zugrunde liegenden Erklärungsmodells aus den präsentierten Datenmustern, so dass das trainierte NN anschließend beim Anlegen von unbekanntem Eingabedaten adäquate Netzwerkausgaben im Sinne der Generalisierung zeigt. Da MLPs als Klasse universeller Approximatoren die Fähigkeit besitzen, jede Funktion zu einem beliebigen Genauigkeitsgrad zu approximieren (vgl. Abschnitt 4.4.1.1), sind sie in der Lage, sowohl den zugrunde liegenden Systemzusammenhang als auch die zufälligen Schwankungen zu lernen. Man spricht in diesem Fall von der Überparametrisierung (engl. *overfitting*) des NN – ein bekanntes Problem herkömmlicher statistischer Verfahren in der Regression und Prognose (vgl. Abschnitt 3.3.3.2). In Abbildung 4.22 ist die für eine Regression die korrekte Funktionsanpassung (b.), im Gegensatz zu einer Unteranpassung ohne Erlernen des funktionalen Zusammenhangs (a.) bzw. einer Überanpassung des funktionalen Zusammenhangs und der zufälligen Schwankungen durch Overfitting (c.) dargestellt.

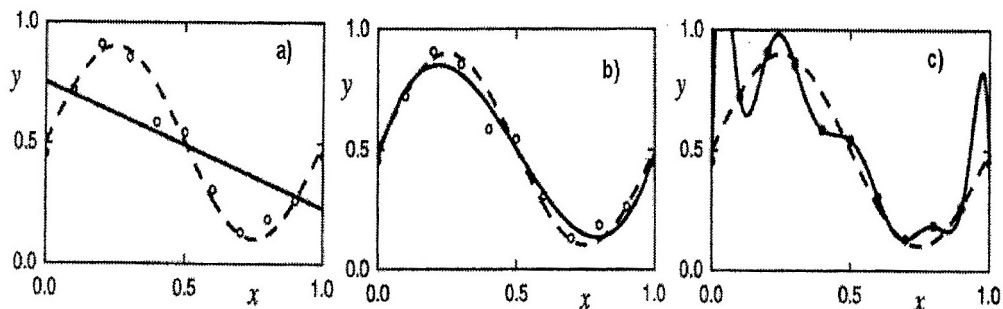


Abbildung 4.22: Underfitting (a.), Fitting (b.) und Overfitting (c.) bei der Approximation von Funktionen durch Neuronale Netze¹

Der Trainingsprozess ist daher durch Kreuzvalidierung (engl. *cross-validation*) so zu kontrollieren, dass ein NN zunächst den zugrunde liegenden Zusammenhang erlernt und den Trainingsprozess beendet, wenn das NN beginnt, unsystematische Zufallsschwankungen abzubilden. Zur Überprüfung wird die Kalibrationsstichprobe (vgl. Abschnitt 3.3.3.2) weiter in mindestens eine Trainings- und eine Validierungsstichprobe unterteilt.² Die Validierungsmenge wird dem NN zur Schätzung der Verbindungsgewichte im Training vorenthalten und

¹ Abbildung entnommen aus BISHOP [1995, S. 11]

² Durch Aufteilung in mehrfache Teilstichproben (engl. *folds*) ergibt sich auch die Möglichkeit zur Abbildung komplexerer Kontrollmechanismen der *k-fold cross validation*, auf die aufgrund ihrer problematischen Anwendung bei zeitlich geordneten Beobachtungswerten der Zeitreihenprognose nicht eingegangen wird.

zur regelmäßigen Überprüfung der Anpassungsgüte verwendet. Die Fähigkeit des NN, das zugrunde liegende Erklärungsmodell aus den Trainingsdaten zu lernen, drückt sich zunächst in einem sinkenden Fehler auf der Trainingsmenge E_{Train} und der Validierungsmenge E_{Valid} mit der wiederholten Präsentation der Datenmuster p aus. Dies wird in Abbildung 4.22 im Zeitablauf der Lernschritte Z dargestellt.

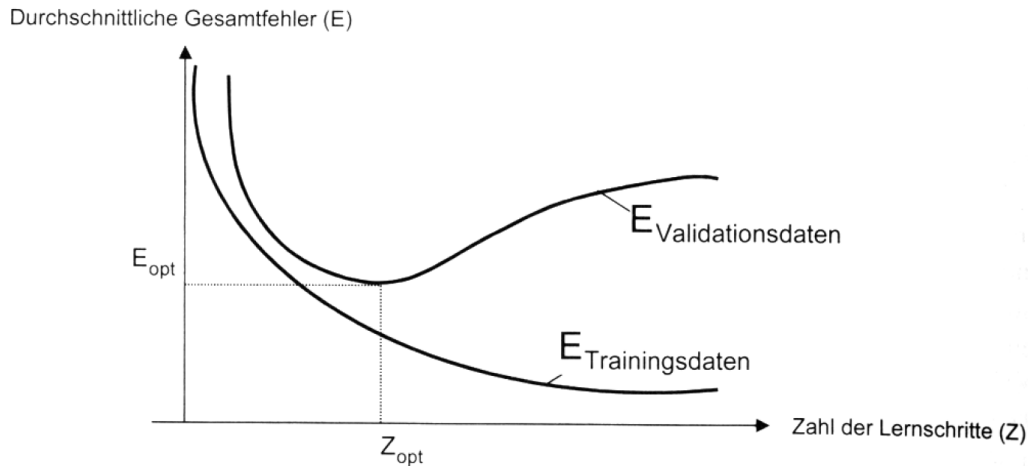


Abbildung 4.23: Verlauf der Trainings- und Validierungsfehler im Training¹

Beginnt das NN hingegen die stochastischen Zufallsschwankungen der Trainingsmenge zu lernen, was eine spätere Generalisierung verhindern würde, steigt der Fehler auf der nicht zum Training verwendeten Validierungsmenge an.² Steigt der Fehler E_{valid} über mehrere Epochen kontinuierlich an, kann der Trainingsprozess abgebrochen werden, weshalb der Prozess auch als frühes Stoppen (engl. *early stopping*) bezeichnet wird. Die Verbindungsgewichte w_{ij} im Zeitpunkt Z_{opt} bezeichnen die Parameter mit der höchsten Generalisierungsleistung durch den niedrigsten Fehler auf den Validierungsdaten E_{opt} . Die so ermittelten Verbindungsgewichte werden abschließend zur Überprüfung des NN auf der Teststichprobe verwendet, die zur Quasi-ex-ante-Evaluation zu keinem Zeitpunkt für die Parametrisierung oder Validierung verwendet wurden.³

Eine vollständige Spezifikation einer Netzwerkarchitektur $A=[V, T, W, K]$ erfordert somit eine Erweiterung um die Entscheidungen zur Parametrisierung im Lernprozess L . Zunächst ist die Anzahl der Initialisierungen I^N der Verbindungsgewichte sowie die Prozedur zur Bestimmung der Intervalle und Verteilungen der zufälligen Startgewichte I^P zu spezifizieren. Insbesondere ist der Lernalgorithmus G , z. B. eine Variante des Backpropagation Algorithmus oder ein Lernverfahren höherer Ordnung, und der vollständige Vektor aller Lernparameter P für jede Phase T des Trainingsprozesses festzulegen. Die Lernparameter können zur Berücksichtigung mehrlagiger Topologien ferner je Schicht aus versteckten Neuronen L spezifiziert werden, als $P_{T,L}$.⁴ Ferner müssen die Methode des Early Stopping zur Steuerung des Abbruchs des Lernprozesses B und die Vorgehensweise zur Auswahl eines NN-Kandidaten aus der Menge der zufälligen Initialisierungen S spezifiziert sein. Darüber hinaus ist die Zielsetzung des

¹ Abbildung entnommen aus BACKHAUS et al. [2003, S. 770]

² Bei einem deterministischen Erklärungsmodell oder bei Beobachtungen mit nur geringen stochastischen Zufallsschwankungen sinken hingegen beide Fehler konsistent, so dass zudem ein zeitliches Abbruchkriterium eingesetzt wird.

³ Die erforderliche Dreiteilung der Daten zum Training von NN wird vielfach als Kritikpunkt gegenüber NN aufgegriffen indem sie - im Gegensatz zur Zweiteilung der Datenmenge für Quasi-Ex-ante-Vergleichsstudien von statistische Verfahren - als „datenhungrig“ bezeichnet werden, obwohl andere Verfahren mit einer universellen Approximationsfähigkeit oder Modelle der dynamischen Regression neben einer Analyse der Residuen ebenso mehrere Datenmengen erfordern.

⁴ Lernparameter können je Schicht variieren, da der Anteil der Fehlerveränderung für Verbindungsgewichte nahe der Eingabeschicht immer geringer wird und dies durch höhere Lernraten kompensiert werden kann.

Lernprozesses durch die Zielfunktion oder die Fehlerfunktion anzugeben, wenngleich nur selten von der quadratischen Zielfunktion des SSE abgewichen wird. Die Lernparameter umfassen daher das Tupel $L = [[I^N, I^P], G, P_{T,L}, B, S, O]$. Somit wird ein NN vollständig durch die Netzwerkarchitektur und das Lernen beschrieben, mit $A = [[N_I - N_{H1}(F_I, F_A, F_O) - N_{H2}(F_I, F_A, F_O) - \dots - N_O(F_I, F_A, F_O)], W, K, [[I^N, I^P], G, P_{T,L}, B, S, O]]$.

4.4 Anwendung von Neuronalen Netzen zur Prognose

4.4.1 Einsatzfähigkeit Neuronaler Netze zur Prognose

4.4.1.1 Eigenschaften von Neuronalen Netzen

Jedes Prognosemodell unterstellt einen den Beobachtungen zugrunde liegenden – bekannten oder unbekanntem – funktionalen Zusammenhang zwischen den unabhängigen Variablen x_i und einer oder mehreren abhängigen Variablen y_i (vgl. Abschnitt 3.1.2ff). Das Erlernen der funktionalen Eingabe-Ausgabe-Beziehung des Erklärungs- und Prognosemodells kann somit allgemein als das Problem der Approximation einer unbekanntem Funktion $f(x)$ aus einer Menge von Datenpunkten betrachtet werden [Poggio und Girosi 1989; zitiert nach: Girosi und Poggio 1994, S. 257]. Ein NN zur Prognose als Äquivalent einer nichtparametrischen Funktion zur Approximierung einer unbekanntem Regressionsfunktion $F(W, x)$ betrachtet werden, mit der Matrix W der Verbindungsgewichte w_{ij} einschließlich der Schwellwerte θ_j als freie Parameter [White 1990, S. 544]. Unter dieser Voraussetzung wird die Fähigkeit eines NN zur Modellierung, Generalisierung und Extrapolation linearer sowie nichtlinearer Prognosemodelle unbekannter Modellform aus verrauschten Daten im Folgenden untersucht.

Die Fähigkeit von NN, die funktionalen Beziehungen des Erklärungsmodells aus den präsentierten Datenmustern durch die Bildung einer adäquaten Modellform zu erlernen, geht auf den theoretischen Nachweis ihrer Fähigkeit zur universellen Funktionsapproximation zurück. HECHT-NIELSEN [1987; 1989] implizierte in einer vergleichenden Analyse, dass nach dem Kolmogorov-Arnold-Sprecher-Theorem ein vierlagiges NN jede kontinuierliche, multivariable Funktion als finite Anzahl einfacher Funktionen beschreiben kann. FUNAHASHI [1989] erweiterte zeitgleich mit anderen Autoren [Cybenko 1989; Irie und Miyake 1988] den Beweis, wonach jede beliebige kontinuierliche Funktion durch ein MLP mit mindestens einer Schicht versteckter Neuronen mit sigmoiden Aktivierungsfunktionen und mindestens einer linearen Informationsverarbeitung in den Neuronen der Eingabe- und Ausgabeschicht realisiert werden kann.¹ HORNIK, STINCHCOMBE und WHITE [Hornik et al. 1989; Hornik et al. 1990; Stinchcombe und White 1990] bewiesen ferner, dass konventionelle MLPs mit nur einer versteckten Schicht aus beliebig vielen Neuronen mit einer beliebigen nichtlinearen Aktivierungsfunktion in der Lage sind, jede BOREL-messbare Funktion aus einem finiten Raum \mathbb{R}^n in einen anderen Raum \mathbb{R}^m abzubilden, mit $\mathbb{R}^n \rightarrow \mathbb{R}^m$, und somit den Zusammenhang zu jedem beliebigen Genauigkeitsgrad zu approximieren. Sie etablierten MLPs somit als Verfahrensklasse universeller Approximatoren, die jede kontinuierliche Funktion zu einem beliebigen, gewünschten Genauigkeitsgrad approximieren können [Girosi und Poggio 1994, S. 257].² Dies stellt eine Erweiterung der durch MCCULLOCH und PITTS bereits nachgewiesenen Eignung von NN zur Modellierung beliebiger logischer Schaltkreise dar [Funahashi 1989, S. 248]. Die theoretische Diskussion wurde seitdem auf die Approximation diskontinuierlicher Funktionen [Selmaie und Lewis 2000] sowie auf alternative Netzwerkarchitekturen mit radialen Basisfunktionen etc. erweitert [Girosi und Poggio 1994]. Eine Übersicht der theoretischen Analysen findet sich in [Gupta und Rao 1994a; Gupta und Rao 1994b]. Es gilt somit als bewiesen, dass NN und insbesondere MLPs jede beliebige funktionale Beziehung eines Erklärungsmodells approximieren können und somit für die zeitreihenanalytische sowie kausalanalytische Prognose einsetzbar sind.

¹ Die ursprüngliche Zielsetzung der Aufsätze war nicht die Prognose im Sinne der Regression, sondern die Funktionsapproximation einer Hyperebene zur Trennung unterschiedlicher Instanzen im Sinne der Klassifikation [Funahashi 1989, S. 249]. Sowohl Regression als auch Klassifikation stellen Fälle der Funktionsapproximation dar.

² Allerdings wiesen GIROSI und POGGIO auch darauf hin, dass sie theoretisch nicht die Eigenschaft der besten Approximation besitzen, wie z. B. RBF-Netze [Girosi und Poggio 1994]. Diese Überzeugung ist inzwischen jedoch sowohl theoretisch als auch praktisch angefochten worden [Sarle 1994].

Im Gegensatz zu traditionellen statistischen Verfahren stellen NN eine rein datengetriebene Methode dar, die nur wenige A-priori-Annahmen über die Modellform erfordern. So wurde unter Abschnitt 3.1.2ff. gezeigt, dass herkömmliche statistische Prognoseverfahren die explizite Formulierung der funktionalen Verknüpfung von abhängigen und unabhängigen Variablen im Sinne einer parametrischen Modellierung erfordern, gefolgt von der Findung der freien Parameter $\beta_0, \beta_1, \dots, \beta_n$ des prädeterminierten Prognosemodells. WHITE bewies dagegen die Lernfähigkeit von NN bei der Approximation beliebiger Funktionen allein aus den empirischen Beobachtungen der Datensätze [White 1990]¹, im Sinne einer nichtparametrischen Modellierung. Demnach sind NN in der Lage, sowohl die funktionale Form des Prognosemodells als auch die geeignete Parametrisierung des Prognosemodells allein aus den angelegten Eingabe- und Ausgabedaten während des Lernprozesses simultan zu erlernen. „This establishes the nonparametric regression capability of multilayer feedforward network models“ [White 1990, S. 535].

In diesem Sinne können NN als nichtparametrische statistische Verfahren betrachtet werden [White 1989a; Cheng und Titterington 1994; Amari et al. 1994], die eine valide Bildung des Erklärungsmodells ohne A-priori-Annahmen zur Modellform und zu den Verteilungsformen der Variablen ermöglichen. Demnach wären NN grundsätzlich in der Lage, beliebige komplexe funktionale Zusammenhänge aus den verrauschten Beispieldaten zu erlernen, sogar wenn die Beziehungen unbekannt oder schwer zu spezifizieren sind [Zhang et al. 1998a]. Dies reduziert die Komplexität des Modellierungsprozesses von NN erheblich, indem iterative Identifikation, Modellbildung und Parametrisierung unter Fortfall des eigenständigen Entscheidungsproblems integriert werden. Dies ist besonders in praktischen Anwendungen vorteilhaft, zumal dann, wenn die Beschaffung zusätzlicher Daten häufig einfacher bzw. effizienter ist als die Ableitung und Überprüfung von Hypothesen über die Struktur des zugrunde liegenden Systems, das die Beobachtungswerte generiert. Daher erscheinen NN für die Lösung empirischer Prognoseprobleme gut geeignet, in denen der tatsächliche Modellzusammenhang der unabhängigen Variablen unbekannt oder schwer zu spezifizieren ist, jedoch eine ausreichende Menge an Beobachtungswerten vorliegt.

Während der theoretische Beweis der universellen Approximationsfähigkeit einer beliebigen Funktion die Abbildung durch ein NN mit beliebiger Komplexität der Architektur erfordert, können in der empirischen Anwendung nur Topologien mit finiter Komplexität durch die ex ante spezifizierte Anzahl versteckter Neuronen implementiert werden. Der Nachweis der universellen Approximationsfähigkeit durch ein MLP beliebiger Komplexität liefert jedoch keinerlei Rückschlüsse auf die notwendige Anzahl versteckter Neuronen oder die Form der sigmoiden Aktivierungsfunktionen [Hornik et al. 1989]. WHITE weist entsprechend auf das Problem der korrekten Modellspezifikation in Abhängigkeit von der Anzahl verfügbarer Datenmuster hin: „Too complex a network (relative to n) learns too much – it memorises the training set and therefore generally performs poorly at generalisation tasks. Too simple a network learns too little – it fails to extract the desired relationships from the training set and thus also performs poorly at generalisation tasks: ... these constitute the dangers of both overfitting and underfitting the training set“ [White 1990, S. 535]. Das empirische Problem der Bestimmung einer optimalen Netztopologie entspricht jedoch der Festlegung der Struktur eines NN durch die Anzahl der freien Architekturparameter

$$A = [[N_I - N_{H1}(F_I, F_A, F_O) - N_{H2}(F_I, F_A, F_O) - \dots - N_O(F_I, F_A, F_O)], W, K] \quad (4.44)$$

¹ Diese Fähigkeit des Auffindens adäquater Parameter zur Funktionsapproximation aus Beobachtungen stellt eine eigenständige Fähigkeit dar, die über die theoretische Fähigkeit einer universellen Approximation hinausgeht.

Die gewählte Komplexität der Netzwerkarchitektur beeinflusst die anschließende Bestimmung der optimalen Parameterausprägungen W der Verbindungsgewichte w_{ij} , der Schwellwerte θ_j und somit die Findung der abzubildenden Modellstruktur unbekannter Komplexität. In der Anwendung muss die Netzwerkarchitektur (vielfach iterativ) so bestimmte werde, das ein Erlernen des Systems möglich ist. Die Ex-ante-Bestimmung der Netzwerk-Netzwerkarchitektur A widerspricht daher der Form eines vollständig nichtparametrischen Ansatzes. Nach DORFFNER [1996] sind MLPs somit treffender als semiparametrische Modelle bezeichnen, die eine zulässige Spezifikation der Netzwerkarchitektur voraussetzen.

Statistische Verfahren besitzen häufig restriktive Annahmen hinsichtlich der zulässigen Modellform, z. B. die Linearität der Verknüpfung der Variablen, Normalverteilung und Homoskedastizität der Residuen etc., die eine valide und reliable Approximation komplexer, realer Systeme erschweren und vielfach iterative Vorverarbeitungen erfordern (vgl. Abschnitt 3.1.4). Traditionell werden Anwendungen der Prognose von linearen statistischen Verfahren dominiert, insbesondere der univariaten Exponentiellen Glättung und von ARIMA-Modellen sowie der linearen dynamischen Regression der kausalanalytischen Modellierung, die von TONG [1990] auch als Familie der „*linear Gaussian models*“ bezeichnet wird. Diese Verfahren unterstellen grundsätzlich, dass die analysierten Zeitreihen von linearen Prozessen erzeugt werden und somit als lineare Prognosemodelle prognostiziert können. Während lineare Modelle diverse verfahrenstechnische Vorzüge aufweisen (höhere Transparenz, vereinfachte Analyse, bessere Erklärbarkeit und leichtere Anwendung), sind sie zur Anwendung in nichtlinearen Problemstellungen ungeeignet [Zhang et al. 1998a]. In empirischen Anwendungen, insbesondere solchen der Wirtschaftswissenschaften, ist es vielfach unrealistisch, a priori eine lineare Struktur des Erklärungs- und somit des Prognosemodells zu unterstellen, insbesondere deshalb, weil gerade Systeme und Prozesse der betrieblichen Realität häufig nichtlinear sind [Granger und Teräsvirta 1993].

Zur Erweiterung der linearen Modelle wurde eine Reihe nichtlinearer Prognoseverfahren entwickelt, die nichtlineare Interaktionen der Variablen z. B. durch nichtlineare Autoregressive (nAR) oder nichtlineare Moving-average-(nMA-)Modelle [Tong 1990], Bilineare Modelle [Granger und Andersen 1978], Threshold Autoregression (TAR) [Tong 1983; Teräsvirta und Lukkonen 1983; Schlittgen 1997a], Smooth Transition Autoregression (STAR) oder bei Heteroskedastizität auch durch ARCH- und generalisiert ARCH-(GARCH-)Modelle abbilden [Engle 1982]. Eine repräsentative Übersicht und Diskussion nichtlinearer Verfahren findet sich bei [De Gooijer und Kumar 1992; Tong 1990]. Allerdings erfordern diese parametrischen Ansätze der nichtlinearen Verfahren die A-priori-Spezifikation der expliziten Modellform des nichtlinearen Erklärungsmodells für eine Zeitreihe von Beobachtungen, vielfach ohne Kenntnis der zugrunde liegenden Gesetzmäßigkeiten und bei einer Vielzahl potenziell anwendbarer Modelle. NN zeigen dagegen die Fähigkeit zur flexiblen Approximation genereller funktionaler Formen [Zhang 2001; Zhang et al. 2001]. Unter Verwendung von nichtlinearen Aktivierungsfunktionen sind NN in der Lage, sowohl lineare als auch nichtlineare funktionale Zusammenhänge zu approximieren. Die Eigenschaft des datengetriebenen Ansatzes der NN erlaubt hingegen eine unmittelbare Approximation der unbekannt linearen oder nichtlinearen Modellstruktur aus den Beobachtungswerten ohne Kenntnis und A-priori-Spezifikation der funktionalen Abhängigkeiten zwischen unabhängigen und abhängigen Variablen [Zhang et al. 1998a, S. 36].

ZIMMERMANN [2006b] ordnet die Anwendungsbereiche der NN gemäß der Anzahl der unabhängigen Variablen und dem Grad der Nichtlinearität der Problemstellung im Vergleich zu etablierten mathematischen Verfahren ein, was in Abbildung 4.24 grafisch dargestellt ist:

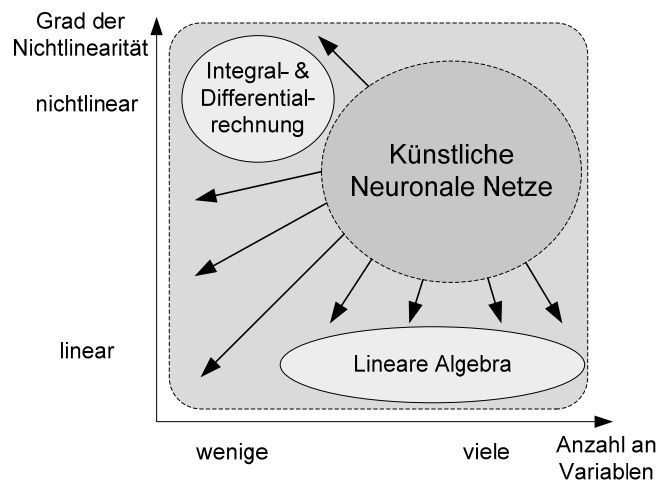


Abbildung 4.24: Anwendungsbereiche von Neuronalen Netzen¹

Nach ZIMMERMANN [2006b] sind NN zur Lösung nichtlinearer Problemstellungen mit einer hohen Anzahl von Variablen besonders geeignet. Sie erweitern somit das Instrumentarium herkömmlicher Verfahren der Linearen Algebra, die zur Lösung linearer Gleichungssysteme mit einer Vielzahl von linearen Variablen geeignet sind, und der Verfahren der Integral- und Differentialrechnung, die zur Lösung nichtlinearer Systeme mit wenigen Variablen geeignet sind. Darüber hinaus sind NN jedoch auch in der Lage, Anwendungen mit nur wenigen Variablen sowie lineare Systeme zu approximieren, so dass ihre Anwendung grundsätzlich auch auf andere Bereiche ausgeweitet werden kann. NN werden jedoch überwiegend in linearen und nichtlinearen Problemstellungen mit unbekannter Modellstruktur eingesetzt und gelten dort wegen ihrer Fähigkeit zur datengetriebenen, semi-parametrischen Modellierung als flexibler einsetzbar als parametrische, statistische Ansätze.

Die Fähigkeit zur universellen Approximation aus Daten attestiert den NN jedoch noch keine grundsätzliche Eignung zur Prognose, die ja eine Generalisierung des Gelernten auf ungesehene Daten erfordert. So teilen NN die Eigenschaft universeller Approximation z. B. mit algebraischen und trigonometrischen Polynomen [Girosi und Poggio 1994, S. 257]², die jedoch häufig Probleme in der Generalisierung aufweisen. Nach der Approximation einer spezifischen Modellstruktur und der Parametrisierung aus einer Datenstichprobe kann ein korrekt spezifiziertes NN die zugrunde liegende Struktur des datengenerierenden Prozesses im Sinne der Inferenz schlussfolgern und korrekt für ungesehene Beobachtungen extrapolieren, selbst wenn die Stichprobe der präsentierten Datenmuster erhebliche Zufallsschwankungen enthält [Zhang et al. 1998a, S. 35]. Da Problemstellungen der Prognose die Inferenz zukünftiger, ungesehener Datenpunkte aus Beobachtungen der Vergangenheit im Sinne der Generalisierbarkeit erfordern, sind NN theoretisch zur Vorhersage und Prognose geeignet.

Zusammenfassend können NN somit als semi-parametrische, universelle Approximatoren charakterisiert werden. Sie erlernen jede beliebige lineare und nichtlineare funktionale Form eines Erklärungs- und Prognosemodells ohne formelle Ex-ante-Modellbildung allein aus verrauschten Beobachtungswerten der zu prognostizierenden Zeitreihe und erlauben eine Generalisierung auf unbekannte Daten der Zukunft im Sinne der Prognose.³ Sie erscheinen

¹ Erweiterte Darstellung in Anlehnung an ZIMMERMANN [2006b, S. 4].

² Hier ist die Eigenschaft der universellen Approximation auch als Weierstrass-Theorem bekannt [Girosi und Poggio 1994, S. 257].

³ Auch wenn diverse Architekturen von NN theoretisch zur universellen, nichtparametrischen Approximation befähigt sind, konzentrieren sich die theoretischen Nachweise und unsere nachfolgenden Analysen allein auf die Architektur des MLP.

somit besonders für die Prognose komplexer, nichtlinearer oder linearer betriebswirtschaftlicher Zeitreihen geeignet [Tang et al. 1991, S. 306]. Voraussetzung für die erfolgreiche Anwendung des semi-parametrischen Verfahrens der NN ist die Spezifikation einer adäquaten Netzwerkarchitektur für eine unbekannte Modellstruktur. Der Nachweis der universellen Approximationsfähigkeit durch ein MLP beliebiger Komplexität liefert jedoch keinerlei Rückschlüsse auf die notwendige Anzahl versteckter Neuronen oder die Form der semilinearen Aktivierungsfunktionen [Hornik et al. 1989]. Allerdings impliziert die universelle Approximationsfähigkeit, dass „any lack of success in applications must arise from inadequate learning, insufficient numbers of hidden units or the lack of deterministic relationships between input and target“ [Hornik et al. 1989, S. 364]. Da NN - unter Berücksichtigung aller relevanten unabhängigen Variablen - theoretisch zur Prognose beliebiger Erklärungsmodelle befähigt sind, können ungenaue Prognosen daher nur auf Fehler in der Bestimmung der Netzwerkarchitektur zurückgeführt werden.

Die Fähigkeit von NN zu einer erfolgreichen empirischen Anwendung in der Prognose erfordert somit ihre systematische Überprüfung für eine gegebene Problemstellung und einen Datensatz. Hierzu werden zunächst empirische Anwendungen und Vergleichsstudien diskutiert und nachfolgend mögliche Richtlinien zur validen und reliablen Modellierung NN vorgeschlagen.

4.4.1.2 Anwendungsbereiche von Neuronalen Netzen

Eine repräsentative Darstellung aller Anwendungsbereiche oder gar Anwendungen von NN übersteigt den Rahmen dieser Untersuchung.¹ “Since neural networks provide a universal methodology for simulating cognitive or ‘intelligent’ functions, there is no limit to the imagination in the design of .. applications” [Levine 2000, S. 370]. Ihre umfassenden theoretischen Fähigkeiten begünstigen eine Erforschung und Erprobung von NN in einer Vielzahl möglicher Anwendungsbereiche der Geistes-, Natur-, Ingenieurs- und Sozialwissenschaften, einschließlich der Wirtschaftswissenschaften. NN werden in betrieblichen Anwendungen sowohl zur Vorhersage von Absatzzahlen, der Höhe von Aktienkursen oder des zukünftigen Bedarfs an Elektrizität als auch zur Vorhersage der Kreditwürdigkeit von Firmen oder der Wahrscheinlichkeit eines zukünftigen Konkurses eingesetzt [Vellido et al. 1999]. Darüber hinaus kann ein NN mit nahezu identischer Architektur ebenso zur Kontrolle von Prozessen in Produktionsanlagen, zur Früherkennung von Krankheiten aus Röntgenbildern, zur Erkennung gesprochener Sprache, von Gesichtern aus laufenden Videobildern oder unvollständiger Fingerabdrücke, von Sprengstoffen an Flughäfen oder sogar zur Zielerkennung in Raketensystemen verwendet werden [Widrow et al. 1994]. Die nachfolgende Diskussion bleibt jedoch zunächst auf betriebswirtschaftliche Anwendungen beschränkt, deren Umfang ebenfalls keinen Anspruch auf Vollständigkeit oder Repräsentativität zulässt.²

¹ Eine Suche in der repräsentativen Literaturliteraturdatenbank ISI Web of Knowledge (ohne Einschränkung auf anwendungsbezogene bzw. wirtschaftswissenschaftliche Publikationen) ergibt 85.131 wissenschaftliche Publikationen zu NN in Fachzeitschriften (bis einschließlich 12/2005).

² Bei der Betrachtung betriebswirtschaftlicher Anwendungen von NN ergeben sich vielfach Überschneidungen mit Modellen und Verfahren der anwendungsorientierten Forschung in den ingenieurwissenschaftlichen Disziplinen, die hier nicht dargestellt werden. Eine Übersicht ingenieurwissenschaftlicher Untersuchungen in [Ham und Kostanic 2000] zeigt eine Konzentration der Anwendungen auf die Bereiche Mustererkennung, Signalverarbeitung und adaptive Kontrolle von dynamischen Prozessen. Eine Sammlung von Veröffentlichungen zur industriellen Modellierung und Kontrolle dynamischer Systeme findet sich in [Hunt et al. 1995; Hrycej 1997; Irwin et al. 1995; Nørgaard 2000; Omidvar und Elliott 1997; Suykens et al. 1996], Ansätze zur adaptiven Kontrolle nichtlinearer, dynamischer Systeme in [Ng 1997] und eine Sammlung von Beiträgen im Kontext der Signalverarbeitung in [Hu und Hwang 2002] sowie zur Zielverfolgung in [Amoozegar 1998]. Eine Sammlung von Beiträgen zur ingenieurwissenschaftlichen Anwendung von NN in kontinuierlichen chemischen Prozessen der Prozessindustrie findet sich [Mujtaba und Hussain 2001]. Darüber hinaus existieren Hunderte von Studien zur Anwendung von NN in der industriellen Fertigung, u. a. in der Konferenzserie Artificial Neural Networks for Intelligent Engineering (siehe [Dagli 1994] bis [Dagli 2002]), den Veröffentli-

Ein Überblick über die Vielfalt betriebswirtschaftlicher Anwendungsbereiche von NN findet sich in Literaturübersichten [Li 1994; Wong et al. 1995; Sharda 1994; Widrow et al. 1994; Wong et al. 1997; Vellido et al. 1999; Krycha und Wagner 1999; Wong et al. 2000; Smith und Gupta 2000; Bakircioglu und Kocak 2000] sowie in Sammelwerken [Lisboa et al. 2000; Zhang 2004; Smith et al. 2002a; Gupta und Smith 2000; Kappen und Gielen 1997] und Monographien [Agarwal 1993; Bourbia 1989; Hou 1989].¹ Die Untersuchungen kategorisieren die publizierten Anwendungen üblicherweise nach betriebswirtschaftlichen Funktionen [Wong et al. 1997; Wong et al. 2000; Wong und Selvi 1998] oder nach Institutionen im Sinne der speziellen Betriebswirtschaftslehren [Vellido et al. 1999].² Einzelne Autoren gliedern die Anwendungen auch historisch nach ihrem zeitlichen Entstehen [Zhang et al. 1998a], nach der Art der verwendeten Datenbasis [Zhang et al. 1998a] oder je nach der verwendeten Modellform in lineare oder nichtlineare Anwendungen [Widrow et al. 1994], was jedoch wenig zur Übersichtlichkeit beiträgt. Darüber hinaus entstanden Literatur- und Anwendungsübersichten innerhalb einzelner spezieller Betriebswirtschaftslehren, insbesondere für Anwendungen in der Industrie [Fogelman-Soulié und Gallinari 1998; Jain und Vemuri 1999; Kappen et al. 1995; Meireles et al. 2003; Babel 1997; Suzuki 2000], im Transportwesen [Dougherty 1995], in der Energiewirtschaft [Kalogirou 2001; Vankayala und Rao 1993] etc. sowie innerhalb spezieller betrieblicher Funktionen, z. B. des Finanzwesens [Wong und Selvi 1998; Goonatilake und Treleaven 1995], des Rechnungswesens [Coakley und Brown 2000; Brown et al. 1995], in der Produktion [Zhang und Huang 1995; Huang und Zhang 1993] und in der Produktionsplanung [Sabuncuoglu 1998] sowie im Marketing [Rittinghaus-Meyer 1993; Venugopal und Baets 1994; Mazanec 1999; Wiedmann und Buckler 2001]. Über eine deskriptive Darstellung der unterschiedlichen Anwendungen hinaus diskutieren einzelne Anwendungsübersichten auch die relative Güte von NN im jeweiligen Anwendungsbereich und versuchen daraus Handlungsempfehlungen für eine erfolgreiche Modellierung der Netzwerkarchitektur und ihrer Parameter abzuleiten.

Aufgrund der Heterogenität der Netzwerkarchitekturen in den jeweiligen Anwendungen erlauben funktionale oder institutionale Systematiken jedoch keine Ableitung konsistenter Modellierungsempfehlungen.³ Vielmehr ist eine Modellierung von NN für spezifische Entscheidungsprobleme der Prognose, der Klassifikation oder der Optimierung erforderlich, über Funktionen, Institutionen und Disziplinen hinweg. Dies wird auch durch die Entwicklung spezifischer Netzwerkarchitekturen für einzelne Anwendungsprobleme belegt (siehe

chungen der IEEE-Konferenzserie der International Joint Conference on Neural Networks (IJCNN), der INNS International Conference on Artificial Neural Networks (ICANN) und diverser weiterer Fachkonferenzen. Viele dieser Veröffentlichungen haben jedoch nur geringe Relevanz für nicht-technische Anwendungen im betrieblichen Umfeld. Ebenso erlauben rein ingenieurwissenschaftliche Publikationen, z. B. zur optischen Zielerkennung, zur Steuerung von Kommunikationssystemen, zur Fahrzeugkontrolle etc. keine unmittelbaren Erkenntnisse für Anwendungen in der betriebswirtschaftlichen Prognose. Für sie sei auf eine Fülle repräsentativer Literaturübersichten verwiesen [Amoozegar 1998; Arul 1996; Burrascano et al. 1999; Cramer 1998; Egmont-Petersen et al. 2002; Elgindy und Palkovics 1993; El-Gindy und Palkovics 1993; Goita et al. 1994; Horne et al. 1990; Hussain 1999; Ibnkahla 2000; Kezunovic 1997; Maglaveras et al. 1998; Maier und Dandy 2000a; Miller et al. 1992; Poulton 2002; Raj et al. 1999; Robert et al. 1998; Sick 2002; Telagi 1996; Wiggins et al. 1992; Wu et al. 1991; Yasdi 2000; Zadeh et al. 1997; Zhang und Friedrich 2003; Dassen et al. 1998; del Aguila et al. 2004; Flood 2001; Gemello et al. 2000; Ghamasae et al. 1997; Greenwood 1990; Liu und Ye 2004; Lorentz und Yuh 1996; Nayak et al. 2001; Qiang et al. 2002; Rizvi und Nasrabadi 1999].

¹ Weitere systematische Literatur- und Anwendungsübersichten von NN zu spezifischen Problemstellungen finden sich in Dissertationen, so z. B. in [Alex 1998; Conrad 1996], die später im Kontext der Prognose diskutiert werden.

² Als betriebswirtschaftliche Funktionen gelten z. B. Absatz, Marketing, Controlling, Finanzen, Organisation, Personal, Produktionswirtschaft, Steuerlehre, Wirtschaftsinformatik usw., während eine institutionelle Gliederung in Handels-, Industrie-, Versicherungs- und Bankenbetriebslehre usw. unterscheidet.

³ Die Vorhersage von nominal skalierten, abhängigen Variablen zur binären Klassifikation erfordert spezifische Modellierungs-Heuristiken, die nur von geringer Relevanz für die Vorhersage von Punktprognosen zur Regression sind und auch nicht zur Bestellmengenrechnung in der Warendisposition verwendet werden können [Crone et al. 2006e].

Abschnitt 4.2.5).¹ Daher ist eine entscheidungsorientierte und anwendungsbezogene Kategorisierung nach den Entscheidungsproblemen der Prognose, der Kategorisierung, der Optimierung und der Simulation nach SHARDA [1994] vorzuziehen. Diese Kategorisierung korrespondiert mit der Einteilung der statistischen Prognosemodelle nach dem Skalenniveau der abhängigen Variablen (siehe BACKHAUS et al. [1994, S. XVII] in Abschnitt 3.1.2.3) und ist in Abbildung 4.25 dargestellt.

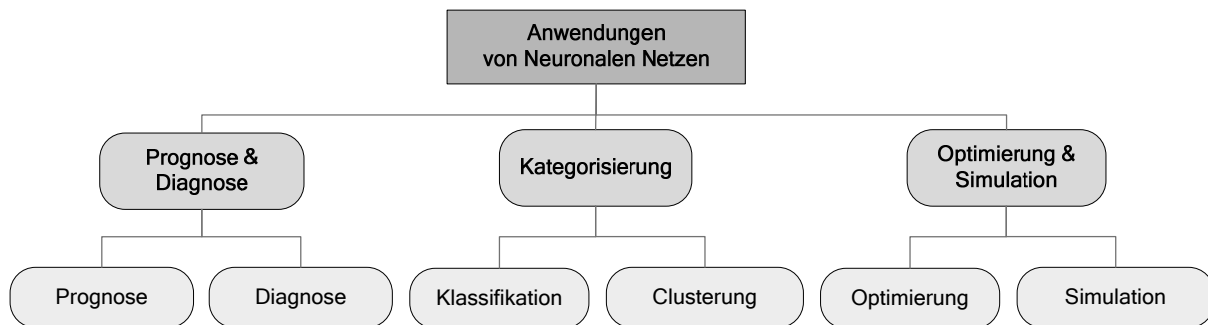


Abbildung 4.25: Anwendungen von Neuronalen Netzen zur Entscheidungsunterstützung²

Die Entscheidungsprobleme der Diagnose, Kategorisierung durch Klassifikation und Clustering sowie der Optimierung und Simulation werden zunächst exemplarisch dargestellt und in den inhaltlichen Zusammenhang dieser Arbeit eingeordnet, um im nachfolgenden Abschnitt die Anwendungen von NN zur Prognose eingehend zu diskutieren.³

Zunächst können Ansätze der Prognose und Diagnose anhand des Zukunftsbezugs der abhängigen Variablen voneinander unterschieden werden. In Abgrenzung zu Anwendungen zur Prognose und zur Klassifikation besitzen Anwendungen der Diagnose keinen Zukunftsbezug der abhängigen Variablen. NN werden zur Diagnose in der Betrugserkennung der Nutzung von Kreditkarten und Mobiltelefonen [Grosser et al. 2005; Burge und Shawe-Taylor 2001], von Versicherungspolice[n] [Viaene et al. 2005; o.V. 1998; He et al. 1997] oder des Betrugs in Jahresabschlüssen [Green und Choi 1997] eingesetzt, anhand der Überwachung laufender oder zurückliegender Transaktionen.

Ferner kann die Vorhersage zukünftiger Tatbestände sowohl als Anwendung der (Punkt-)Prognose zur Regressionsanalyse oder als Anwendung der Klassifikation zur Diskriminanzanalyse einer oder mehrere Klassen im Sinne einer überwachten Kategorisierung erfolgen.⁴ Anwendungen von NN zur Klassifikation repräsentieren einige der frühesten und noch immer populärsten Anwendungsbereiche von NN⁵, und dienen vielfach der Erklärung der methodischen Grundlagen klassischer Architekturen und Lernverfahren [Levine 2000, S. 213; Rumelhart et al. 1986; Werbos 1994]. Dabei sind insbesondere finanzwissenschaftliche Anwendungen zur Vorhersage der Kreditwürdigkeit von Individuen im Sinne des *credit*

¹ So sind NN zur Aktienkursprognose anderen Anwendungen der Prognose und Regression ähnlicher als jene, die zur Portfoliooptimierung oder zur Vorhersage von Insolvenzen innerhalb des Finanzwesens oder von Banken oder Versicherungen verwendet werden.

² Erweiterte Darstellung nach [Levine 2000, S. 200].

³ Viele der Entscheidungsprobleme der Klassifikation und Clustering können aufgrund des Zukunftsbezuges der aus ihnen abgeleiteten Entscheidungen fälschlicherweise als Ansätze der Prognose interpretiert werden. Nach der engeren Definition im Sinne einer (Punkt-)Prognose oder Regression mit einer metrischen abhängigen Variablen sind sie hier jedoch zu vernachlässigen, da sie nicht zur Absatzprognose in der Warendisposition geeignet sind.

⁴ Die Unterscheidung erfolgt anhand des metrischen oder nominalen Skalenniveaus der abhängigen Variablen. Zur Klassifikation zählen danach auch Anwendungen, die mittels metrisch skalierten, abhängiger Variablen die Klassifikation als Vorhersage von Wahrscheinlichkeiten der Klassenzugehörigkeit kodierten, z. B. im Data Mining und in der Kreditwürdigkeitsprüfung des „credit scoring“ (vgl. Abschnitt 3.1.2.3).

⁵ Anwendungen der Klassifikation gehen bis zu den Arbeiten ROSENBLATTS in der Entwicklung des ersten Perceptrons zurück [Levine 2000, S. 212].

scoring [Timothy und Sung-Chang 1999; Lee und Chen 2005; West 2000; Desai et al. 1996], zur Konkurs- bzw. Insolvenzvorhersage von Unternehmen [Atiya 2001; Charalambous et al. 2000; Zhang et al. 1999; Perez 2006; Lee et al. 2005; Pendharkar 2005] und der Insolvenz von Banken hervorzuheben [Wilkinson 1993; Tam und Kiang 1992]. Ferner werden NN zur Klassifikation von Kunden, z. B. von Käufern und Nichtkäufern im Direktmarketing und Customer Relationship Management [Rygielski et al. 2002; Cao et al. 2005b; Crone et al. 2006a; Crone und Soopramanien 2005; Crone et al. 2004], zur Klassifikation des Umsatzpotenzials von Standorten und Ladeneinrichtungen für Einzelhandelsfilialen eingesetzt [Stahlbock et al. 2005; Stahlbock und Crone 2005]. Sie dienen auch einer Vereinfachung von Punktprognosen, beispielsweise durch die Aktienkursprognose als nominalen „steigt“ oder „fällt“-Klassifikation [Ye et al. 2004; Schierholt und Dagli 1996; Rehkugler und Poddig 1991; Saad et al. 1998; Desai und Bharati 1998; Cao et al. 2005a]. Anwendungen von NN zur Klassifikation von Zeitreihen sind ebenfalls als nominale Vorhersage einzuordnen. Im Gegensatz zu anderen Forschungsdisziplinen [Schwenker et al. 2003; Husken und Stagge 2003; Tsuji et al. 2000; Kehagias und Petridis 1997; Petridis und Kehagias 1996] sind jedoch noch keine betriebswirtschaftlichen Anwendungen veröffentlicht.

Darüber hinaus werden NN zu Clusteranalyse von Daten mit einer unbekanntem abhängigen Variable im Sinne einer unbekanntem Klassenzugehörigkeit eingesetzt. Das NN bestimmt die Klassenzugehörigkeit von Objekten, so dass die Attribute der Objektinstanzen innerhalb eines Clusters möglichst homogen und zwischen den Clustern möglichst heterogen sind. Anhand der Attributausprägungen neuartiger Objekte können diese einzelnen Clustern zugeordnet, und somit Entscheidungen getroffen werden.¹ NN werden insbesondere zur Bildung homogener Marktsegmente [Balakrishnan 1992b; Hruschka 1992] und Kundensegmente eingesetzt [Balakrishnan 1992a; Laudeman]. Ansätze der Clusteranalyse können durch die Bestimmung der Klassenzuordnung auch als Vorhersagen betrachtet werden [Sfetsos und Siriopoulos 2004; Gutteridge et al. 2003; Kehagias und Petridis 2002; Bougaardt 2002; Chang et al. 2002; Potvin et al. 1993; Hruschka 1992]. So werden NN als Clusteranalyse vereinzelt auch zur Konkursvorhersage von Firmen und Banken eingesetzt [Alam et al. 2000]. Im inhaltlichen Zusammenhang der Prognose sind Ansätze zur Clusterung von Zeitreihendaten durch die Kategorisierung ähnlicher Zeitreihen in homogene Klassen hervorzuheben [Kehagias und Petridis 2002; Oh et al. 2005].

Ansätze der Optimierung nutzen NN, um die Parameter eines komplexen, meist nichtlinearen Systems zu finden, die eine gegebene Zielfunktion minimieren oder maximieren.² Anwendungen von NN zur linearen, nichtlinearen oder kombinatorischen Optimierung gehen auf eine frühe Veröffentlichungen von HOPFIELD und TANK [1985] zurück. Ihr folgte eine Vielzahl von Anwendungen zur Lösung des Problems des Handlungsreisenden durch NN zur Kombinatorischen Optimierung [Smith 1999; Smith et al. 2002b; Smith 1996], zur Wege- und Routenoptimierung und zur Produktionsplanung [Kamgar-Parsi und Kamgar-Parsi 1987; Schäfer 1991; Zuperl und Cus 2003; Shadbolt und Taylor 2002]. SHARDA gibt einen Überblick über 45 frühe Veröffentlichungen [Sharda 1994; ebenso Ashenayi 1991], gefolgt von weiteren Literaturübersichten [Ashenayi 1991]. Ansätze von NN zur Simulation zielen dagegen auf die Modellierung, Implementierung und Analyse eines dynamischen Systems in einem formalen Modell der NN zur Gewinnung von Erkenntnissen über das reale System.

¹ Die Clusteranalyse betrachtet Inputdaten als Zustandsvektoren in einem diskreten Zustandsraum möglicher Eingabemuster im Sinne der Mustererkennung, wobei alle Eingabemuster einer Region die gleiche Kategorie erhalten [Anderson 1995, S. 210]. Sie sind von Ansätzen der Clusteranalyse zur Extraktion von Wissen aus den Netzwerkverbindungen wie bei [Wilbert 1995] zu unterscheiden.

² Damit wird nicht die Optimierung einer NN-Netzwerkarchitektur, sondern die Anwendung von NN zur Optimierung bezeichnet. Ähnlich wird nicht die Simulation der Funktionsweise eines NN, sondern die Nutzung von NN zur Simulation und zur formalen Analyse dynamischer Modelle und Systeme bezeichnet.

Neben einzelnen natur- und ingenieurwissenschaftlichen Anwendungen [Mihalakakou et al. 2000; Sharma 2002; Mihalakakou et al. 2002; Klemenc und Fajdiga 2002; Al-Shayji und Liu 2002; Hussain 1999; Müller 2002] existieren nur vereinzelte wirtschaftswissenschaftliche Anwendungen, z. B. zur Produktionsablaufplanung [Laguna und Martâi 2002] oder zur Simulation von Kapitalmärkten [Zimmermann et al. 2001].

Clusteranalyse, Optimierung und Simulation werden ebenfalls zur Informationsbereitstellung für betriebliche Entscheidungsprobleme im Sinne einer Vorhersage angewendet, verwenden hierzu jedoch spezifische Netzwerkarchitekturen, z. B. selbstorganisierende Karten (SOM). Ihre Netzwerkausgaben eignen sich nicht zur Ableitung von Vorhersagen einer metrischen abhängigen Variablen im Sinne der Regression oder der Punktprognose. Somit können sie nicht zur Bestimmung zukünftiger Bedarfe in der Warendisposition angewendet werden, weshalb sie nachfolgend vernachlässigt werden. Prognose und Klassifikation werden von SHARDA [1994] auch als statistische Anwendungen von NN bezeichnet, in Abgrenzung zu Ansätzen der Optimierung und Simulation. NN sind aufgrund ihrer linearen und nichtlinearen Approximationsfähigkeit überall dort anwendbar, wo statistische Verfahren der Regressionsanalyse und der Diskriminanzanalyse bereits erfolgreich eingesetzt werden. Entsprechend kann die Tabelle 3.1 um Verfahren und Architekturen von NN erweitert werden, um die Anwendungsbereiche von NN im Vergleich zu statistischen Verfahren zu verdeutlichen und bereits diskutierte Netzwerkarchitekturen einzuordnen. Sie sind in Tabelle 3.1 aufgeführt.¹

Tabelle 4.2: Kategorisierung Neuronaler Netze nach den Skalenniveaus der Variablen²

	Skalierung der abhängigen Variablen	Skalierung der unabhängigen Variablen		
		Metrisch	Ordinal	Nominal
abhängige Variable bekannt	Metrisch	Regressionsanalyse NN: MLP, RBF, CCNN, DBDNN, GRNN, TDNN* Support Vector Regression Entscheidungsbäume	↻	Varianzanalyse
	Ordinal	↻	↻	↻
	Nominal	Diskriminanzanalyse NN: MLP, LVQ, PNN* Support Vector Machines Entscheidungsbäume	↻	Kontingenzanalyse
abhängige Variable unbekannt	Keine	Hauptkomponentenanalyse NN: MLP (mit Umformung)	↻	Clusteranalyse NN: SOM, BAM* Fuzzy Clustering

↻, ↻ = Richtung einer zulässigen Skalierung der Werte auf ein niedrigeres Skalenniveau zur Prognoserechnung

Die Verfahrensklasse der NN bieten folglich umfangreiche Anwendungsmöglichkeiten zur betrieblichen Entscheidungsunterstützung [Crone 2003a], wovon die Prognose nur einem der möglichen Anwendungsbereiche darstellt.

¹ Die Aufzählung und Einordnung der Architekturen von NN erhebt keinen Anspruch auf Vollständigkeit oder Repräsentativität. Zudem werden einzelne Verfahrensklassen aufgrund ihrer geringen Bedeutung vernachlässigt. Obwohl NN auch vereinzelt zur Hauptkomponentenanalyse eingesetzt werden [Diamantaras und Kung 1996; Jiang et al. 1996; Petranek et al. 1997; Sudjianto 1996], konnten in Übereinstimmung mit COAKLEY und Brown [2000] keine betriebswirtschaftlichen Anwendungen ermittelt werden. Ferner konnten aufgrund der deskriptiven Problemstellung der Faktoranalyse bzw. der Varianzanalyse auch in diesem Bereich keine Anwendungen von NN ermittelt werden, was aufgrund der Struktur der Netzwerkarchitekturen von NN plausibel erscheint.

² Verwendete Abkürzungen: MLP = Multilayer Perceptron; RBF = Radiale Basisfunktionsnetze; CCNN = Cascade-Correlation NN; DBDNN = Delta-Bar-Delta-NN; GRNN = Generalised Regression NN; TDNN = Time Delay NN; LVQ = Learning Vector Quantization; PNN = Probabilistic NN; SOM = Self-Organising-Map; BAM = Bidirectional Associative Memory.

4.4.1.3 Anwendungen von Neuronalen Netzen zur Prognose

Die Anwendungen von NN zur Prognose besitzen im Sachzusammenhang der vorliegenden Arbeit eine besondere Bedeutung. Daher wird auf sie im nachfolgenden Abschnitt gesondert und detailliert anhand einer repräsentativen Literaturrecherche eingegangen.

In Erweiterung vorangegangener Studien zur Anwendung von NN zur Prognose [Adya und Collopy 1998; Crone und Graffaille 2004; Crone und Preßmar 2006] wurde eine systematische Literaturrecherche durchgeführt, deren Ergebnisse in Abbildung 4.26 geordnet nach dem Jahr der Veröffentlichung dargestellt sind.¹

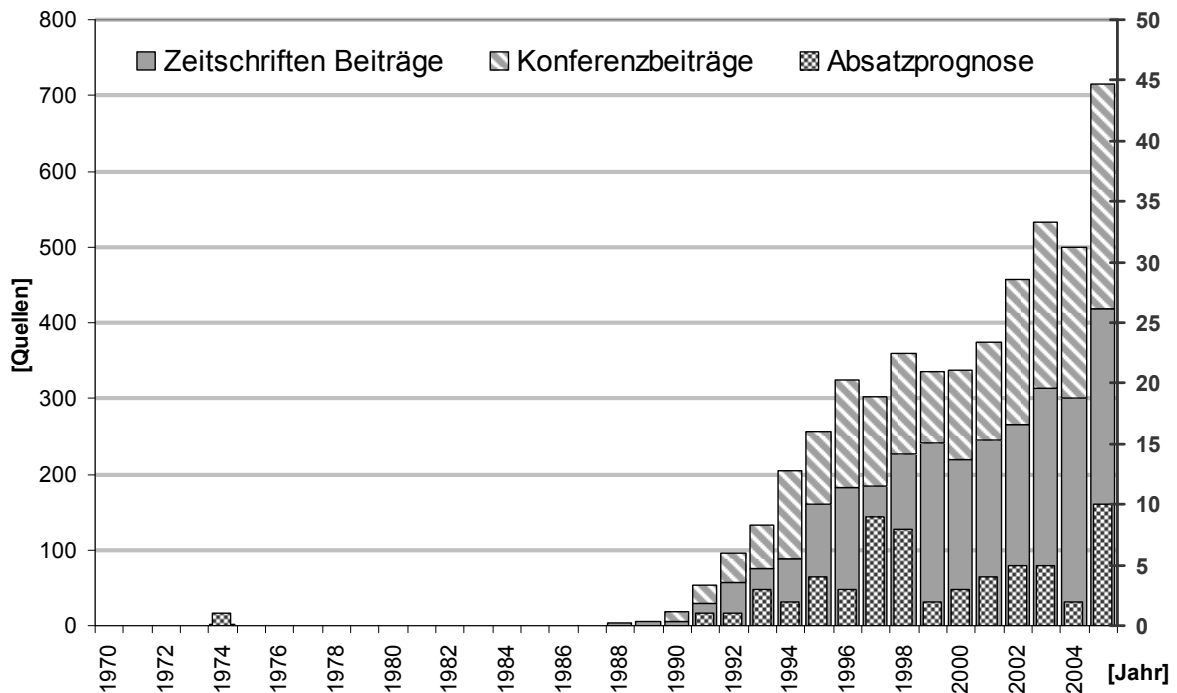


Abbildung 4.26: Wissenschaftliche Veröffentlichungen zur Prognose mit Neuronalen Netzen

Die Literaturrecherche² ergab 5009 Publikationen zur Prognose mit NN, davon 3024 Veröffentlichungen in wissenschaftlichen Zeitschriften und 1985 Konferenzbeiträge. Sie bestätigt ein ungebrochenes und wachsendes Interesse an den Verfahren der NN zur Prognose, gemessen an der kontinuierlichen Zunahme von wissenschaftlichen Publikationen in

¹ Die Literaturrecherche wurde im Dezember 2006 in der Literaturdatenbank des „ISI Web of Knowledge“ durchgeführt. Die Literaturdatenbank umfasst im Dezember 2006 22000 Zeitschriften, 23 Millionen Patente, 12000 Konferenzbände, 5500 Webseiten und 5000 Bücher mit über 40 Millionen Veröffentlichungen aus dem Science Citation Index Expanded (SCI-Expanded, von 1945 bis 12/2006), dem Social Sciences Citation Index (SSCI, von 1956 bis 12/2006) und dem Arts & Humanities Citation Index (A&HCI, von 1975 bis 10/2006). Der ISI Catalogue of Science & Technology Proceedings enthält über 5 Millionen Konferenzbeiträge (von 1990 bis 12/2006). Da nur wenige Konferenzpublikationen zur Prognose mit NN im Zeitraum zwischen der Bekanntmachung des Backpropagation-Algorithmus im Jahr 1986 und dem Beginn der Indizierung durch ISI zu erwarten sind, kann die ISI-Recherche als repräsentativ erachtet werden. Die Analyse vernachlässigt Publikationen aus dem Jahr 2006, da ISI Zeitschriften und Konferenzbeiträge des Vorjahres vielfach erst im Laufe des Folgejahrs indiziert und so keine zuverlässige Übersicht zulässt. Darüber hinaus enthält sie nur vereinzelte deutschsprachige Veröffentlichungen, zudem weder Dissertationen noch Monographien, die aufgrund ihrer geringen Anzahl nachfolgend separat diskutiert werden können.

² Der Suchbegriff lautet [title=((neural AND net*) OR perceptron*) AND (forecast* OR predict* OR timer* OR time w/2 ser* OR timeser* OR time ser*)]. Abweichende Platzhalter * wurden für zusätzliche Suchmaschinen verwendet. Die Literaturrecherche wurde im Dezember 2006 durchgeführt, um das Jahr 2005 vollständig zu erfassen. Die Recherche ist eine Erweiterung einer vorangegangenen Literaturstudie [Crone und Graffaille 2004], die in [Crone und Preßmar 2006] um Konferenzveröffentlichungen erweitert wurde. Die komplette Literaturrecherche ist online im Internet verfügbar unter: <http://www.neural-forecasting.com>.

Fachzeitschriften und Konferenzbänden seit 1989.¹ Die Literaturrecherche bestätigt darüber hinaus, dass die Problemstellung der Punktprognose als Forschungs- und Anwendungsbereich von NN im Vergleich zu Klassifikation und Clusteranalyse verhältnismäßig jung ist. So wurde die Fähigkeit von NN zur Regression theoretisch als Reaktion auf die frühe Kritik von MINSKY und PAPERT [1969] entwickelt und praktisch erst durch die Entwicklung des Backpropagation-Algorithmus durch WERBOS [1974] und RUMELHART et al. [1986] überprüft [Gupta und Rao 1994b, S. 245].²

Veröffentlichungen zur Vorhersage mit NN gehen darüber hinaus bis in die Anfänge neuronaler Modellierung zurück. HU [1964] verwendete WIDROWS einlagige ADALINE für eine multivariate Wettervorhersage in San Francisco. Diese vielfach als erste Prognose zitierte Publikation erfolgte jedoch als binäre Klassifikation zur Vorhersage ob fallen wird oder nicht [Widrow und Smith 1963; zitiert nach Neuralware 1991, S. 53-54]. Die Mehrschritt-Vorhersage auf der Basis aktueller Barometerwerte und Veränderungen des Barometers zum Vortag erfolgte somit nicht als Prognose einer metrisch skalierten, abhängigen Variablen im Sinne der Definition dieser Untersuchung. Im Sinne der vorangegangenen Definition werden nachfolgend Anwendungen zu Vorhersage durch Klassifikation, Clusterung, Optimierung und Simulation ausgeschlossen.³

WERBOS [1974; erneut veröffentlicht in Werbos 1994] demonstriert bereits 1974 das von ihm entwickelte Backpropagation-Verfahren anhand der Vorhersage volkswirtschaftlicher Zeitreihen im Vergleich zu zeitreihenanalytischen Regressions- und ARIMA-Modellen. Er demonstriert somit zeitgleich die erste Anwendung mehrlagiger NN bzw. MLPs sowie die ersten wirtschaftswissenschaftlichen Anwendungen von NN zur Prognose.⁴ Erst nach Verbreitung des Backpropagation-Algorithmus durch RUMELHART et al. [1986] folgten weitere Publikationen im Anwendungsbereich der Prognose, wie Abbildung 4.26 verdeutlicht. LAPEDES und FARBER [1987; 1988] bestätigen die Eignung von MLPs zur Prognose nichtlinearer, chaotischer Zeitreihen am Beispiel der MACKEY-GLASS-Gleichung, später ebenso Li [1990]. WHITE [1988] versuchte erfolglos die Kursentwicklung der IBM-Aktie mit NN zu prognostizieren, gefolgt von weiteren Versuchen zur Prognose von Aktienkursen [Kimoto und Kazuo 1990; Schöneburg 1990].⁵ PENG [1990] demonstriert einen ersten Ansatz zur Vorhersage der Elektrizitätsnachfrage mit NN, die heute einen der wenigen kommerziell bedeutenden Anwendungsbereiche von NN zu Prognose darstellt. Zeitgleich entstanden erste theoretische Weiterentwicklungen von NN für unterschiedliche Anwendungsbereiche der Prognose [Weigend et al. 1990; Weigend et al. 1991] sowie erste Vergleiche mit etablierten statistischen Verfahren auf einer Auswahl empirischer Zeitreihen, insbesondere mit ARIMA-Modellen [Sharda und Patil 1990; Khotanzad und Lu 1990; Tang et al. 1991; Sharda und Patil 1992a]. KANG [1991] unternahm eine umfassende Untersuchung von NN zur Prognose im Vergleich zu ARIMA-Modellen auf synthetischen und empirischen Daten.

Seit der Verbreitung des Backpropagation-Algorithmus erhalten NN zunehmende Aufmerksamkeit in unterschiedlichen Anwendungen und Disziplinen (vgl. Abschnitt 4.4.1.2),

¹ Neben der Anzahl von Publikationen kann auch die Anzahl von Zitaten der Artikel als Maßstab des Interesses verwendet werden. Insgesamt wurden die 5009 Artikel 27124 mal zitiert, mit einem zunehmenden Aufwärtstrend seit 1990.

² Dies wurde jedoch auch durch die geringen Computerressourcen, die geringe Datenverfügbarkeit und die eingeschränkte Auswahl an theoretischen Netzwerkarchitekturen unterstützt.

³ Dies verdeutlicht, dass eine Literaturrecherche durch Suchbegriffe aufgrund der uneinheitlichen Verwendung der Begriffe im Bereich Prognose und Vorhersage keine zuverlässige Trennung in Anwendungen zur Regression bzw. zur Klassifikation erlaubt. Die Analyse wurde entsprechend manuell überprüft.

⁴ Die erhebliche Innovationsleistung dieser Arbeit wird in Abbildung 4.26 anhand der nachfolgenden Zeitspanne ohne Veröffentlichungen deutlich.

⁵ Die von SHARDA [1994] zu den frühen Veröffentlichungen gezählte Publikation von SUTTON zur Prognose von Zeitreihen mit der sog. *temporal difference method* erscheint aufgrund der nur weitläufigen Verwandtschaft zu NN als wenig relevant [Sutton 1988].

einschließlich der Betriebswirtschaftslehre. Aus der Abbildung 4.26 wird deutlich, dass die Forschungsanstrengungen zur theoretischen Weiterentwicklung von NN und ihrer Anwendung in der Prognose seit 1991 erheblich ansteigen. Die umfassende Literatur zur Prognose mit NN erschwert eine systematische Literaturanalyse oder auch nur eine repräsentative Darstellung für alle potenziellen Anwendungsbereiche der Prognose. Trotz der erheblichen Forschungsanstrengungen und der Vielzahl an Veröffentlichungen entstanden nur vereinzelte Metastudien und Literaturübersichten, die Berichte über Anwendungen von NN zur Prognose für spezielle Anwendungsbereiche zusammentragen und ihre Ergebnisse hinsichtlich Prognosegüte und Modellierung evaluieren. MARQUEZ, HILL, O'CONNER und REMUS [Marquez et al. 1992; Hill et al. 1994b] geben erste Literaturübersichten über Anwendungen von NN zur kausalen und zeitreihenanalytischen Prognose. ZHANG et al. [1998a] geben eine umfassende Literaturübersicht früher Anwendungen von NN zur Prognose, analysieren die Prognosegüte von 24 Veröffentlichungen und diskutieren die verwendeten Architekturen zur Ableitung von Modellierungsempfehlungen.¹ ADYA und COLLOPY [1998] geben eine Übersicht über die relevante Literatur und untersuchen die Prognosegüte und Validität von 48 Veröffentlichungen zur Vorhersage, davon 19 zur Punktprognose. Darüber hinaus entstanden Literaturübersichten zu spezifischen betrieblichen Prognoseanwendungen, insbesondere zur Vorhersage von Wechselkursen [Wei et al. 2004], zur Kreditwürdigkeit von Firmen [Atiya 2001], zum Bedarf an betrieblichen Ressourcen der elektrischen Energie [Hobbs et al. 1998; Hippert et al. 2001] oder des Wassers [Maier und Dandy 2000b], zur Steuerung und Kontrolle industrieller Prozesse [Tai et al. 1990; Tai et al. 1991; Hunt et al. 1992; Zorriassatine und Tannock 1998; Chowdhury et al. 2001] sowie zur Erstellung volkswirtschaftlicher Prognosen [Moody 1998; Salzano 1999].² CRONE et al. geben eine erste Übersicht zur Absatzprognose in Industrie und Handel [Crone und Graffelle 2004; Crone und Preßmar 2006], die nachfolgend erweitert wird.

Da die Vielzahl an Veröffentlichungen ab 1991 eine Diskussion aller relevanten Publikationen verhindert, erscheint eine Einschränkung auf relevante Anwendungen der Punktprognose in den Wirtschaftswissenschaften als Betrachtungsgegenstand dieser Untersuchung als notwendig. Es existieren signifikante Unterschiede in der Modellierung von NN zur Prognose in unterschiedlichen Anwendungsbereichen, die eine Verallgemeinerung der Erkenntnisse über Anwendungsbereiche hinweg verbieten. Aufgrund der datengetriebenen Modellierung und der hohen Freiheitsgrade eines NN bestimmt die Länge der Zeitreihen, die Struktur der Zeitreihen hinsichtlich des Anteils der Zufallskomponente sowie des Auftretens von Ausreißern und fehlenden Beobachtungswerten, die Anzahl und Verfügbarkeit der abhängigen Variablen, die Qualität der Beobachtungswerte sowie die Zielsetzung der Anwendung nachhaltig die tatsächliche Modellierung eines NN zur Prognose.³ Zeitreihen der Absatzprognose von Produkten unterscheiden sich insbesondere von Zeitreihen naturwissenschaftlicher

¹ Auf die Aussagen zur Prognosegüte von NN wird erst im nachfolgenden Abschnitt 4.4.1.4 eingegangen.

² Veröffentlichungen zu ingenieurwissenschaftlichen Anwendungen der intelligenten Kontrolle und Steuerung industrieller Prozesse können aufgrund der Äquivalenz von Finite-Impulse-Response-(FIR-)Filtern und Infinite-Impulse-Response-(IIR-)Filtern auf Zeitreihendaten mit $AR(p)$ - bzw. $ARIMA(p,d,q)$ -Modellen ebenfalls dem Anwendungsbereich der Prognose zugerechnet werden. Sie werden entsprechend ebenfalls berücksichtigt.

³ Die Erklärungs- und Prognosemodelle von NN zur Vorhersage von Aktienkursen oder der Nachfrage nach elektrischer Energie durch dynamische Regressionsmodelle sind signifikant verschieden von jenen zur Prognose verrauschter Absatzzeitreihen in Industrie und Handel. Absatzzeitreihen von Produkten umfassen häufig nur wenige konsistente Jahre von monatlichen oder wöchentlichen Beobachtungswerten, insbesondere aufgrund von Niveauverschiebungen und Strukturbrüchen durch Produktneueinführungen, Relaunches, Ein- und Auslistungen durch Handelspartner etc. Ferner weisen sie, je nach Produktkategorie, Marktgröße, Anzahl der Marktteilnehmer und Konkurrenzsituation verhältnismäßig hohe Zufallsschwankungen und somit einen hohen Variationskoeffizienten.

Phänomene.¹ Die relative Güte von NN im Vergleich zu statistischen Verfahren besitzt somit nur vor dem Hintergrund der zu untersuchenden Anwendungsdomäne uneingeschränkte Gültigkeit. Daher beschränken sich die nachfolgende Literaturanalyse, die Beurteilung der empirischen Güte von NN und die abzuleitenden Gestaltungshinweise zu ihrer Modellierung auf relevante Publikationen von NN zur Absatzprognose in Industrie und Handel. Insbesondere erfolgt eine Ausgrenzung von Anwendungen der Natur-, Human- und weiterer Sozialwissenschaften, um die Validität der Erkenntnisse zu gewährleisten.²

Die Verfeinerung der Literaturrecherche erfordert eine schrittweise, manuelle Beurteilung und Kategorisierung aller Publikationen anhand von Titel, Abstract und Inhalt [Crone und Graffaille 2004].³ Anhand der genannten Kriterien wurden so 271 Veröffentlichungen zur zeitreihen- oder kausalanalytischen Prognose identifiziert und den betriebswirtschaftlichen Funktionen Marketing, Finanzwesen, Produktion und Absatz zugeordnet. Ihre Verteilung ist in Abbildung 4.27 dargestellt.

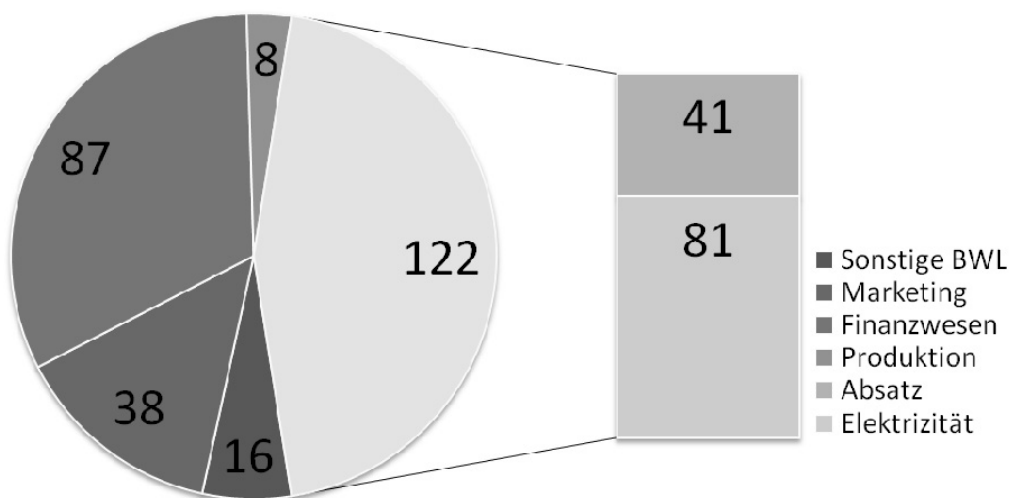


Abbildung 4.27: Betriebswirtschaftliche Veröffentlichungen zur Prognose mit Neuronalen Netzen

Ein Großteil der betriebswirtschaftlichen Publikationen entstand zu Fragestellungen der Finanzwissenschaften, insbesondere der Prognose von Aktien- und Währungskursen auf Kapitalmärkten sowie zur Prognose auf Absatzmärkten. Von den Publikationen können 122

¹ Sie bilden unterschiedliche Kategorien von Zeitreihen, die sich durch ein geringes Rauschen, das Vorliegen einer großen Anzahl von Beobachtungswerten und Muster eines nichtlinearen deterministischen Prozesses auszeichnen. Die Zeitreihen umfassen z. B. die Vorhersage von Sonnenflecken, Wasserständen in Flüssen, dem Schwingungsverhalten physikalischer Systeme oder die Prognose von synthetischen, nichtlinearen Benchmark-Zeitreihen, die kaum Gemeinsamkeiten mit der Problemstellung der Absatzprognose zur betrieblichen Warendisposition besitzen, und somit auch nur wenig Relevanz. Selbst die objektiven Ergebnisse empirischer Vergleichsverfahren wie z. B. der Santa Fe Competition [Weigend und Bhansali 1994] besitzen aufgrund der Betrachtung weniger nichtlineare Zeitreihen naturwissenschaftlicher Herkunft nur eine geringe Aussagekraft für die Anwendungsdomäne der Absatzprognose.

² Eine Vielzahl früherer Publikationen widmete sich beispielsweise der Vorhersage von Proteinstrukturen [Holley und Karplus 1989; McGregor et al. 1989; Pascarella und Bossa 1989; Qian 1999; Wilcox et al. 1989; Bohr et al. 1990; Kneller et al. 1990] sowie weiterer chemischer und physikalischer Problemstellungen der Naturwissenschaften [Elrod et al. 1989; Elrod et al. 1990; Ferrano und Wong 1990; Thibault et al. 1990], die jedoch kaum Erkenntnisse in Bezug auf die Modellierung von NN in wirtschaftswissenschaftlichen Anwendungen bieten. Weitere Publikationen und Dissertationen zur Anwendung von NN bleiben hier unberücksichtigt, z. B. ZIMMERER [1997], der NN im Vergleich zu ökonometrischen und zeitreihenanalytischen Verfahren zur Vorhersage kointegrierter Zeitreihen von US-\$/DM-Wechselkursen untersucht. Ebenso WIDMANN [2001], die die Beziehung von NN zu Verfahren der Statistik untersucht, jedoch ohne empirische Anwendung der NN. Entsprechend ist eine individuelle Beurteilung jeder Veröffentlichung zur Beurteilung ihrer Relevanz erforderlich.

³ Dabei wurden auch Veröffentlichungen einbezogen, die zumindest teilweise Anwendungen der Regression und einzelne betriebswirtschaftliche Zeitreihen untersuchen, so z. B. [Adya und Collopy 1998]. Die ursprüngliche Übersicht aus [Crone und Graffaille 2004] wurde für [Crone und Preßmar 2006] sowie für diese Analyse im Hinblick auf Veröffentlichungen aus den Jahren 2005 und 2006 entsprechend aktualisiert.

Veröffentlichungen der absatzmarktorientierten Prognose zugerechnet werden. Aus den Publikationen zur Absatzprognose wurden anschließend jene Veröffentlichungen entfernt, die nicht zur Vorhersage von Gütern im Sinne diskreter Absatzeinheiten dienen. Insbesondere wurde die Prognose von Werbeerfolgswerten, Marktanteilen, Preisindizes und Preisen den Funktionen des Marketing zugeordnet.¹ Aus den verbleibenden 41 Veröffentlichungen zur Bedarfs- und Absatzprognose von Waren und Gütern mittels NN wurden zusätzliche Literaturquellen und relevante Autoren anhand der Literaturverzeichnisse identifiziert und die Recherche durch erweiterte Literaturrecherchen nach Autoren erweitert.² Es wurden auch Veröffentlichungen zu theoretischen Weiterentwicklungen von NN berücksichtigt, die zur Überprüfung der Methoden zumindest eine Absatzzeitreihe verwenden. Die Literaturrecherche identifiziert abschließend 63 Publikationen in der Anwendungsdomäne der Absatzprognose von Gütern, über die nachfolgend ein Überblick gegeben wird.³

Die Mehrzahl früher Veröffentlichungen untersucht die grundsätzliche Eignung von NN zur Absatzprognose auf wenigen Referenz-Zeitreihen. TANG et al. [1991] beispielsweise untersuchen die Prognosegüte von NN auf drei populären Benchmark-Zeitreihen, einschließlich der bekannten Airline-Passenger-Zeitreihe⁴, und erweitern diese später um den Einfluss unterschiedlicher Architekturparameter [1993a; später erschienen als Tang und Fishwick 1993b]. KOLARIK und RUDORFER [1994] untersuchen die Auswirkungen von ausgewählten Lern- und Architekturparametern auf die Prognosegenauigkeit im Vergleich zu ARIMA-Modellen, ebenfalls anhand der Airline-Passenger-Zeitreihe sowie der IBM-Aktienkurse in Anlehnung an WHITE [1988]. FARAWAY und CHATFIELD [1995; 1998] untersuchen die Fähigkeiten von NN bei der Absatzprognose und das Problem des Overfitting als Reaktion auf die Veröffentlichung von TANG et al., ebenfalls unter Verwendung der Airline-Passenger-Zeitreihe.⁵ NAM und SCHAEFER [1995; ebenso erschienen als Nam et al. 1997] prognostizieren ebenfalls den Bedarf an internationalen Flügen im Vergleich zu ARIMA-Verfahren und Exponentieller Glättung, jedoch nicht anhand der Zeitreihe von BROWN. ALEKSEEV und SEIXAS [2002] untersuchen den Einfluss von Architekturparametern der NN auf die Prognosegüte des Flugpassagieraufkommens in Brasilien.

Eine Vielzahl weiterer Veröffentlichungen beschränkt sich auf die Prognose einzelner Zeitreihen aus unterschiedlichen Absatzbereichen. JAGLIESKA und JACOB [1993] entwickeln ein NN zur kausalanalytischen Absatzprognose für Lotterielose in Australien unter Berücksichtigung von Werbeaktionen. KONG und MARTIN [1995] untersuchen die Prognosegüte von NN auf monatlichen Absatzzeitreihen eines Lebensmittelgroßhändlers. LUXHOJ et al. [1996] analysieren die Eignung von NN und eines erweiterten hybriden Verfahrens zur Vorhersage der monatlichen Abverkäufe einer Zeitreihe eines Elektronikartikel-Herstellers. YIP et al.

¹ Ferner wurde die Vorhersage von elektrischen Lasten zur Ermittlung des kontinuierlichen Bedarfs der Energieversorger ausgegrenzt. Die Vorhersage elektrischer Lasten zeichnet sich durch besondere Strukturen von Zeitreihen, oft in kurzen Zeitintervallen von Stunden, mit mehrfacher Saisonalität, erklärenden Variablen und Marktinteraktionen aus, die eine Vergleichbarkeit zur Absatzprognose diskreter Waren erschweren.

² Die erweiterte Literaturrecherche umfasste die Datenbanken des ISI, IEEE Xplore, ScienceDirect, EBSCO, Citeseer, Proquest, OCLC und ABI/Inform.

³ Entsprechend der Literaturrecherche werden nur Publikationen bis einschließlich 2005 berücksichtigt. Einige technische Reports und unveröffentlichte Konferenzbeiträge waren nicht vollständig zu beziehen [Patterson et al. 1997; Muller 1997; Gorzalczyk 1998; Orta und Serpen 2005; Zekic-Susac et al. 2005; Ou und Chen 2006; Jiang 2003; Sztandera et al. 2003]. Da dies auf eine eher geringe Relevanz der jeweiligen Veröffentlichung schließen lässt, werden diese Publikationen aus der nachfolgenden Betrachtung ausgeschlossen. Die Veröffentlichung von [Dong et al. 2005] ist nicht verfügbar, wurde jedoch 2006 als [Dong und Wen 2006] in einer Fachzeitschrift veröffentlicht und ist aufgrund der zeitlichen Einschränkung der Recherche zu vernachlässigen.

⁴ Die Airline-Passenger-Data-Zeitreihe wurde bereits von BROWN zur Überprüfung der Exponentiellen Glättung sowie von BOX und JENKINS zur Überprüfung der Methodologie saisonaler ARIMA-Verfahren verwendet [Brown 1956; Box und Jenkins 1970].

⁵ CHATFIELD [Chatfield 1998] fasst die Erkenntnisse als grundsätzliche Kritik an NN zusammen, die jedoch nicht spezifisch auf Veröffentlichungen der Absatzprognose ausgerichtet ist.

[1997] vergleichen die Absatzprognose eines Produkts mit unterschiedlichen gleitenden Mittelwerten und der einfachen Exponentiellen Glättung. KUO und XUE [1997; ebenso nahezu unverändert in: Kuo 2001; 2002; 1998a; 1998b; 1999] untersuchen die Erweiterung von NN zur dynamischen Regression unter Kodierung von Werbeaktivitäten als Fuzzy-Eingabeinformationen auf einer Zeitreihe von Papaya-Milch beim US-amerikanischen Pharmazie-Einzelhändler CVS. TCHABAN et al. [1997; 1998] stellen die Prognosefähigkeit eines NN anhand einer Zeitreihe des Handels dar. ONG und FLITMANN [1998] vergleichen die Prognosegenauigkeit von NN mit Exponentieller Glättung und der Regression auf einer unspezifizierten Zeitreihe. BABA und SUTO [2000] untersuchen die Eignung von Genetischen Algorithmen für die Bestimmung der Inputvariablen für eine Zeitreihe wöchentlicher Absatzmengen von technischen Einzelteilen. CHANG et al. [2005] entwickeln ein mit Evolutionären Algorithmen konfiguriertes NN zur Bedarfsvorhersage von Halbleiterplatinen anhand einer Zeitreihe aus Taiwan.

Die Beurteilung von NN auf mehreren Zeitreihen erfolgte überwiegend unter Verwendung der Daten der populären M-Competition [Makridakis et al. 1982] mit monatlichen, vierteljährlichen oder jährlichen Beobachtungsintervallen oder auf synthetischen Zeitreihendaten. SHARDA und PATIL [1990; 1992b] untersuchen 75 Zeitreihen der 111 Zeitreihen der M-Competition im Vergleich zu ARIMA-Verfahren. TANG und FISHWICK [1993b; 1993a] vergleichen die Prognosegüte von NN mit der von ARIMA-Modellen auf 14 monatlichen und vierteljährlichen Zeitreihen der M-Competition. FOSTER et al. beurteilen die Prognosefähigkeit von NN auf 384 ökonomischen Zeitreihen der M-Competition im Vergleich zu Regression und Exponentieller Glättung [1992]. HILL, O'CONNOR und REMUS [1996] untersuchen die Prognosegüte von NN auf allen 111 Zeitreihen der M-Competition. Darauf aufbauend untersuchen die Autoren die Notwendigkeit der Saisonbereinigung für die Anwendung von NN auf einer Auswahl saisonaler, monatlicher Zeitreihen [Nelson et al. 1994; Nelson et al. 1999], ZHAO et al. [2003] untersuchen die positiven Ergebnisse von [1994a], scheitern jedoch bei dem Versuch einer Replikation, ebenso wie DHAWAN und O'Connor [2005]. BALKIN und ORD [2000] nehmen als einzige mit einer systematischen Methodologie zur automatischen Modellierung von NN an der Prognose der 3003 Zeitreihen der M3-Competition teil, jedoch nur mit geringem Erfolg. LIAO und FILDES [2005] analysieren die Prognosegüte von NN im Vergleich zu etablierten statistischen Verfahren auf 261 nicht-saisonalen Zeitreihen der Telekommunikations-Industrie. ZHANG et al. [2001] untersuchen die theoretischen Fähigkeiten von NN zur Prognose anhand von synthetischen Zeitreihen aus 240 linearen Prozessen und 240 nichtlinearen Prozessen [Zhang et al. 1998c; Zhang et al. 2001]. CRONE et al. [Crone et al. 2006a; Crone et al. 2006b; Crone et al. 2006c; Crone et al. 2006d] untersuchen die Fähigkeit und Sensitivität unterschiedlicher Datenvorverarbeitungen von NN in der Prognose typischer Zeitreihenmuster der Absatzprognose im Vergleich zu statistischen Prognoseverfahren auf 12 und 36 synthetischen, die aus eine monatlichen Zeitreihe des Handelsabsatzes dekomponiert wurde.

Im Gegensatz zur Prognose von monatlichen Zeitreihen auf aggregiertem Niveau sind nur wenige Publikationen der Absatzprognose wöchentlicher oder täglicher Daten gewidmet, wie sie für den Absatz im Einzelhandel typisch sind.¹ POH und JASIC [1995] untersuchen die Anwendung zweier Netzwerkarchitekturen zur kausalanalytischen Prognose des Absatzes in Warenhäusern auf täglichen und wöchentlichen Zeitreihen unter Verwendung von Marketingaktionen in unterschiedlichen Produktkategorien und mit variierenden Werbebudgets. ANSUI et al. [1996] modellieren NN als dynamische Regressionsmodelle und vergleichen die

¹ Die Meta-Studien zur Qualität einzelner Publikationen zur Anwendung von NN in der Absatzprognose [Crone und Graffelle 2004; Crone und Preßmar 2006] werden entsprechend nicht zu den originären Publikationen gezählt, auch wenn sie dasselbe Thema behandeln.

Prognosegüte mit ARIMA-Modellen mit Interventionsvariablen auf saisonalen Absatzzeitreihen. ZABEL [2000] untersucht die Fähigkeit von NN u. a. an 28 Artikeln mit wöchentlichen Abverkaufszahlen. BANSAL et al. [1998] evaluieren NN auf 20 wöchentlichen Zeitreihen eines US-amerikanischen Pharma-Einzelhändlers und betrachten die Auswirkungen auf die Warenbestände hinsichtlich Fehlmengen und Überbeständen direkt auf der prognostizierten Menge, ohne Berechnung von Sicherheitsbeständen. Die Untersuchung wird in [Shanmugasundaram et al. 2002] auf tägliche Absatzdaten und die Erprobung unterschiedlicher Rekurrenter Netzwerkarchitekturen für 10 Produkte erweitert, jedoch nur mit geringem Erfolg. JIANG et al. [2000] untersuchen die Prognosegüte auf täglichen Scanner-Daten von 14 Produkten aus 3 Produktkategorien im Vergleich zu statistischen Verfahren. THIESING und VORNBERGER [1995a; 1995b; 1995c; 1995d; 1995; 1997a; 1997b; 1997c; 1998] publizieren eine Reihe weitgehend identischer Beiträge aus einem Projekt zur Absatzprognose an Supermärkten der Firma Tegut. Während diese Konferenzbeiträge lediglich die Prognosegüte für eine einzelne Zeitreihe wiedergeben, zeigt die Dissertationsschrift THIESINGS [1998] die Analyse darüber hinausgehender Zeitreihen, jedoch ohne Bezug zu Fragen des Bestandsmanagements oder der Bestellmengenrechnung. CONRAD [1996] untersucht die Einsatzmöglichkeiten von NN zur Absatzprognose im Vertrieb von Publikumszeitschriften im Vergleich zu ARIMA-Verfahren.

CRONE [2002] überprüft die Fähigkeit von NN zur Absatzprognose und Warendisposition anhand einer Abverkaufszeitreihe an Warenautomaten. Diese dient später zur Entwicklung und Überprüfung einer Methodologie zur Modellselektion von NN in der Absatzprognose [Crone 2004; 2005]. THOMASSEY et al. [2002; 2005] entwickeln ein Prognosesystem zur kurz-, mittel und langfristigen Absatzprognose für einen französischen Textileinzelhändler und überprüfen die Güte von NN anhand von Zeitreihen für 42.000 Produkte in 322 Produktkategorien im Vergleich zu ARIMAX-Modellen. RAGG et al. [2002] entwickeln ein NN mit Regularisierung zur Kontrolle des Overfittings in der Prognose von Remissionsquoten der Bild-Zeitung an 112.000 Abverkaufsstellen. YOHDA et al. [2002] prognostizieren den Absatz von zwei neuen Produkten ohne Absatzhistorie anhand von 36 kausalen Erklärungsvariablen mit NN unter Kodierung der Eingangsdaten durch Fourier-Transformationen. KOTSIALOS et al. [2005] untersuchen die Prognosegüte von NN zur mittel- bis langfristigen Absatzprognose anhand von 74 wöchentlichen Zeitreihen zweier Produktionsunternehmen

Darüber hinaus existiert eine Reihe von Untersuchungen im Anwendungsgebiet des Handels, die jedoch einen monatlichen Absatz aus makroökonomischer Sicht betrachten. Die Publikationen von ALON et al. [2001] untersuchen die Prognose des aggregierten, monatlichen Absatzes im Einzelhandel der USA. CHU und ZHANG [2003] untersuchen unterschiedliche Ansätze der Datenvorverarbeitung durch Trend- und Saisonbereinigung ebenfalls anhand des monatlichen Einzelhandelsabsatzes der USA. Als Erweiterung untersuchen ZHANG und QI [2003; 2005] weiter die Fähigkeiten von NN in der Absatzprognose von saisonalen und trendbehafteten Zeitreihen anhand makroökonomischer Zeitreihen des Handels und der Industrie von jeweils fünf unterschiedlichen Produkten.

Viele der relevanten Veröffentlichungen sind aufgrund mehrfacher Publikation als Duplikate zu bewerten, insbesondere von KUO et al. und THIESING et al. sowie von FARAWAY et al., TANG et al., NAM et al. Somit reduziert sich die Anzahl der für diese Untersuchung relevanten Veröffentlichungen um 22 auf nur noch 41 Beiträge, wovon nur wenige die Komplexität hochfrequenter Zeitreihen des Handels abbilden. In Tabelle 4.3 wird eine Übersicht über die Publikationen nach der Form der Publikation und der Anzahl der verwendeten Zeitreihen in den Experimenten gegeben.

Tabelle 4.3: Publikationen zur Absatzprognose mit Neuronalen Netzen

	Konferenzbeiträge & Arbeitspapiere	Aufsätze in Zeitschriften & Dissertation
Einzelne Zeitreihe	[Jagielska und Jacob 1993; Faraway und Chatfield 1995; Kong und Martin 1995; Thiesing et al. 1995a; Thiesing et al. 1995b; Thiesing et al. 1995c; Thiesing et al. 1995d; Thiesing und Vornberger 1995; Kuo 1997; Thiesing und Vornberger 1997a; Thiesing und Vornberger 1997b; Thiesing und Vornberger 1997c; Muller 1997; Patterson et al. 1997; Yip et al. 1997; Gorzalczany 1998; Thiesing 1998; Thiesing und Vornberger 1998; Tchaban et al. 1998; Ong und Flitman 1998; Baba und Suto 2000; Alekseev und Seixas 2002; Crone 2002; Jiang 2003; Sztandera et al. 2003; Orra und Serpen 2005; Zekic-Susac et al. 2005; Ou und Chen 2006]	[Nam und Schaefer 1995; Luxhoj et al. 1996; Nam et al. 1997; Faraway und Chatfield 1998; Tchaban et al. 1998; Kuo und Xue 1998a; Kuo und Xue 1998b; Kuo und Xue 1999; Kuo 2001; Alon et al. 2001; Kuo et al. 2002; Chu und Zhang 2003; Chang et al. 2005]
Mehrere Zeitreihen	[Sharda und Patil 1990; Tang und Fishwick 1993a; Kolarik und Rudorfer 1994; Nelson et al. 1994; Poh und Jasic 1995; Zhang et al. 1998b; Zabel 2000; Shanmugasundaram et al. 2002; Thomassey et al. 2002; Yohda et al. 2002; Dhawan und O'Connor 2005; Crone et al. 2006a; Crone et al. 2006b; Crone et al. 2006c; Crone et al. 2006d]	[Tang et al. 1991; Foster et al. 1992; Tang und Fishwick 1993b; Sharda 1994; Conrad 1996; Hill et al. 1996; Ansuji et al. 1996; Bansal et al. 1998; Nelson et al. 1999; Jiang et al. 2000; Zhang et al. 2001; Ragg et al. 2002; Thomassey et al. 2005; Kotsialos et al. 2005; Liao und Fildes 2005; Zhang und Qi 2005]

Die vergleichsweise geringe Anzahl von dedizierten Publikationen zur ökonomischen Absatzprognose mit NN bestätigt die Ergebnisse vorangegangener Literaturstudien [Adya und Collopy 1998]. Absatzprognosen mit NN nehmen somit eine untergeordnete Rolle in den Forschungsanstrengungen zur Prognose mit NN ein – sowohl innerhalb der Wirtschaftswissenschaften als auch im Vergleich über wissenschaftliche Disziplinen hinweg. Dies wird auch in Abbildung 4.26 deutlich, in der die zeitliche Entstehung der Publikationen zur Absatzprognose dargestellt ist. So lassen sich pro Jahr nur durchschnittlich vier Veröffentlichungen zur Absatzprognose ermitteln.¹

Wie aus Tabelle 4.3 ersichtlich wird, besteht ein Großteil der Publikationen aus Konferenzbeiträgen und unveröffentlichten Arbeitspapieren zu einzelnen Zeitreihen, die eine geringere Validität bzw. Reliabilität als Veröffentlichungen in etablierten Fachzeitschriften und zudem einen geringeren Erkenntnisfortschritt als Veröffentlichungen auf einer Stichprobe mehrerer Zeitreihen vermuten lassen. Unter Berücksichtigung möglicher Unterschiede in der Validität und Reliabilität der Experimente werden nachfolgend ausgewählte Veröffentlichungen hinsichtlich ihrer Aussagefähigkeit und der festgestellten Güte von NN im Vergleich zu statistischen Verfahren in der Absatzprognose analysiert.

4.4.1.4 Empirische Güte der Anwendung

Die Verfahrensklasse der NN konkurriert mit diversen statistischen und heuristischen Verfahren um den empirischen Einsatz zur Prognose von Zeitreihen. Während Verfahrensvergleiche auf synthetischen Zeitreihen die grundsätzlichen theoretischen Eigenschaften und Fähigkeiten eines Verfahrens dokumentieren, zeigen Vergleiche der Prognosegüte auf einer repräsentativen Auswahl von empirischen Zeitreihen den tatsächlich zu erwartenden Zuegewinn durch ihre Anwendung in der Praxis auf. Somit müssen neuartige Prognoseverfahren

¹ Dies erklärt auch die große Aufmerksamkeit, die einzelne Publikationen innerhalb einer Anwendungsdomäne erhalten. So zählt zum Beispiel das „International Journal of Forecasting“ Veröffentlichungen zu NN zu den am zweithäufigsten bzw. fünfhäufigsten nachgefragten Veröffentlichungen in den Jahren 2001 bzw. 2002 [Elsevier 2003] und rechnet seit 2000 kontinuierlich zwei Veröffentlichungen zu NN in der Prognose zu den 15 meistgefragten Veröffentlichungen seiner Bestsellerliste.

wie das der NN anhand ihrer Prognosegüte auf empirischen Zeitreihen beurteilt werden, um sich in Theorie und Praxis zu etablieren. ADYA und COLLOPY bezeichnen die Überprüfung neuartiger Verfahren in Vergleichsstudien als „well-established tradition in forecasting research“ [Adya und Collopy 1998].¹

Der Großteil aller vorangehend dargestellten Veröffentlichungen der Absatzprognose mit NN identifizieren NN in ihren Experimenten als überlegene oder zumindest gleichwertige Verfahrensklasse.² Allerdings variiert die Aussagekraft einzelner Publikationen, so z. B. durch Publikation von fehlerhaften Experimenten, die Verwendung nur einer einzigen Zeitreihe, den Vergleich anhand verzerrter Fehlermaße oder den mangelnden Vergleich zu etablierten statistischen Benchmark-Verfahren der Prognose [Crone und Preßmar 2006]. Somit sollte der Schwerpunkt nicht auf einer Analyse der relativen Güte von NN über alle Publikationen liegen, sondern vielmehr bedeutende Publikationen mit hoher Validität und Reliabilität identifiziert und diese Ergebnisse im Anwendungskontext der Warendisposition geprüft werden. ADYA und COLLOPY [1998] identifizieren und analysieren in einer Meta-Studie 48 Publikationen zur Vorhersage mit NN in Anwendungen der Betriebswirtschaftslehre, davon 28 zur Klassifikation und 19 zu Anwendungen der Regression bzw. Punktprognose. Sie bewerten die Studien hinsichtlich ihrer Validität und Implementierung und identifizieren nur sechs zuverlässige Studien, die die Überlegenheit von NN in der Punktprognose unterstützen: [Chen et al. 1992; Ho et al. 1992; Foster et al. 1992; Connor et al. 1994; Park et al. 1991; Tang et al. 1991]. Diese stammen überwiegend aus dem Bereich der Prognose elektrischer Lasten, nur zwei befassen sich mit der Absatzprognose. Die verbleibenden 13 Publikationen lieferten dagegen keine validen Ergebnisse. Während im Fall der Klassifikation 57,1% aller Publikationen als valide eingeschätzt wurden, bestätigten für die Regression nur 31,6% ein valides Experimentdesign und somit zulässige und verlässliche Ergebnisse.

Die Erkenntnisse über die geringe Validität vieler Publikationen zur Regression werden durch zwei Meta-Studien für Publikationen der Absatzprognose bestätigt [Crone und Preßmar 2006; Crone und Graffaille 2004]. Unter Erweiterung des Beurteilungsrahmens von ADYA und COLLOPY evaluieren sie die Reliabilität der Experimente durch ausreichende Dokumentation von Modellaufbau und Ergebnissen, der Validität der Experimente durch die Berücksichtigung bedeutender Aspekte empirischer Prognoseexperimente sowie der Parametrisierung von NN. Die Ergebnisse sind enttäuschend und somit vergleichbar mit denen von ADYA und COLLOPY [ADYA UND COLLOPY 1998]: Von den 47 Publikationen zur Absatzprognose dokumentieren über 50% der Beiträge weniger als die Hälfte der getroffenen Modellierungsentscheidungen ihrer Experimente; nur 19% der Veröffentlichungen dokumentieren zwei Drittel aller für eine Replikation notwendigen Experimentparameter. Dies bestätigt grundsätzlich, dass die Mehrzahl der Autoren die Experimente mit NN nicht hinreichend dokumentiert und somit eine erfolgreiche Anwendung der publizierten NN-Architektur in ähnlichen Problemstellungen verhindert. Ferner berücksichtigten über 80% aller Veröffentlichungen weniger als die Hälfte aller relevanten Aspekte für ein objektives Experimentdesign von NN zur Prognose, z. B. unter Vernachlässigung mehrerer Initialisierungen oder einer Trennung der Datensätze.³ Nur vier Studien demonstrierten eine akzeptable Validität, was die Wissen-

¹ Nach dem International Institute of Forecasting (IIF) haben empirische Vergleichsstudien besondere Bedeutung für den Wissenszugewinn von Prognoseverfahren.

² Dies wird auch kritisch als „File-drawer“-Problem neuartiger Verfahren bezeichnet [Chatfield 2005; Armstrong 2005], das darin besteht, dass lediglich positive Studien zur Veröffentlichung gelangen und somit ein positive Verzerrung tatsächlicher Ergebnisse erfolgt. Da dieses Problem jedoch wohlmöglich alle neuartigen Verfahren und eine Vielzahl von wissenschaftlichen Forschungsbereichen und Publikationen betrifft, soll hier nicht weiter darauf eingegangen werden.

³ Während eine geringe Reliabilität häufig auf Seitenbeschränkungen in der Veröffentlichung von wissenschaftlichen Beiträgen zurückgeführt werden kann, stellt die geringe Validität der Experimente die Qualität der Veröffentlichung fundamental in Frage.

schaftlichkeit der Vorgehensweise und die generelle Qualität von Veröffentlichungen zu NN in der Absatzprognose signifikant in Frage stellt.¹

Somit können vorangegangene Veröffentlichungen nur begrenzt zur Beurteilung der Prognosegüte von NN im Vergleich zu etablierten statistischen Verfahren verwendet werden. Aus den analysierten Publikationen zeigen lediglich die folgenden Veröffentlichungen eine akzeptable Validität und Reliabilität [Alekseev und Seixas 2002; Faraway und Chatfield 1995; Faraway und Chatfield 1998; Hill et al. 1996; Liao und Fildes 2005]. ALEKSEEV und SEIXAS [2002] prognostizieren die Nachfrage nach Flügen in Brasilien im Vergleich zu konventionellen, multivariaten ökonometrischen Modellen. NN reduzieren den Prognosefehler signifikant von 4,6% MAPE auf unter 1% MAPE. FARAWAY und CHATFIELD [1995; 1998] analysieren die Fähigkeit und Robustheit von NN zur Prognose der bekannten Airline-Passenger-Data. NN erzielen demzufolge nur eine mit ARIMA-Modellen vergleichbare Prognosegenauigkeit, weisen jedoch vielfältige Probleme in der robusten Parametrisierung auf, was ihre Anwendung insbesondere für CHATFIELD in Frage stellt [1993b; 1998; 2005].

HILL, O'CONNER und REMUS [1996] analysieren 111 Zeitreihen aus der M1-Competition (vgl. Abschnitt 3.3.3). Die NN zeigten im Vergleich zu statistischen Verfahren sowie subjektiven Expertenprognosen signifikant bessere Ergebnisse in der Vorhersage monatlicher und vierteljährlicher Daten und eine vergleichbare Prognosegüte für jährliche Daten. Eine detaillierte Analyse zeigte, dass NN nicht nur bei Zeitreihen mit Diskontinuitäten und nichtlinearen Daten konventionelle Verfahren übertrafen, sondern auch bei linearen Zeitreihen. Während die Studie bzgl. der mehrfachen Analyserichtungen der Ergebnisse und des Umfangs der empirischen Zeitreihen viele vorangegangene Verfahrensvergleiche hinsichtlich der Validität übertrifft, zeigt sie erhebliche Schwächen in der Dokumentation, die eine Reproduktion der positiven Ergebnisse unmöglich machen [Zhao et al. 2003]. LIAO und FILDES [2005] untersuchen 261 Zeitreihen der Domäne über einen Prognosehorizont von 18 Beobachtungen und 5 Prognoseursprüngen in einem hoch validen Experimentdesign über mehrere Zeitreihenursprünge mit robusten Fehlermaßen und unterschiedlichen Architekturparametern der NN und mit hinreichender Reliabilität. Sie fanden eine signifikante Verringerung der Prognosefehler von 56% durch NN im Vergleich zum vorherigen Benchmark des gedämpften Trends.

Diesen positiven Ergebnissen stehen jedoch auch negative Resultate gegenüber. Im Zuge der M3-Competition, des bislang umfassendsten empirischen Verfahrensvergleichs statistischer Verfahren, lieferte der einzige Teilnehmer der Verfahrensklasse NN durch BALKIN und ORD [2000] nur unterdurchschnittliche Ergebnisse: NN waren 3,4% weniger genau als das Verfahren der Exponentiellen Glättung mit gedämpftem Trend und 4,2% weniger genau als kombinierte Prognoseverfahren [Makridakis und Hibon 2000].

Die geringe Validität und Reliabilität der gefundenen Studien und die mehrdeutigen Erkenntnisse hinsichtlich der Prognosegüte von NN erfordern eine objektive, hoch valide und reliable Beurteilung der Eignung von NN zur Absatzprognose durch ein entsprechend repräsentatives Experimentdesign in dieser Untersuchung. Die Prognosegüte von NN ist somit zusätzlich zur Untersuchung und Diskussion des Einflusses kostenminimaler Zielfunktionen auf die Disposition mit NN im empirischen Verfahrensvergleich in Abschnitt 6 festzustellen.

¹ Für nähere Ausführungen der Ergebnisse sei auf CRONE und PREBMAR [2006] verwiesen.

4.4.2 Netzwerkarchitekturen zur Prognose

4.4.2.1 Zeitreihenanalytische Modellierung

NN besitzen erhebliche Freiheitsgrade in der Modellierung, die eine problem- und datenspezifische Konfiguration ihrer Architektur in vielfältigen Anwendungen unüberwachter und überwachter, homoassoziativer sowie heteroassoziativer Lernaufgaben erlauben.¹ Ebenso gestatten MLP die Modellierung dynamischer Erklärungs- und Prognosemodellen als heteroassoziative Lernaufgabe der kausalanalytischen oder zeitreihenanalytischen Regression. Nachfolgend werden zunächst Architekturen zur Modellierung von Zeitreihenprognosen vorgestellt, gefolgt von kausalanalytischen Architekturen im nachfolgenden Abschnitt.²

Die Modellierung von NN zur zeitreihenanalytischen Prognose beruht auf der Vorhersage einer abhängigen Variablen \hat{y}_{t+h} aus n zeitlich verzögerten Realisationen derselben Variablen $y_t, y_{t-1}, \dots, y_{t-n+1}$ aus den n vorhergehenden Zeitpunkten $t, t-1, t-2, \dots, t-n+1$. Dies modelliert ein MLP in Analogie zu einem nicht-linearen autoregressiven AR(p)-Modell [Zhang et al. 1998a], wobei die Anzahl der autoregressiven Terme p der Anzahl der Eingabeneuronen n entspricht, $p = n$, und die Anzahl der abhängigen Variablen \hat{y}_{t+h} der Anzahl von Ausgabeneuronen. Im einfachsten Fall erfolgt eine Einschrittprognose $t+1$ für einen Prognosehorizont $h = 1$:

$$\hat{y}_{t+1} = f(y_t, y_{t-1}, \dots, y_{t-n+1}) \quad (4.45)$$

Die korrespondierende Architektur eines MLP zur zeitreihenanalytischen Einschrittprognose ist in Abbildung 4.28 dargestellt.

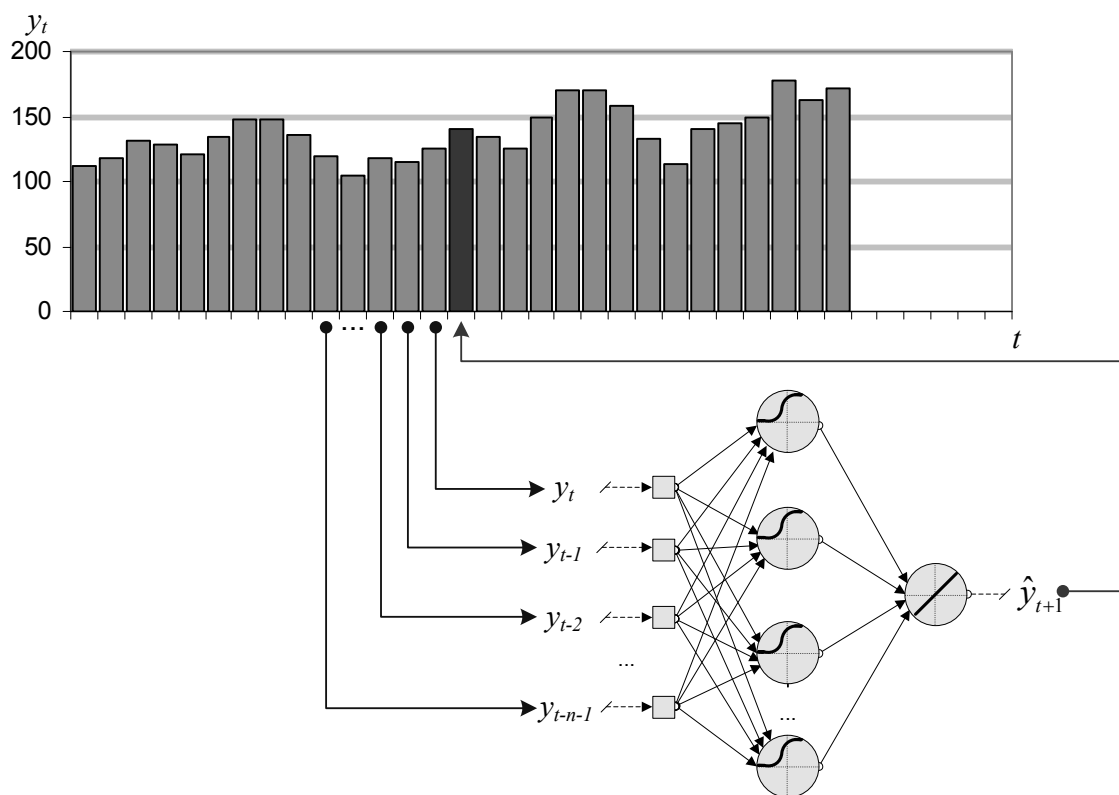


Abbildung 4.28: Architektur eines MLP zur zeitreihenanalytischen Einschritt-Prognose

¹ Vgl. die Ausführungen in Abschnitt 4.2.

² Die nachfolgenden Abschnitte verwenden die Notation und Terminologie, die bereits in Kapitel 3 eingeführt wurde. Auf eine neuerliche Definition der Variablen und Indizes wird daher verzichtet.

Abbildung 4.28 stellt ein MLP mit drei Schichten von Neuronen zur Abbildung eines nichtlinearen $AR(p)$ -Prozesses zur Einschrittprognose des nächsten Beobachtungswertes während der Parametrisierung auf bekannten Trainingsdaten dar.¹ Die Architektur besteht aus $p = n$ Inputneuronen, m versteckten Neuronen mit der logistischen Aktivierungsfunktion und einem Ausgabeneuron für $t+h$, ebenfalls mit der logistische Aktivierungsfunktion $A=[[n-m(\text{Sum}, \text{Log}, \text{Id})-1(\text{Sum}, \text{Log}, \text{Id})], \text{FF}, \text{F}]$. Die dargestellte Architektur zeigt die am häufigsten verwendete Modellierungsform aus wissenschaftlichen Veröffentlichungen zur Absatzprognose. Die Literaturanalyse von CRONE und PREßMAR [2006] belegt, dass alle Publikationen zur Absatzprognose mit NN zumindest ein MLP mit einer Schicht versteckter Neuronen in ihren Experimenten verwendeten. In 93% aller Veröffentlichungen wird die logistische Aktivierungsfunktion (Log) in den Neuronen der versteckten Schicht verwendet. 94% verwenden ein einzelnes Ausgabeneuron für Einschrittprognosen. Die Neuronen der Ausgabeschicht verwenden in 50% aller Experimente die Log Aktivierungsfunktion, in 42% die lineare Identitätsfunktion (Id) und in 8% den TanH [Crone und Preßmar 2006].

Da das MLP als nichtlinearer $AR(p)$ -Prozess keinerlei zeitliche Kontextinformationen über die Zeitreihe besitzt, müssen zur Vorhersage des nächsten Beobachtungswertes alle Strukturinformationen des datengenerierenden Prozesses im Zeitfenster des Eingangsvektors der Länge n vorhanden sein. Im Training werden dem NN die verzögerten Eingabedaten als Vektor mit den zugehörigen Ausgabedaten als als gleitendes Fenster über die Zeitreihe präsentiert. MLPs werden dazu in 84% aller Veröffentlichungen durch den Backpropagation-Algorithmus trainiert [Crone und Preßmar 2006].

Zur Mehrschrittprognose mit mehreren Prognosehorizonten $t+h$, $h = 1, 2, \dots, H$ ist unter Verwendung eines einzelnen Ausgabeneurons analog zu ARIMA-Modellen eine iterative Berechnung unter Verwendung vorangegangener Prognosewerte erforderlich, mit

$$\begin{aligned} \hat{y}_{t+1} &= f(y_t, y_{t-1}, \dots, y_{t-l}) \\ \hat{y}_{t+2} &= f(\hat{y}_{t+1}, y_t, \dots, y_{t-l+1}) \\ &\dots \\ \hat{y}_{t+H} &= f(\hat{y}_{t+H-1}, \hat{y}_{t+H-2}, \dots, y_{t-l+H}) \end{aligned} \quad (4.46)$$

Darüber hinaus können MLPs aufgrund ihrer flexiblen Architektur aus beliebig kombinierbaren, einfachen Verarbeitungselementen auch Prognosemodelle zur simultanen Mehrschrittprognose abbilden. Danach berechnet ein MLP simultan H Prognosewerte $\hat{y}_{t+1}, \hat{y}_{t+2}, \dots, \hat{y}_{t+H}$ unter Verwendung von H Ausgabeneuronen bei einem konstanten Inputvektor von n verzögerten Realisationen $y_t, y_{t-1}, \dots, y_{t-n+1}$, mit

$$\begin{aligned} \hat{y}_{t+1} &= f(y_t, y_{t-1}, \dots, y_{t-l}) \\ \hat{y}_{t+2} &= f(y_t, y_{t-1}, \dots, y_{t-l}) \\ &\dots \\ \hat{y}_{t+H} &= f(y_t, y_{t-1}, \dots, y_{t-l}) \end{aligned} \quad (4.47)$$

Da die direkte Mehrschrittprognose nicht durch mögliche Prognosefehler aus vorangegangenen Iterationen $\hat{y}_{t+1}, \hat{y}_{t+2}, \dots$ beeinflusst wird, ist theoretisch eine Verringerung der

¹ Dies ist äquivalent zur iterativen Einschritt-Prognose in der Berechnung von ARIMA-Verfahren und der Exponentiellen Glättung.

Prognosefehler bei langen Prognosehorizonten zu erwarten [Zhang et al. 1998a].¹ Insofern erweitert das MLP konventionelle $AR(p)$ -Modelle durch die Mehrschrittprognose und geht somit über die Fähigkeiten der Modellierung autoregressiver Prozesse hinaus.

Neben der gebräuchlichen Architektur zur Absatzprognose in Abbildung 4.28 kann eine individuelle Informationsverarbeitung V je Neuron $f_{u_i}(\bullet)$ oder je Schicht durch die Kombination unterschiedlicher Eingabe-, Aktivierungs- und Ausgabefunktionen erzielt werden. In der praktischen Anwendung wird davon meist kein Gebrauch gemacht, sondern eine einheitliche Informationsverarbeitung für Neuronen einer Schicht ausgewählt [Crone und Preßmar 2006]. Eine grafische Darstellung der allgemeinen Modellform eines MLP zur zeitreihenanalytischen Ein- oder Mehrschritt-Prognose ist in Abbildung 4.29 dargestellt.

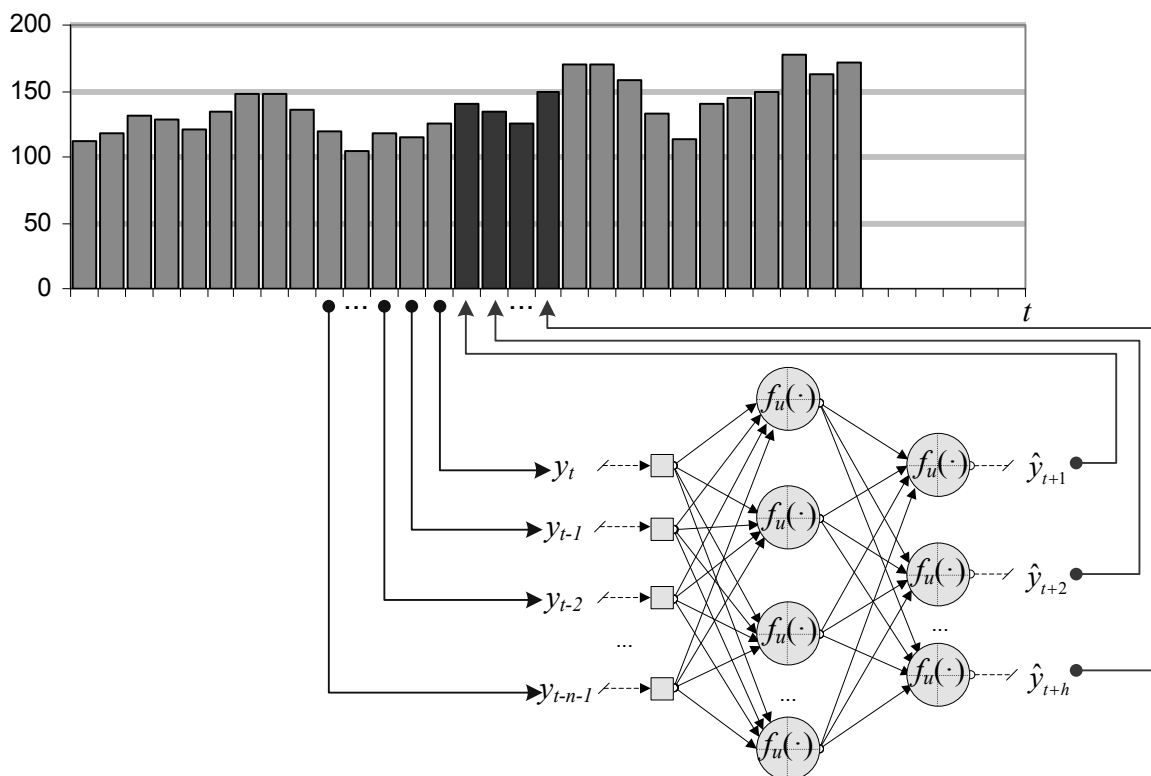


Abbildung 4.29: Architektur eines MLP zur zeitreihenanalytischen Mehrschritt-Prognose

Darüber hinaus können MLPs ebenso zur Modellierung kausalanalytischer Prognosemodelle konfiguriert werden, die im nachfolgenden Abschnitt vorgestellt werden.

4.4.2.2 Kausalanalytische Modellierung

Die flexible funktionale Abbildung beliebiger Inputdaten auf heteroassoziative Ausgabedaten erlaubt ebenso die Modellierung von MLPs zur kausalanalytischen Prognose. Dies modelliert ein MLP in Analogie zu einer nicht-linearen Regression, unter Abbildung der funktionalen Beziehung zwischen einer oder mehreren unabhängigen Erklärungsvariablen x_i mit metrischem oder nominalem Skalenniveau und der abhängigen Variablen y in der Form

$$\hat{y}_t = f(x_{1,t}, x_{2,t}, \dots, x_{i,t}) \quad (4.48)$$

¹ Dies wurde bislang jedoch lediglich postuliert und erfordert eine empirische Überprüfung in einem separaten Vergleich.

Die Anzahl der Eingabeneuronen n entspricht der Anzahl i unabhängiger Variablen x_i , die Anzahl der abhängigen Variablen y_j der Anzahl der Ausgabeneuronen. Im seltenen Fall einer univariaten Regression durch NN wird der funktionale Zusammenhang zwischen einer einzelnen unabhängigen Variablen x und einer abhängigen Variablen y durch ein MLP mit einem Eingabeneuron und einem Ausgabeneuron modelliert.¹ Ein MLP zur univariaten Regression ist in Abbildung 4.31 dargestellt, wobei die Eingabe- und Ausgabedaten x_t und y_t nicht notwendigerweise zeitlich geordnete Beobachtungen, sondern auch zeitlose, ungeordnete Realisationen einer stationären Verteilung darstellen können.

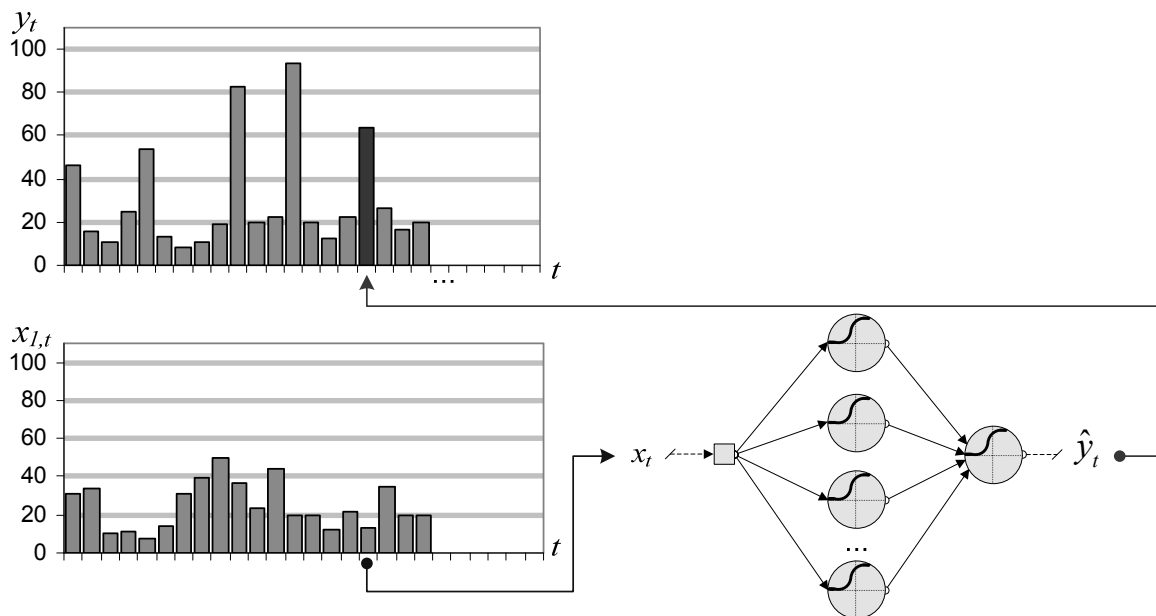


Abbildung 4.30: Architektur eines MLP zur univariaten, nichtlinearen Regression

Die Flexibilität der Verfahrensklasse der MLP erlaubt eine problemlose Erweiterung der univariaten Modellierung auf weitere kausalanalytische Modellformen. So können durch Anlegen weiterer unabhängiger Variablen x_i an zusätzlichen Eingabeneuronen Modelle der multivariaten Regression abgebildet werden. Unter Erweiterung des Eingabevektors um zeitlich verzögerte Realisationen der abhängigen und der unabhängigen Variablen kann das MLP zur Abbildung allgemeiner Modelle der dynamischen Regression formuliert werden. Das MLP erlernt die funktionale Abbildung zwischen der abhängigen Variablen y und einer oder mehreren unabhängigen Variablen x_i , sowie zeitlich verzögerten Realisationen der unabhängigen $x_{i,t-l}$ und der abhängigen Variablen y_{t-l} , mit²

$$\hat{y}_{t+h} = f(x_{1,t+h}, x_{1,t+h-1}, \dots, x_{1,t+h-l}, \dots, x_{i,t+h}, x_{i,t+h-1}, \dots, x_{i,t+h-l}, \dots, y_t, y_{t-1}, y_{t-2}, \dots, y_{t-l}) + \varepsilon_{t+h}. \quad (4.49)$$

Die Verwendung mehrerer Ausgabeneuronen erlaubt eine simultane Mehrschrittprognose einer abhängigen Variablen zu mehreren Prognosehorizonten, $\hat{y}_{t+1}, \hat{y}_{t+2}, \dots, \hat{y}_{t+H}$.³ In Abbil-

¹ Der Sonderfall der Regression gegen die Zeit als unabhängige Variable stellt einen Sonderfall dessen dar, wobei die Zeit t als Zeitreihe eines laufenden Zeitindexes $t = \{1, 2, \dots, T\}$ als einzige Information an das Netz angelegt wird.

² Vgl. Formel 3.7 in Abschnitt 3.1.2.4. Für eine nähere Interpretation der Erklärungsmodelle wird auf Kapitel 3 verwiesen.

³ Darüber hinaus ist theoretisch auch die simultane Schätzung mehrerer unabhängiger Variablen $\hat{y}_{1,t+h}, \hat{y}_{2,t+h}, \dots, \hat{y}_{j,t+h}$ im Sinne von Modellen der Vektor-Autoregression möglich. Diese ist jedoch bislang nicht hinreichend untersucht worden [Jiang et al. 2000; Moshiri und Cameron 2000; Binner et al. 2004; Moshiri 1998] und unterbleibt daher an dieser Stelle.

dung 4.31 wird die allgemeine Modellform eines MLP unter Verwendung zeitlich verzögerter Realisationen zur Mehrschrittprognose einer abhängigen Variablen abgebildet.

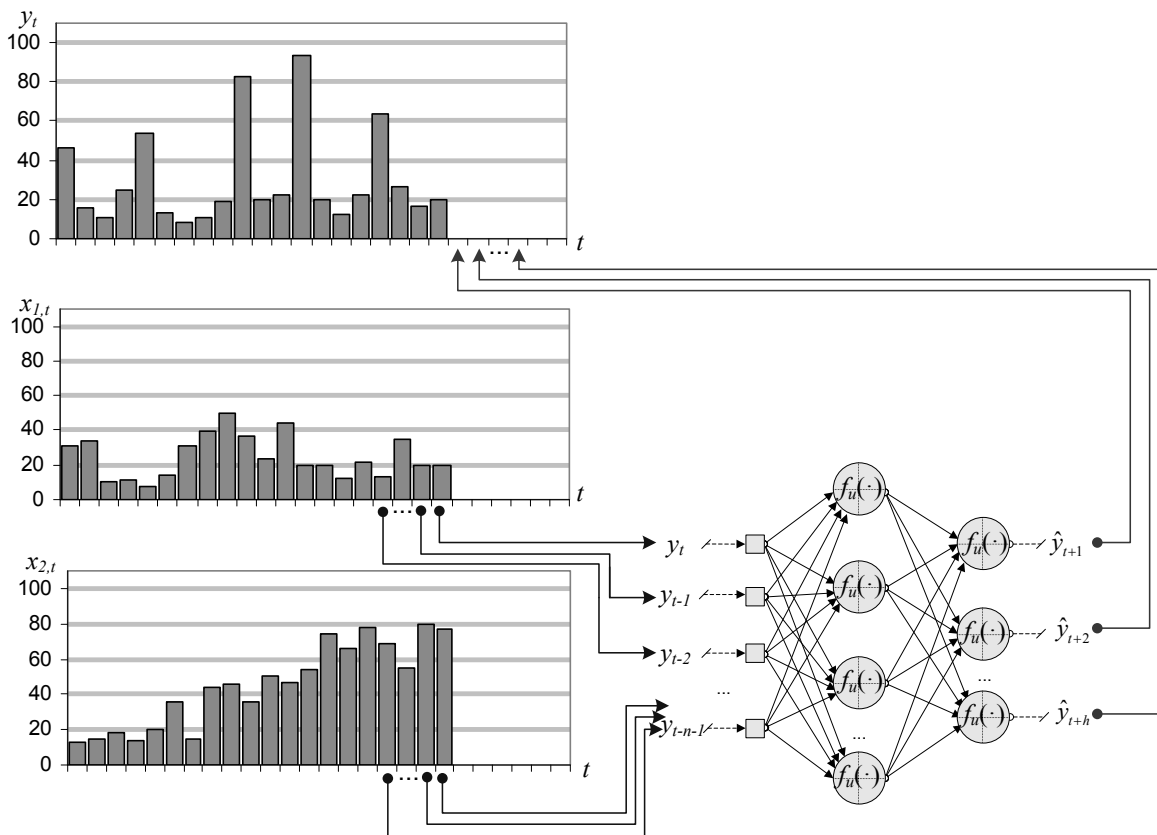


Abbildung 4.31: Architektur eines MLP zur dynamischen Regression

Als Sonderfall der dynamischen, kausalanalytischen Modellierung sind auch Interventions- oder Eventmodelle durch MLPs realisierbar. Die Verwendung von Erklärungszeitreihen aus binären Dummy-Variablen $x_i \in \{0;1\}$ erlaubt die Kodierung besonderer Ereignisse wie Marketingaktionen (in Form von Werbeaktivitäten, Preisaktionen etc.), Kalendereffekte (Weihnachten, Ostern, Schulferien und Feiertage) sowie Ausreißer, fehlende Datenpunkte usw. und ihre explizite Einbeziehung in die Modellbildung.¹ Grundsätzlich sind auch hier unterschiedliche Modellierungen realisierbar, sowohl unter Nutzung einer einzelnen unabhängigen Dummyvariablen zum Prognosezeitpunkt $t+h$ in der Einzschrittprognose von \hat{y}_{t+h} , mehrerer Dummy-Variablen für alle zukünftigen Prognosezeitpunkte $t+1, t+2, \dots, t+H$ in der Mehrschrittprognose von $\hat{y}_{t+1}, \hat{y}_{t+2}, \dots, \hat{y}_{t+H}$, sowie zusätzlicher Dummy-Variablen zu den lags der verzögerten Variablen $t, t-1, \dots, t-1-n$ in der Vergangenheit, was den Übergang zur Modellen der dynamischen Regression darstellt. In der einfachsten Form x_{t+h} entspricht die Architektur einem nichtlinearen ARX(p)-Modell zur Interventionsmodellierung. Abbildung 4.32 stellt eine allgemeine Modellarchitektur für eine einzelne Zeitreihe von Dummy-Variablen dar.

¹ Entsprechend wird für jede Erklärungsvariable eine separate, binärkodierte Zeitreihe benötigt. Darüber hinaus dienen Dummy-Variablen auch zur Modellierung zeitlicher Komponenten wie Monaten, Kalenderwochen oder Wochentagen zur expliziten Abbildung saisonaler Komponenten des Zeitablaufes in autoregressiven Modellansätzen. Für eine Darstellung dieser sowie weiterer typischer Anwendungsbereiche wird auf Literatur zur dynamischen Regression verwiesen [Harvey 1993; Pankratz 1991].

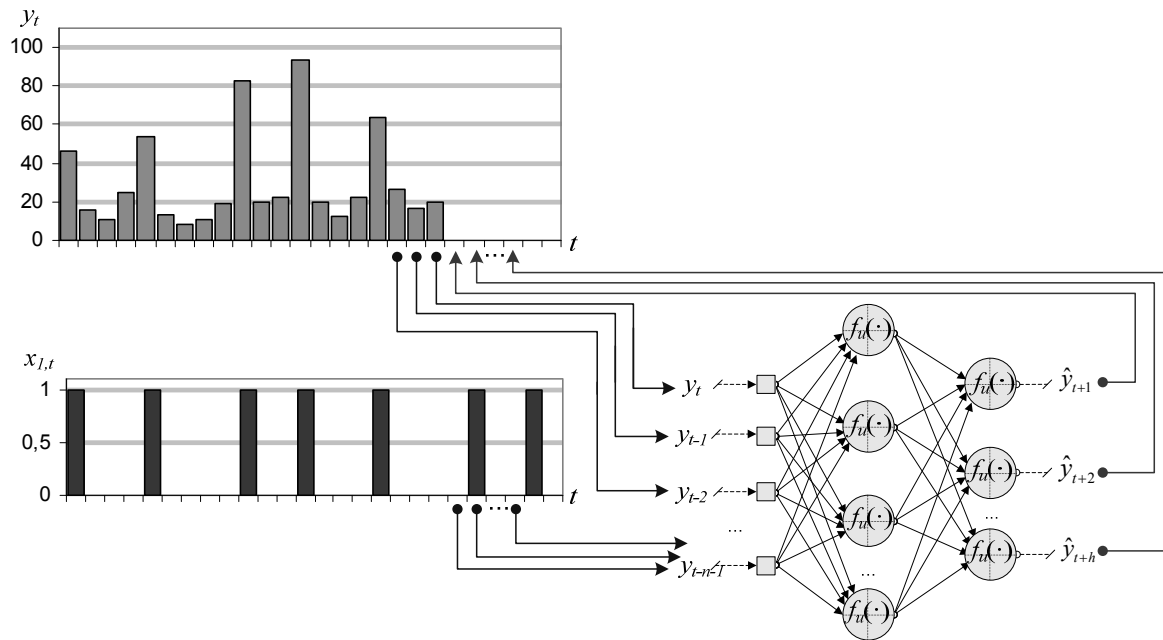


Abbildung 4.32: Architektur eines MLP zur kausalen Prognose mit Interventionsvariablen

Das MLP erlernt die funktionale Abhängigkeit binärer Interventionsvariablen ebenso wie der metrisch skalierten unabhängigen Variablen oder zeitlich verzögerten abhängigen Variablen eigenständig aus den vergangenen Beobachtungen und erlaubt so eine flexible Modellierung und Informationsverarbeitung unter Nutzung von Erklärungszeitreihen unterschiedlicher Skalenniveaus und Informationsgehalte.¹ Ihre Funktionsweise wird nachfolgend im Vergleich zu statistischen Verfahren diskutiert, um Erkenntnisse über ihr Anwendungspotenzial im Vergleich zu statistischen Verfahren abzuleiten.

4.4.2.3 Äquivalenz zu statistischen Prognoseverfahren

Die Komplexität der mathematischen Formulierung eines MLP sowie die grafische Notation verhindern vielfach die Interpretation ihrer Informationsverarbeitung sowie den Vergleich ihrer Funktion mit bestehenden statistischen Verfahren. Als Konsequenz werden MLPs vielfach als Blackbox-Verfahren bezeichnet, obwohl sie einer statistischen bzw. deskriptiven Analyse durchaus zugänglich sind. Der nachfolgende Abschnitt untersucht die Informationsverarbeitung des MLP im Vergleich zu etablierten Prognoseverfahren, deren statistische Eigenschaften und Funktionsweise hinreichend bekannt sind. Hierzu werden zunächst die Eigenschaften einzelner Neuronen im Vergleich zu statistischen Verfahren betrachtet und diese anschließend zu Netzwerk-Architekturen von Neuronen kombiniert, um so Erkenntnisse über ihre statistische Funktion und ihre Fähigkeiten abzuleiten.

Im einfachsten Fall empfängt ein einzelnes Neuron u_j mit linearer Aktivierungsfunktion (Id) zur Abbildung der abhängigen Prognosevariablen \hat{y}_t Eingabedaten von einer einzelnen, zeitlich verzögerten Realisation der abhängigen Variablen über ein einzelnes Eingabeneuron u_i . Dies entspricht der Informationsverarbeitung in einem versteckten Neuron als Teil einer umfassenden MLP-Architektur oder auch einem MLP ohne versteckte Neuronen der Form

¹ Neben diesen Ansätzen zur Prognose mit NN existiert eine Vielzahl von Anwendungsmöglichkeiten von NN im Kontext der Prognose [Janetzke und Falk 2006]. So werden NN zur gewichteten Kombination von ARI-MA-Prognosen [Shi 2000], oder auch zur Selektion geeigneter Prognoseverfahren verwendet [Sohl und Venkatachalam, 2000]. Diese Anwendungen erlauben die Erweiterung der NN zur Abbildung kompletter Expertensysteme auf Basis der Konzepte neuronaler Informationsverarbeitung.

$A=[[1-0-1(\text{Sum}, \text{Id}, \text{Id})], \text{FF}, \text{F}]$.¹ Das Neuron formuliert so das Prognosemodell eines linearen autoregressiven AR(1)-Prozesses, welches in Abbildung 4.33a in grafischer Notation abgebildet ist. Wird die Eingabeinformation des Neurons auf p zeitlich verzögerte Realisationen der abhängigen Variablen $y_t, y_{t-1}, \dots, y_{t-n-1}$ erweitert, die durch n Eingabeneuronen modelliert werden, wird ein konventionelles lineares AR(p)-Modell formuliert (siehe Abbildung 4.33 b.), mit $A=[[n-0-1(\text{Sum}, \text{Id}, \text{Id})], \text{FF}, \text{F}]$ und $n = p$.

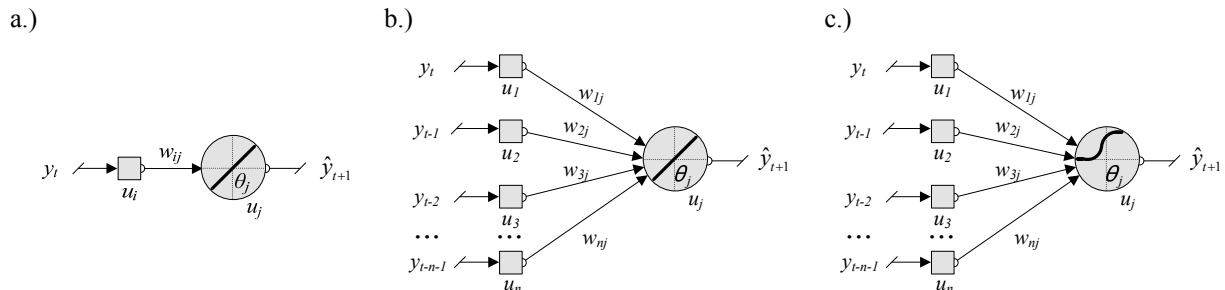


Abbildung 4.33: Architektur eines MLP zur Modellierung eines linearen AR(1)-Prozesses (a.), eines linearen AR(p)-Prozesses (b.) und eines nichtlinearen AR(p)-Prozesses (c.)

Die mathematische Notation der Informationsverarbeitung des Neurons in Abbildung 4.33b verdeutlicht die Äquivalenz zu einem AR(p)-Prozess:²

$$\begin{aligned} \hat{y}_{t+1} &= f(y_t, y_{t-1}, y_{t-2}, \dots, y_{t-n-1}) \\ &= y_t w_{1j} + y_{t-1} w_{2j} + y_{t-2} w_{3j} + \dots + y_{t-n-1} w_{nj} - \theta_j \end{aligned} \quad (4.50)$$

Die Verwendung einer nichtlinearen Aktivierungsfunktion im Ausgabeneuron erweitert das Prognosemodell weiter zu einem nichtlinearen AR(p)-Prozess, auch als NAR(p)-Prozess bezeichnet. Es wird in Abbildung 4.33c dargestellt und entspricht einem MLP der Architektur $A=[[n-0-1(\text{Sum}, \text{Log}, \text{Id})], \text{FF}, \text{F}]$. Da statistische NAR(p)-Prozesse kein leicht interpretierbares Äquivalent als Prognoseverfahren besitzen, erlauben sie jedoch keine intuitive Interpretation. Die Informationsverarbeitung in linearen Neuronen zeigt hingegen alle statistischen Eigenschaften, die von der Anwendung konventioneller autoregressiver Prozesse bekannt sind (vgl. Kapitel 3.2.2) und in der Literatur bereits hinreichend analysiert und dokumentiert wurden [siehe u.a. Harvey 1993; Pankratz 1991].

Die Analyse von Neuronen bei kausalanalytischer Modellierung erfolgt in Übereinstimmung mit der zeitreihenanalytischen Interpretation eines Neurons als Prognosemodell eines linearen oder nichtlinearen AR(p)-Prozesses. Die Ausgabe des Prognosemodells y erfolgt durch ein einzelnes Neuron u_j mit linearer Aktivierungsfunktion, welches die abhängige Variable repräsentiert. Das Neuron u_j erhält Eingabedaten einer unabhängigen Variablen x_l durch ein Eingabeneuron u_i und berechnet so eine konventionelle, univariate lineare Regression. Die Informationsverarbeitung entspricht einem MLP mit der Architektur $A=[[n-0-1(\text{Sum}, \text{Id}, \text{Id})], \text{FF}, \text{F}]$ und ist in Abbildung 4.34a dargestellt. Unter Erweiterung der Modellform auf mehrere unabhängige Variablen x_1, x_2, \dots, x_n und n Eingabeneuronen $u_{i1}, u_{i2}, \dots, u_{in}$, kann ein Neuron eine multiple lineare Regression berechnen, deren Varianten in Abbildung 4.34b dargestellt sind.

¹ Ein einzelnes Neuron kann bereits als Architektur eines MLPs aufgefasst werden. Dies entspricht auch einer MLP-Architektur ohne versteckte Neuronen mit einem einzelnen Eingabeneuron u_i und einem Ausgabeneuron u_j , die in der universellen Notation hinreichend beschrieben werden kann, mit $A=[V, T, W, K]$.

² In der Originalnotation von AR(p)-Modellen entspricht w_{ij} den Parametern ϕ_i sowie θ_j der Konstanten c (vgl. Kapitel 3). Auf eine Darstellung der Notation aus Abbildung 4.33 a.) als Sonderfall aus b.) wird verzichtet.

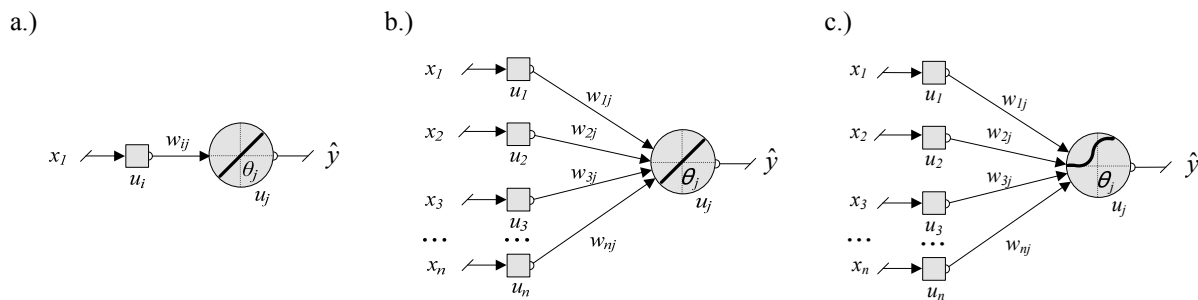


Abbildung 4.34: Architektur eines MLP zur Modellierung einer univariaten linearen Regression (a.), einer multivariaten linearen Regression (b.) und einer nichtlinearen logistischen Regression (c.)

Die Identität von linearer Regression und Informationsverarbeitung in einem Neuron mit linearen Verarbeitungsfunktionen wird in der mathematischen Formulierung der grafischen Notation in Abbildung 4.34b unter Verwendung der Symbole der Diskussion von NN belegt:

$$\hat{y}_t = x_{1,t}w_{1j} + x_{2,t}w_{2j} + x_{3,t}w_{3j} + \dots + x_{i,t}w_{ij} - \theta_j \quad (4.51)$$

Darüber hinaus kann durch die Verwendung einer Architektur mit nichtlinearer Aktivierungsfunktion ein abweichendes Modellverhalten erzielt werden, welches in Abbildung 4.34c als $A=[[n-0-1(\text{Sum}, \text{Log}, \text{Id})], \text{FF}, \text{F}]$ dargestellt ist. Unter Verwendung der logistischen Aktivierungsfunktion realisiert das Neuron eine logistische Regression zur Vorhersage der Wahrscheinlichkeit, ob eine abhängige Variable mit binärem Skalenniveau $[0, 1]$ den Wert 0 oder 1 annimmt, im Sinne einer binären Klassifikation.¹

Ein einzelnes Neuron hat somit die Fähigkeit, sowohl lineare als auch nichtlineare autoregressive Prozesse $\text{AR}(p)$ der Zeitreihenanalyse abzubilden, zudem Modelle der univariaten, der multivariaten und der logistischen Regression zur Kausalanalyse.² Die Fähigkeit zur Approximation beliebiger autoregressiver oder kausaler Formen der Regression erlaubt auch die Modellierung ausgewählter Sonderfälle eines $\text{AR}(p)$, die als eigenständige Verfahren der Zeitreihenanalyse besondere Aufmerksamkeit und Anwendung in der Praxis gefunden haben. Hierzu zählen insbesondere die Naiven Verfahren für stationäre, trendbehaftete und saisonale Zeitreihen, NF1, NF2 und NF3, gleitende sowie gewichtete Mittelwerte sowie näherungsweise auch Ansätze der einfachen Exponentiellen Glättung (vgl. Kapitel 3.2.2).

Ein Neuron kann durch die Anpassung der Verbindungsgewichte w_{ij} und des Schwellwertes θ_j unterschiedliche Modellformen abbilden, die eine Interpretation als konventionelle statistische Prognoseverfahren erlauben.³ Durch Annäherung der Parameter an die Verbindungsgewichte $\{w_{1j} \doteq 1; w_{2j}, w_{3j}, \dots, w_{nj} \doteq 0\}$ und die Annäherung des Schwellwertes an 0, mit $\theta_j \doteq 0$, entspricht die Informationsverarbeitung eines Neurons mit linearer Aktivierungsfunktion dem Naiven Prognoseverfahren NF1, welches als bedeutender Benchmark für stationäre Zeitreihen gilt (siehe Abbildung 4.37 a.).

¹ Die gleichzeitigen oder zeitlosen Ausprägungen der abhängigen und unabhängigen Variablen in (4.51) können darüber hinaus um zeitlich verzögerte Realisationen der erklärenden Variablen $x_{1,t}, x_{1,t-1}, \dots, x_{1,t-n-1}$ sowie zeitlich verzögerte Realisationen der abhängigen Variablen y_{t-n-1} erweitert werden; somit kann die Informationsverarbeitung eines Neurons zu Modellformen der dynamischen Regression erweitert werden. Auf eine explizite Abbildung wird jedoch verzichtet.

² Die Anwendung konventioneller Lernverfahren wie des BackpropagationAlgorithmus zur Approximation der Prognosemodelle erfolgt dabei unabhängig von der spezifischen Ausgestaltung der linearen oder nichtlinearen Netzwerkarchitektur aus einem oder mehreren Neuronen. Dies erlaubt die beliebige Modellierung von MLPs in einem konsistenten Modellierungsansatz und unter Verwendung einer Software.

³ Dies kann sowohl automatisch im Trainingsprozess gesteuert durch die Trainingsfehler erfolgen als auch gezielt durch die Ex-ante-Konfiguration der Netzwerkarchitektur unter Anwendung von A-priori-Strukturwissen.

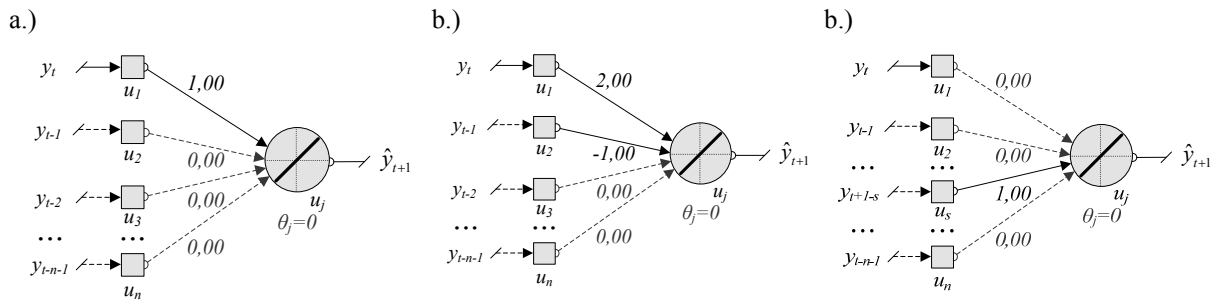


Abbildung 4.35: Architektur eines MLP zur Modellierung der Naiven Verfahren NF1 (a.), NF2 (b.) und NF3 (c.)

Unter Abbildung einer zeitlichen Verzögerung von einer Saison $t+L-s$ erlaubt ein einzelnes Neuron darüber hinaus auch die Abbildung des Naiven Prognoseverfahrens NF3 für saisonale Zeitreihen (siehe Abbildung 4.37c.). Unter Verwendung einer vorangegangenen Beobachtung $t-1$ ist darüber hinaus eine Abbildung trendbehafteter Verläufe durch das Naive Verfahren NF2 möglich, welches in Abbildung 4.37b dargestellt ist. Ein einzelnes Neuron kann somit allein durch die Anpassung seiner Parameter im Lernprozess sowohl stationäre, trendbehaftete als auch saisonale Prognosemodelle approximieren. Die Fähigkeit zur Abbildung des NF2 zeigt ferner, dass ein Neuron der versteckten Schicht durch eine geeignete Wahl der Verbindungsgewichte auch erste oder saisonale Differenzen einer Zeitreihe bilden kann, im Sinne der Integration in ARIMA(0,d,0)-Prozessen bzw. saisonalen Integration in SARIMA(0,d,0) (0,D,0)_s-Prozessen. Somit ist ein einzelnes Neuron explizit zur Modellierung linearer und nichtlinearer ARI(p,d) als auch zur Modellierung saisonaler SARI(p,d)(P,D)_s-Prozesse fähig.

Darüber hinaus kann ein Neuron auch Verfahren der gleitenden Mittelwerte abbilden. Unter Annäherung an identisch verteilte Verbindungsgewichte w_{ij} , mit $w_{1j} \cong w_{2j} \cong \dots \cong w_{nj}$, w_{ij} , $\sum_{i=1}^n w_{ij} = 1$, und unter Annäherung des Schwellwertes an $\theta_j \doteq 0$ bildet ein Neuron das Prognosemodell eines gleitenden Mittelwertes ab, welches in Abbildung 4.36a. dargestellt ist.

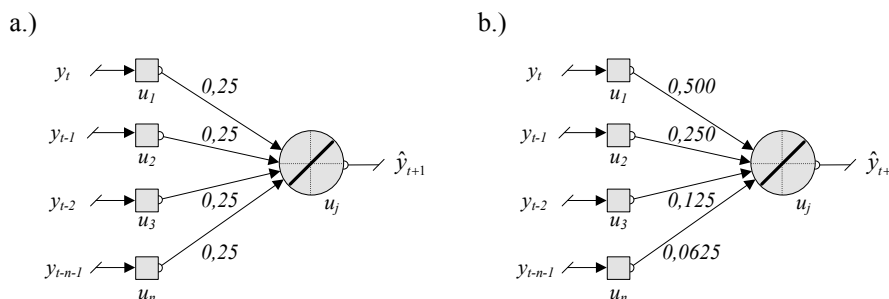


Abbildung 4.36: Architektur eines MLP zur Modellierung eines gleitenden Mittelwertes (a.) und eines exponentiell gewichteten gleitenden Mittelwertes (b.)

Unter Abbildung unterschiedlicher Verbindungsgewichte w_{ij} , mit $\sum w_{ij} = 1$ und $\theta_j = 0$, können ferner beliebig gewichtete gleitende Mittelwerte als Sonderfall eines AR(p)-Prozesses abgebildet werden. Bei exponentieller Verteilung der Verbindungsgewichte w_{ij} , mit $\sum w_{ij} = 1$ und $\theta_j = 0$, kann ein Neuron bei ausreichender Länge des Eingabevektors n das Prognoseverfahren der einfachen Exponentiellen Glättung (S.ES) approximativ modellieren. Seine Architektur ist in Abbildung 4.36b für $\alpha = 0,5$ dargestellt.¹ Unter Aufgabe der bisherigen

¹ Diese Formulierung kann jedoch nur approximativ erfolgen, da die einfache S.ES einem ARIMA(0,1,1)-Prozess entspricht. Während dieser nur als MA(1)-Prozess der ersten Differenzen einer Zeitreihe effizient modelliert werden kann, erfordert seine Abbildung als AR(p)-Prozess einen unendlich langen Eingabevektor verzögerter Realisationen von y . Siehe hierzu auch die anschließende Diskussion zur Approximation von MA(q)-Prozessen durch AR(p)-Prozesse endlicher Länge. Andernfalls kann aufgrund der exponentiellen Abnahme der Gewichte im S.ES selbst für ein niedriges $\alpha=0,1$ bereits ein Eingabevektor der Länge $n=12$ eine hinreichende Approximation gewährleisten; für höhere Alpha-Faktoren entsprechend noch weniger.

Annahme der Addition aller Verbindungsgewichte zu 100%, $\sum w_{ij} = 1$, kann ein Neuron darüber hinaus auch Zeitreihen mit einem positivem Trend durch $\sum w_{ij} > 1$ sowie mit negativem Trend durch $0 < \sum w_{ij} < 1$ als gleitende Mittelwerte prognostizieren.¹ Nichtlineare Architekturen der Mittelwertbildung entziehen sich jedoch einer solchen Interpretation.

Ein Neuron mit nichtlinearer Aktivierungsfunktion erlaubt jedoch auch die Abbildung eines annähernd linearen Systemverhaltens mit linearer Aktivierungsfunktion. Durch die Bestimmung des Schwellwertes θ_j in Relation zu den Verbindungsgewichten w_{ij} und den Eingabewerten y_i, x_i kann die Informationsverarbeitung auf den linearen Bereich der semilinearen Aktivierungsfunktion beschränkt werden [Neuneier und Zimmermann 1998]; somit können auch nichtlineare Architekturen von MLPs die vorangehend gezeigten Ausprägungen linearer Prognosemodelle approximieren. Zudem können im Lernprozess die Informationen einzelner Eingabeneuronen durch Erlernen von Verbindungsgewichten nahe Null, $w_{ij} \doteq 0$, ausgeschaltet werden. Somit bieten MLPs umfangreiche Möglichkeiten zum Erlernen einer Vielzahl von Prognosemodellen, die in der statistischen Zeitreihenanalyse als Benchmarks gelten.

Ein einzelnes Neuron erlaubt so durch die Parametrisierung der Verbindungsgewichte w_{ij} und der Schwellwerte θ_j eine selektive Abbildung elementarer Prognosemodelle aus den Ansätzen der Naiven Verfahren NF1, NF2 und NF3, gleitender Mittelwerte, linearer und nichtlinearer autoregressiver Ansätze sowie der univariaten, multivariaten und logistischen Regression. Darüber hinaus ist jedes einzelne Neuron in der Lage, durch Minimierung der Fehler im Lernprozess gemäß der präsentierten Datenmuster unterschiedliche, datensatzspezifische Prognosemodelle abzubilden. Dies entspricht einer simultanen Modellselektion und Modellparametrisierung, was eine datensatz- und problemadäquate Modellierung eines MLP allein durch Vorgabe einer einfachen, flexiblen Struktur zur Informationsverarbeitung zulässt, ganz im Sinne der semi-parametrischen Modellierung nach DORFFNER [1996].

Die Interpretation einzelner Neuronen als lineare oder nichtlineare autoregressive Modelle erlaubt auch die weiterführende Interpretation von mehrlagigen MLPs. Dazu wird zunächst die Informationsverarbeitung in den einzelnen Neuronen und Schichten eines mehrlagigen MLP bei zeitreihenanalytischer Modellierung diskutiert, siehe zum Beispiel Abbildung 4.28 oder auch Abbildung 4.37 für ein MLP mit rein linearen Aktivierungsfunktionen.² Wie zuvor dargestellt berechnet jedes versteckte Neuron $u_{n+1}, u_{n+2}, \dots, u_{n+m}$ ein lineares AP(p)- oder nichtlineares NAR(p)-Modell aus den Eingabedaten $y_t, y_{t-1}, \dots, y_{t-n-1}$, unabhängig von den anderen Neuronen einer Schicht. Durch die Initialisierung der Verbindungsgewichte mit unterschiedlichen Startwerten zu Beginn des Trainingsprozesses können die Gewichte eines Eingabeneurons u_n für jedes Neuron der versteckten Schicht u_m in abweichende lokale oder globale Minima konvergieren. Jedes versteckte Neuron bildet so im Lernprozess unterschiedliche Verbindungsgewichte zu den jeweiligen Neuronen der Eingabeschicht, so dass die versteckte Schicht m unterschiedliche Regressionsmodelle zur Abbildung unterschiedlicher Systemcharakteristika der Zeitreihe mit abweichenden Prognoseausgaben enthält.³ Das Neuron der Ausgabeschicht u_{n+m+1} kombiniert anschließend die Prognosen aller versteckten

¹ Dies erlaubt in mehrschichtigen MLPs theoretisch auch die Bildung von Mittelwerten aus mehreren Trendberechnungen und somit die Abbildung einer Reihe von Trendmodellen.

² Die nachfolgende Interpretation der Funktionsweise lässt sich entsprechend auch auf Anwendungen der Kausalanalyse übertragen. Dabei können die versteckten Neuronen zudem als latente Variable betrachtet werden, deren Interaktion anhand der Verbindungsgewichte w_{ij} interpretiert werden kann [Buckler 2001].

³ Dies wird weiter durch die zufällige Reihenfolge der präsentierten Datenmuster in iterativen Online-Lernverfahren und die vielfach suboptimale Parametrisierung der lokalen Lernverfahren unterstützt. Somit bilden auch versteckte Neuronen mit linearen Aktivierungsfunktionen unterschiedliche Verbindungsgewichte.

Neuronen durch eine lineare oder nichtlineare Regression und gibt sie als Prognosewert der abhängigen Variablen aus.¹

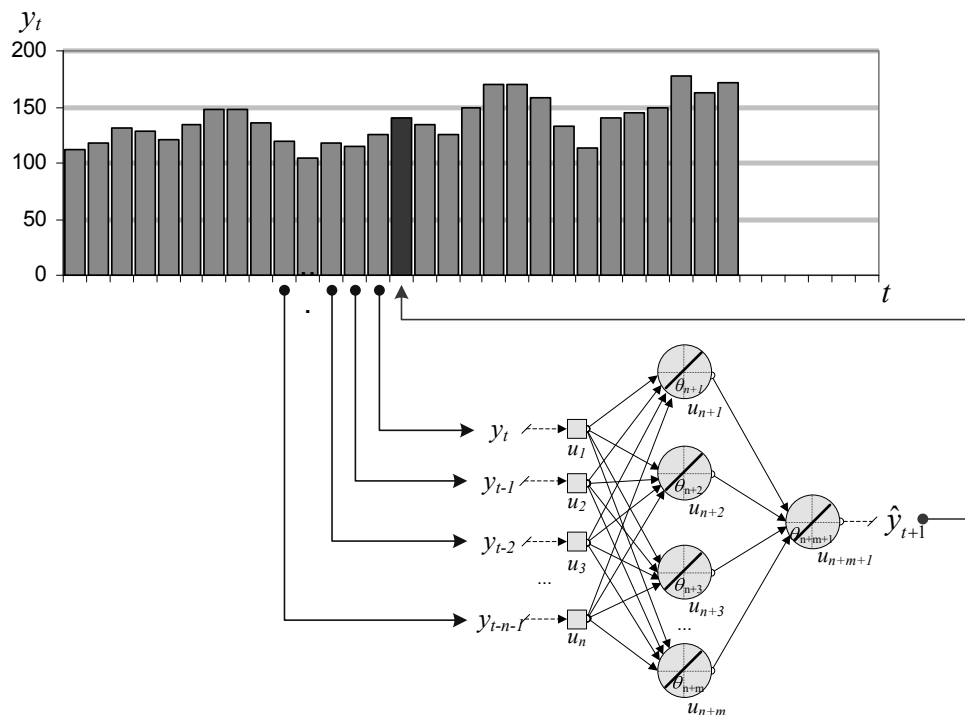


Abbildung 4.37: Architektur eines MLP zur Modellierung eines linearen $AR(p)$ -Prozesses

Unter Verwendung einer linearen Aktivierungsfunktion im Ausgabeneuron und unter Annäherung seines Schwellwertes an 0, $\theta_{n+m+1} \doteq 0$, bildet dies eine gewichtete Summe der m Prognosewerte der vorgelagerten, versteckten Neuronen im Sinne eines gewichteten Mittelwertes. Dies erlaubt auch die Interpretation eines MLP als interne Abbildung eines gewichteten Expertengremiums aus nichtlinearen autoregressiven $NAR(p)$ -Prognosen.² Diese Interpretation erscheint deshalb besonders interessant, weil in unterschiedlichen Anwendungen vielfach belegt werden konnte, dass die Kombination einzelner Prognoseverfahren zu einer erhöhten Prognosegenauigkeit führt [Armstrong 2001d; Chan et al. 2004; de Menezes et al. 2000; Fang 2003; Tang et al. 2003; Terui und van Dijk 2002; Zou und Yang 2004]. Darüber hinaus existieren vielfältige Ansätze zur Kombination einzelner NN [Hayashi und Setiono 2002; Sharkey 1999] sowie von NN und statistischen Ansätzen [Hansen und Nelson 1997; Jerez-Aragones et al. 2003; Tseng et al. 2001; Wedding und Cios 1996].

Alternativ kann die Kombination mehrerer Neuronen zu einem mehrlagigen MLP auch als Kombination disjunkter Prognosemodelle interpretiert werden. Das MLP kann theoretisch innerhalb einer Architektur simultan unterschiedliche Prognosemodelle in unterschiedlichen Neuronen der versteckten Schicht abbilden, beispielsweise das Naive Verfahren NF1 und mehrere Mittelwerte unterschiedlicher Länge und Gewichtung für eine stationäre Zeitreihe, und diese Einzelprognosen anschließend durch eine gewichtete Gremienbildung kombinieren.

¹ Bei Mehrschrittprognosen werden die Ausgaben der versteckten Neuronen entsprechend für jeden Prognosehorizont h unabhängig voneinander in einem Neuron der Ausgabenschicht kombiniert.

² ZELL [2000, S. 89f.] weist dementsprechend darauf hin, dass bei der Verwendung ausschließlich linearer Aktivierungsfunktionen in den versteckten und den Ausgabeneuronen eine identische Berechnungsvorschrift durch eine kombinierte Linearkombination ohne versteckte Neuronen erzielt werden kann, unter verringertem Rechenaufwand. Durch das Training mit dem Backpropagation-Algorithmus unter Verwendung einer quadratischen Zielfunktion stimmt die Berechnung des MLP mit der eines konventionellen $AR(p)$ -Modells überein. Erst durch Erweiterung der Aktivierungsfunktionen der versteckten Schicht auf nichtlineare Aktivierungsfunktionen, wie im MLP in Abbildung 4.28 dargestellt, bildet sich ein komplexeres Systemverhalten.

Alternativ können auch versteckte Neuronen mit unterschiedlichen Modellen zur stationären, trendbehafteten oder saisonalen Prognose zu einem der neun archetypischen Zeitreihenmuster kombiniert werden (vgl. Abschnitt 3.3), womit eine Zerlegung und Rekombination einzelner Zeitreihen-Komponenten erreicht werden kann.¹

Es wird deutlich, dass mehrlagige MLPs eine Vielzahl von Prognosemodellen als Unterklasse besitzen und sowohl die Modellierung linearer sowie nichtlinearer Modelle der Autoregression, der univariaten, multivariaten und logistischen Regression als auch einfacher Prognoseverfahren als deren Sonderfälle enthalten. Darüber hinaus zeigt ZHANG [1998b; 2001] in einem empirischen Verfahrenvergleich die Fähigkeit von MLPs zur validen und reliablen Approximation und Prognose von bilinearen, nichtlinearen Threshold-Autoregressive-(TAR-), Smooth-Transition-Autoregressive-(STAR-) oder auch heteroskedastischen ARCH- und GARCH-Prozessen auf. MLPs können daher als übergeordnete Klasse herkömmlicher, statistischer Prognoseverfahren betrachtet werden.²

Die Modellierung beliebiger autoregressiver $AR(p)$ - und $NAR(p)$ -Prozesse erlaubt MLPs jedoch noch keine explizite und effiziente Abbildung von $ARIMA(p,d,q)$ -Prozessen. MLPs sind explizit nicht in der Lage, die Komponenten der $MA(q)$ - Fehlerterme im Sinne kombinierter $ARMA(p,q)$ -, $ARIMA(p,d,q)$ - oder auch saisonaler $SARMA$ oder $SARIMA(p,d,q)(P,D,Q)_s$ -Modelle abzubilden [Werbos 1974; Dorffner 1996]. Hierzu sind rekurrente Netzwerkarchitekturen erforderlich, die den Prognosefehler des MLP unmittelbar als tatsächlichen Fehlerwert oder in Form des Prognosewertes für jedes anliegende Muster als Information an vorgelagerte Schichten des NN als Eingabe zurückführen und so einen $MA(q)$ -Prozess explizit abbilden. Somit repräsentiert die Verfahrensklasse der NN eine Oberklasse linearer und nichtlinearer $SARIMA$ -Modelle, nicht jedoch die überwiegend betrachteten MLPs. Dies erscheint insofern besonders relevant, als die Verfahren der Exponentiellen Glättung aufgrund ihrer hohen Prognosegenauigkeit und Fehlerrobustheit in unterschiedlichen Anwendungsbereichen in der Praxis am weitesten verbreitet sind. Verfahren der Exponentiellen Glättung stellen jedoch Sonderfälle von integrierten Moving-average-Prozessen $IMA(0,d,q)$ dar.³ Aus der Erfahrung empirischer Anwendungen erscheint es daher vielmehr plausibel, NN-Architekturen auf der Basis von $IMA(d,q)$ -Prozessen zu entwickeln, als MLPs zur Approximation von $ARI(p,d)$ zu überprüfen. Entsprechend empfiehlt ZIMMERMANN auch die Vernachlässigung vorwärtsgerichteter NN wie der MLPs und die ausschließliche Anwendung rekurrenter Netzwerkarchitekturen [Zimmermann 2006a; ebenso implizit in Zimmermann und Neuneier 2000b; Zimmermann et al. 2005; Schafer und Zimmermann 2006].

Umso erstaunlicher ist angesichts dieser fundamentalen theoretischen Nachteile die nahezu ausschließliche Verwendung von MLPs in Publikationen zur Absatzprognose sowie ihre generelle Dominanz in anderen Disziplinen.⁴ Ferner stellt ZHANG [2001] in einer empirischen Untersuchung der Prognose von linearen $ARIMA$ -Prozessen mit nichtlinearen MLPs fest, dass MLPs ebenfalls einfache $MA(1)$ -, $MA(2)$ - und kombinierte $ARMA(p,q)$ -Prozesse valide approximieren und prognostizieren können. Dies ist deshalb möglich, weil ein $MA(q)$ -Prozess

¹ Die Komplexität der Interaktion von Verbindungsgewichten, Schwellwerten und Nichtlinearitäten erlaubt jedoch eine Extraktion dieser Regeln aus trainierten MLPs, so dass diese Ausführungen nur zur verbesserten Interpretation postuliert werden.

² WIDMANN [2001] untersucht darüber hinaus die Verfahren der Projection Pursuit Regression in Verbindung zu NN anhand einer statistischen Analyse. Da diese jedoch in der Absatzprognose kaum Beachtung gefunden haben, werden sie nachfolgend vernachlässigt.

³ Siehe auch Abschnitt 3.2.2.4.3.

⁴ Die verbreitete Aussage, wonach 80% aller Anwendungen von NN durch MLPs erfolgen, geht auf eine frühe Studie durch CAUDIL [Caudil 1992] und KAASTRA und BOYDT [Kaastra und Boyd 1996] zurück. Sie wird durch die Literaturanalyse von CRÖNE und PREBMAR [2006] im Anwendungsbereich der Absatzprognose auch 15 Jahre später noch bestätigt.

durch einen endlich langen $AR(p)$ -Prozess hinreichend approximiert werden kann.¹ Dies erfordert eine Erweiterung der autoregressiven Terme p des Eingabevektors, was bei der Prognose saisonaler $IMA(d,q)$ -Prozesse durch viele Eingabeneuronen $n = p$ zu vergrößerten Netzwerkarchitekturen mit vielen freien Parametern führt. Theoretisch erfordert dies eine Erweiterung der universellen Approximationsfähigkeit von MLPs, da neben einer unendlichen Anzahl von versteckten Neuronen zur Approximation beliebiger Funktionen zur Abbildung aller zeitreihenanalytischen Prozesse auch eine unendliche Anzahl von Eingabeneuronen erforderlich sein kann. Die Modellierung von vorwärtsgerichteten Netzwerkarchitekturen der MLPs kann somit zwar nicht immer als ein effizienter, wohl aber als ein effektiver Modellierungsansatz zur Prognose allgemeiner $SARIMA(p,d,q)(P,d,Q)_s$ -Prozesse angesehen werden. Dies erklärt auch den erheblichen Erfolg von MLPs in theoretischen und empirischen Anwendungen und erlaubt eine Vernachlässigung rekurrenter Netzwerkarchitekturen in den nachfolgenden Kapiteln.

Zusammenfassend bleibt festzuhalten, dass NN und insbesondere MLPs eine Klasse flexibler Modelle zur linearen und nichtlinearen semiparametrischen Regression darstellen. MLPs erlauben aufgrund der hohen Freiheitsgrade in der Modellierung die Bildung unterschiedlicher Netzwerkarchitekturen zur Prognose, spezifiziert durch die Bestimmung des Tupels $A = [[N_I - N_{H1}(F_I, F_A, F_O) - N_{H2}(F_I, F_A, F_O) - \dots - N_O(F_I, F_A, F_O)], W, K]$. Innerhalb einer Architektur kann bereits ein einziges lineares Neuron im Trainingsprozess durch Konvergenz der Verbindungsgewichte und Schwellwerte die Funktion unterschiedlicher statistischer Verfahren annehmen, die in einer nachgelagerten Einheit situativ kombiniert werden können. Über die flexible, semi-parametrische Abbildung eines linearen oder nichtlinearen Prognosemodells allein aus den gelernten Datenmustern bieten sie auch eine hohe Flexibilität in der Verarbeitung metrisch skaliertes sowie nominal skaliertes Eingabedaten zur Abbildung nahezu beliebiger Eingabe- Ausgabebeziehungen zwischen unabhängigen und abhängigen Variablen. Sie besitzen ferner die Fähigkeit zur Mehrschrittprognose durch multiple Ausgabeneuronen und stellen somit eine wertvolle Erweiterung des zeitreihenanalytischen Instrumentariums dar.

Allerdings folgen NN ebenso wie alle linearen Methoden der Regression einem explizit in der Fehlerfunktion abgebildeten Modellierungsziel. Unter Verwendung der etablierten quadratischen Zielfunktion modelliert so auch ein nichtlineares MLP herkömmliche Kleinste-Quadrate-Schätzer zur Approximation des Mittelwertes einer normalverteilten Zufallsvariablen. Die Untersuchungen im nachfolgenden Abschnitt zeigen zunächst, dass diese Beschränkung auf quadratische Zielfunktionen nicht notwendig ist und darüber hinaus in vielen betriebswirtschaftlichen Anwendungsfällen der Prognose auch nicht den tatsächlichen Entscheidungskosten der Prognose entspricht. Entsprechend wird im nachfolgenden Abschnitt die Methodologie zur Modellierung von MLP zur Prognose um beliebige asymmetrische Kostenfunktionen erweitert und ihre Eigenschaften sowie ihre Anwendbarkeit analysiert.²

¹ Dieses ist auch als WOLDS Repräsentations-Theorem bekannt und findet in der sparsamen Modellierung rational verteilter Lags in $ARIMA(p,d,q)$ -Modellen Anwendung [siehe u.a. Diebold 2001]. Ein $MA(q)$ -Prozess kann demnach durch einen infiniten $AR(p)$ -Prozess approximiert werden und umgekehrt, jedoch ist die explizite Modellierung als $MA(q)$ -Term effizienter und sparsamer, da weniger freie Verfahrensparameter benötigt werden.

² Die nachfolgenden Ausführungen zu asymmetrischen Kosten- bzw. Zielfunktionen zum Training von NN beschränken sich nicht auf Architekturen der MLPs, sondern sind theoretisch auch auf rekurrente Architekturen der NN anwendbar. Vor dem Hintergrund der dominanten Verbreitung und Anwendung von MLPs in der Prognose wird jedoch auf eine Ausweitung auf alternative Architekturen verzichtet.

5 Neuronale Netze zur Disposition

Wie in Abschnitt 4.3.2 gezeigt wurde, minimiert ein Lernalgorithmus implizit eine spezifische Zielfunktion¹, die das Lernziel des Neuronalen Netzes definiert und unabhängig von dem Lernalgorithmus bestimmt werden kann. Die quadratische Fehlerfunktion nimmt im Training von Neuronalen Netzen eine zentrale Rolle ein, wird aber in wissenschaftlichen Untersuchungen bisher weitgehend vernachlässigt. Sie wird nachfolgend genauer analysiert, um ihre besondere Bedeutung im Lernprozess und in der Anwendung von NN hervorzuheben und die Freiheitsgrade ihrer Modellierung aufzuzeigen. Ferner wird die Nutzung weitere Zielfunktionen zum Training Neuronaler Netze in der wissenschaftlichen Literatur untersucht.

Ausgehend von diesen Erkenntnissen werden anschließend bestehende Erweiterungen von symmetrischen statistischen Zielfunktionen und asymmetrischen betriebswirtschaftlichen Kostenfunktionen untersucht. Aufbauend auf der theoretischen Fähigkeit zur Minimierung der tatsächlichen Entscheidungskosten in asymmetrischen Kostenfunktionen werden diese anschließend als Zielfunktionen im Training von MLPs auf synthetischen Zeitreihen experimentell analysiert. Die vielversprechenden Ergebnisse erlauben sodann die weiterführende Analyse und Beurteilung asymmetrischer Kostenfunktionen auf empirischen Zeitreihen im Vergleich zu etablierten symmetrischen Fehlerfunktionen und statistischen Verfahren im darauffolgenden Kapitel 6.

5.1 Lernziele und Zielfunktionen von Neuronalen Netzen

Alle Lernverfahren von NN dienen der Approximation eines datengenerierenden Prozesses aus einzelnen Mustern p von bekannten Eingabedaten x_p und bekannten Ausgabedaten t_{pj} eines Datensatzes P , $p = 1, \dots, P$. Das NN erlernt die heteroassoziative Beziehung zwischen Eingabe- und Ausgabewerten durch die Bestimmung der Verbindungsgewichte w_{ij} und der Schwellwerte θ_j für jedes Ausgabeneuron j , so dass die Distanz der tatsächlichen Netzwerkausgabe o_{pj} und der Soll-Ausgabe bzw. Lerneingabe t_{pj} über alle Datenmuster p minimiert wird.² Das Lernziel eines NN ist somit die Minimierung³ des Ausgabefehlers $E(\bullet)$ über alle Datenmuster p durch die Bestimmung fehlerminimaler Ausgabewerte o_{pj} . Dies erfolgt durch die Bestimmung der freien Parameter w_{ij} und $\theta_j = w_{0j}$ des NN in der Verbindungsmatrix W für eine gegebene Netzwerkarchitektur und in Abhängigkeit von den Eingabedaten x_p .⁴ Das Lernziel wird entsprechend durch die Zielfunktion $O(W)$ (engl. *objective function*) des Netzwerkfehlers formuliert:

$$O(W) = \sum_{p=1}^m E(t_p, o_p(w_{p,0}, w_{p,1}, \dots, w_{p,n}; x_p)) \rightarrow \text{MIN!} \quad (5.1)$$

Die Minimierung des Netzwerkfehlers $E(W)$ erfordert eine Bewertung des Abstands zwischen der Netzwerkausgabe o_{pj} und der Soll-Ausgabe t_{pj} durch die Fehlerfunktion $E(t_{pj}, o_{pj})$. Hierzu

¹ Die Zielfunktion (engl. *objective function*) wird vielfach synonym auch als Fehlerfunktion (engl. *error function*), Kostenfunktion (engl. *cost function* bzw. *cost of error function*) oder Verlustfunktion (engl. *loss function*) bezeichnet. Diese Begriffe sollten jedoch nicht synonym verwendet werden; ihre Bedeutung wird im Folgenden eingehend diskutiert und hergeleitet.

² Solange ein NN nicht fehlerfrei auf allen Mustern eines Datensatzes trainiert ist, wird die Netzwerkausgabe vom Zielwert der Ausgabe abweichen. Bei stochastischen Zeitreihen mit verrauschten Eingabemustern kann daher kein Minimum von 0 gefunden werden, sondern nur ein geringer Zielfunktionswert. Die Probleme des Overfitting durch Minimierung auf einem Trainingsdatensatz werden in Abschnitt 4.3.2 diskutiert.

³ Sinnvolle weiterführende Zielsetzungen einer schnellen oder robusten Minimierung werden vernachlässigt.

⁴ In Anlehnung an Abschnitt 4.3.2 werden die Schwellwerte θ_j als Verbindungsgewichte w_{0j} aller Neuronen der versteckten Schichten und der Ausgabeschicht zu einem Bias-Neuron u_0 modelliert. Sie können somit als reguläre Verbindungsgewichte mittrainiert werden und sind daher in der Verbindungsmatrix W enthalten.

können unterschiedliche Berechnungsvorschriften verwendet werden, die der Distanz von zwei Punkten eines Raums $d(o_{pj}, t_{pj})$ einen numerischen Wert zuordnen. Auf diese Weise kann die Bedeutung der Abweichung für die zugrundeliegende Anwendung bewertet werden.

Nach der Bestimmung der Fehlerfunktion in der Zielfunktion werden die Abweichungen der Netzwerkausgabe bewertet und mit dem Distanzmaß minimiert.¹ Die Gewichtsänderung Δw_{ij} erfolgt proportional zur partiellen Ableitung des Fehlers der Netzausgabe, welcher über die Zielfunktion Richtung und Intensität der Gewichtsänderung definiert. Somit determiniert die Fehlerfunktion die Lernsignale, aus denen das NN den funktionalen Zusammenhang der Trainingsdaten erlernt. Im Hinblick auf die übliche geometrische Interpretation des Minimierungsproblems bestimmt die Wahl der Fehlerfunktion die Form der Fehlerfläche [1988, S. 348], die anschließend mittels Gradientenabstieg auf der Fehlerfläche minimiert wird (vgl. auch Abschnitt 5.3.2.3). Der Bestimmung der Fehlerfunktion kommt daher eine besondere Bedeutung für die problemadäquate Bestimmung der Lernaufgabe zu. Die Fehlerfunktion kann je nach Problemstellung (Regression bzw. Klassifikation) sowie der Anwendungsparameter individuell bestimmt werden.

In Anwendungen der Regression können unterschiedliche Distanzmetriken zur Berechnung der Fehlerwerte verwendet werden, welche die Distanz zwischen Netzwerkausgabe o_{pj} und Soll-Ausgabe t_{pj} , ausgedrückt durch den einfachen Fehler e_{pj} (siehe Formel 3.81), bewerten,

$$e_{pj} = t_{pj} - o_{pj} \quad (5.2)$$

In der Notation der Prognosefehler entspricht t_{pj} dem tatsächlichen Beobachtungswert y_t zum Zeitpunkt t und die Ausgabe des NN o_{pj} dem Prognosewert \hat{y}_t .² Korrespondierend mit der weiten Verbreitung von Kleinste-Quadrate-Schätzern in der Statistik wird überwiegend die Fehlerfunktion der Summe der quadratischen Fehler (engl. *sum of squared errors*, *SSE*) als Zielfunktion verwendet:

$$E(e_p) = \frac{1}{2} SSE(e_p) = \frac{1}{2} \sum_j (t_{pj} - o_{pj})^2 \quad (5.3)$$

Sie unterstellt eine quadratische Distanz der Fehler eines Musters p zwischen der Netzwerkausgabe o_{pj} und der Lerneingabe t_{pj} gemäß der linearen Regression. Entsprechend können alle statistischen Fehler zur Ex-post-Beurteilung der Prognosegüte in Abschnitt 3.3 als Fehlerfunktionen verwendet werden.

Neben der Regression zur Punktprognose kann die Lernaufgabe eines NN – unter Verwendung identischer Daten – jedoch auch als Wendepunktprognose bestimmt werden.³ In der Prognose von Wendepunkten mit einer metrisch skalierten, abhängigen Variablen wird dann ein Fehler zugewiesen, wenn die Veränderung der Netzwerkausgabe Δo_{pj} nicht der Veränderung des tatsächlichen Beobachtungswertes Δt_{pj} entspricht, ungeachtet des Betrags der Vorhersage. Die Kodierung entspricht einer Klassifikation im Sinne einer „Steigt-oder-fällt“-Vorhersage der Zeitreihe durch Verwendung einer abweichenden Zielfunktion mit

¹ Diese Formulierung des Minimierungsproblems zum Auffinden optimaler Parameter eines NN entspricht der konventionellen Minimierung des quadratischen Fehlers in statistischen Regressionsverfahren.

² Die Berechnung stimmt mit den statistischen Fehlermaßen in Abschnitt 3.3 überein.

³ Dies erfolgt vielfach in Anwendungen zur Vorhersage von Aktienkursen, die sowohl als Klassifikation als auch als Regression formuliert werden können. Die Vorhersage dient dem Bilden von Kauf- und Verkaufssignalen, die anzeigen, ob ein Aktienkurs in der Folgeperiode im Vergleich zum Vortag steigt oder fällt. Ziel der Lernaufgabe der Klassifikation ist die Zuweisung eines Fehlers im Fall einer inkorrekten Vorhersage der Klassenzugehörigkeit sowie die Zuweisung keines Fehlers bei korrekter Klassifikation.

$$E(e_p) = \begin{cases} 0 & \text{für } \text{sign}(\Delta t_p) = \text{sign}(\Delta o_p) \\ 1 & \text{für } \text{sign}(\Delta t_p) \neq \text{sign}(\Delta o_p) \end{cases}, \text{ mit } o_{pj}, t_{pj} \in \{0;1\} . \quad (5.4)$$

Ferner kann eine Vorhersage aus metrischen Beobachtungswerten auch als rein binäre Klassifikation modelliert werden, unter entsprechender Reskalierung der Datenmuster p in eine nominal skalierte abhängige Variable $t_p \in \{0;1\}$. Eine korrespondierende Fehlerfunktion spezifiziert die Vorhersage steigender oder fallender Beobachtungswerte so vereinfacht als¹

$$E(e_p) = \begin{cases} 0 & \text{für } o_p = t_p \\ 1 & \text{für } o_p \neq t_p \end{cases}, \text{ mit } o_{pj}, t_{pj} \in \{0;1\} . \quad (5.5)$$

Die Bestimmung der jeweiligen Zielfunktion eines NN bietet somit eine Vielzahl von Freiheitsgraden, was eine erhebliche Flexibilität in der Bestimmung der Lernaufgabe und des Netzwerkverhaltens zur Folge hat.² Alle Fehlerfunktionen und Distanzmetriken weisen einer Abweichung der Netzwerkausgabe unterschiedliche Fehlerwerte zu, so dass diese im Lernprozess ein unterschiedliches relatives Gewicht erhalten.³ Dies unterstreicht die besondere Bedeutung der Fehlerfunktion in der Anwendung von NN.

Eine Beurteilung der Relevanz der Fehlerfunktion kann nur relativ, im Vergleich zu den anderen Parametern der Architektur und des Lernalgorithmus eines NN erfolgen, nach $MLP=[A, L, O]$. Die funktionale Form der Zielfunktion bestimmt zunächst das Lernziel des NN auf den Daten durch Modellierung einer Klassifikation oder Regression (wie in (5.3), (5.4) und (5.5)). Darüber hinaus bestimmt sie durch eine Bewertung der Distanz zwischen Ist- und Soll-Netzwerkausgabe, welche Muster ein NN aus den Daten erlernen soll. Die Parameter der Netzwerkarchitektur und des Lernalgorithmus sind anschließend so zu bestimmen, dass das Lernziel auf den Trainingsdaten valide und reliabel erlernt werden kann.⁴ Somit kann nachfolgend explizit in Lernverfahren und Lernaufgabe unterschieden werden. Während ein Lernverfahren einen Algorithmus zur Findung einer Parameterkombination repräsentiert, die ein Minimum auf einer Fehlerfläche findet, bestimmt die zugrunde liegende Lernaufgabe mittels der Zielfunktion die Form des Fehlergebirges selbst. Die Spezifikation von Netzwerkarchitektur und Lernverfahren ist somit lediglich Voraussetzung zum Erreichen eines Lernziels und als solche der Bestimmung der Zielfunktion unterzuordnen. Kann die Zielsetzung der Anwendung von NN in Form einer Zielfunktion quantifiziert werden, kann ein NN

¹ Dies kann neben der Re-Skalierung der abhängigen Variablen auch durch eine Anpassung der Netzwerkarchitektur, z. B. zur Ausgabe binärer Ausgabewerte durch ein Ausgabeneuron mit binärer Aktivierungsfunktion, unterstützt werden.

² Darüber hinaus existieren vielfältige Möglichkeiten zur Modellierung von NN für Klassifikationsaufgaben durch unterschiedliche Fehlerfunktionen und korrespondierende Netzwerkarchitekturen, je nachdem, ob eine oder mehrere Klassen durch abhängige Variablen t_{pj} mit nominalem oder binärem Skalenniveau abgebildet werden sollen. Vorhersagen der Wahrscheinlichkeit einer Klassenzugehörigkeit können durch entsprechende Lernziele und eine metrische Fehlerfunktion abgebildet werden, unter Anpassung der Netzwerkarchitektur an eine Netzwerkausgabe o_{pj} mit einem adäquaten Ausgabeintervall. Darüber hinaus existieren vielfältige Repräsentationsformen der Klassifikation, die unterschiedlichen Klassen oder Netzwerkfehlern unterschiedliche Fehlerwerte zuordnen. Da die vorliegende Untersuchung jedoch Fragestellungen der Regression untersucht, wird auf eine nähere Analyse von Zielfunktionen zur Klassifikation verzichtet und auf BISHOP [1995] verwiesen.

³ Eine quadratische Fehlerfunktion weist einer Abweichung $t_p - o_p = 1$ einen Fehler $E_p(t_p, o_p) = 1$ und einer Abweichung $t_p - o_p = 2$ einen Fehler $E_p(t_p, o_p) = 4$ zu. Dagegen weist die absolute Fehlerfunktion den Abweichungen $t_p - o_p = 1$ den Fehler $E_p(t_p, o_p) = 1$ und $t_p - o_p = 2$ den Fehler $E_p(t_p, o_p) = 2$ zu und gewichtet einzelne Abweichungen somit unterschiedlich.

⁴ Auch wenn die Bestimmung der Lernaufgabe unabhängig von der Architektur erfolgen kann, kann die Auswahl einer geeigneten Architektur die Minimierung der Zielfunktion und so das Erreichen des Lernziels unterstützen.

auch abweichende Lernaufgaben lösen. Somit sind theoretisch auch spezifische, betriebswirtschaftlich motivierte Kostenfunktionen in Bezug auf Entscheidungen zur Warendisposition unmittelbar in den Trainingsprozess von NN integrierbar.

Die Herleitungen der LMS- und der Backpropagation-Lernregel (vgl. Abschnitt 4.3.2) zeigen, dass die Bestimmung der Zielfunktion eines NN unabhängig von der Berechnungsvorschrift des Lernalgorithmus sowie von möglichen Erweiterungen der Lernverfahren erfolgt. Sie kann somit beliebig gewählt werden.¹ Darüber hinaus verdeutlicht die Analyse des Aufbaus und der Funktionsweise von NN (vgl. Abschnitt 4.2), dass alle Architekturentscheidungen zur Signalverarbeitung im Neuron sowie die Informationsverarbeitung in Netzwerken von Neuronen unabhängig von der Zielfunktion des Lernens getroffen werden können.² Die Untersuchung der methodischen Grundlagen von NN (vgl. Abschnitt 4.1) zeigte ferner, dass theoretische, anwendungsbezogene oder biologische Aspekte keine spezifische, quadratische Zielfunktion erfordert. Vielmehr ist die Verwendung des quadratischen Prognosefehlers auf die historische Entwicklung der Lernalgorithmen zurückzuführen. Die Bestimmung der Zielfunktion bzw. der Fehlerfunktion zum Training von NN kann somit losgelöst von anderen Entscheidungen hinsichtlich der Architektur und der Lernverfahren getroffen werden. Die Unabhängigkeit von Zielfunktion, Lernverfahren und Architektur gestattet daher eine problemangepasste Steuerung des Verhaltens eines NN in der Anwendung.³

Im Gegensatz zu den theoretischen Fähigkeiten von NN zur Minimierung und somit Nutzung unterschiedlicher Zielfunktionen wird seit ihrer populären Beschreibung in der Originalpublikation des Backpropagation-Algorithmus durch RUMELHART, HINTON und WILLIAMS [1986] nahezu ausschließlich die Fehlerfunktion der Summe der quadratischen Fehler *SSE* verwendet. Sie wird in der wissenschaftlichen Forschung ebenso wie in praktischen Anwendungen universell sowohl zur Regression als auch zur Klassifikation eingesetzt.⁴ Nur vereinzelte Publikationen berücksichtigen die Anwendung einer alternativen Fehlerfunktion.

Dies gilt insbesondere für die Publikationen zur Absatzprognose mit NN aus Abschnitt 4.4.1.3 [Crone und Preßmar 2006]. In allen 63 Veröffentlichungen wird ausschließlich die quadratische Fehlerfunktion des *SSE* verwendet. THIESING [1998, S. 20; 33-35] erwähnt zwar die Möglichkeit der Nutzung einer linearen Fehlerfunktion der l_1 -Norm im Backpropagation-Algorithmus zur Absatzprognose für 28 Artikel eines Supermarktes in Anlehnung an NEUNEIER und ZIMMERMANN [1998], präsentiert oder diskutiert jedoch keine entsprechenden Ergebnisse [Thiesing 1998, S. 212-214]. RAGG [2002] nutzt dagegen eine erweiterte Zielfunktion mit einem Weight-decay-Term zur Regularisierung, jedoch ebenfalls in Verbindung mit der konventionellen quadratischen Fehlerfunktion des *SSE*. Alle anderen Publikationen verwenden die quadratische Fehlerfunktion. Darüber hinaus kann sogar eine weitgehende Unkenntnis der Existenz und Bedeutung von Fehlerfunktionen für die Lernaufgabe festge-

¹ Die funktionale Form der Zielfunktion muss jedoch zur Bildung der partiellen Ableitung der Fehlerfunktion vollständig differenzierbar sein. Dies ist nachfolgend zu berücksichtigen.

² Die nachfolgenden Ausführungen zu alternativen Zielfunktionen und die Erweiterung auf allgemeine Kostenfunktionen erfolgt somit unabhängig von der Ausgestaltung der Netzwerkarchitektur von MLPs, RBF-NN, TDNN, GRNN sowie Topologien der vorwärts- oder rückgekoppelten Netze. Die Erkenntnisse können somit auch auf andere Netzwerktypen bezogen werden.

³ Zusätzlich kann die Zielfunktion um so genannte Regularisierungs-Komponenten zur Kontrolle der Lern- und Generalisierungsleistung im Trainingsprozess erweitert werden, z. B. in [Seghouane und Fleury 2001; Lappalainen 1998; Suarez-Romero et al. 2003]. Diese haben jedoch keinen Einfluss auf die Formulierung der Fehlerfunktion und können in der nachfolgenden Diskussion vernachlässigt werden.

⁴ Vielmehr scheint sich der Großteil wissenschaftlicher Veröffentlichungen auf die Entwicklung von Heuristiken zur Ermittlung „optimaler“ Lernverfahren, robuster Lernparameter und Netzwerkarchitekturen zu beschränken und sich nicht mit der Findung problemadäquater Zielfunktionen zu beschäftigen.

stellt werden. So erwähnen nur 9 der 63 Veröffentlichungen¹ explizit die Auswahl des *SSE* als Fehlerfunktion [Faraway und Chatfield 1995; Zekic-Susac et al. 2005; Kuo 2001; Kuo et al. 2002; Kuo und Xue 1998a; Kuo und Xue 1998b; Kuo und Xue 1999; Nam und Schaefer 1995; Tang und Fishwick 1993], obwohl 53 Veröffentlichungen den Backpropagation-Algorithmus anwenden und alle die Ex-post-Beurteilungskriterien der Prognosegüte dokumentieren.² Die mangelnde Transparenz hinsichtlich der Verwendung von Zielfunktionen wird unter anderem in CHANG et al. [2005] deutlich, die ein MLP mit Backpropagation und der Zielfunktion des *SSE* trainieren, gleichzeitig jedoch die Netzwerkarchitektur durch einen genetischen Algorithmus zur Maximierung der Fitness anhand des *MAPE* optimieren, den sie als eigentliches Zielkriterium ihrer Anwendung identifiziert haben. Die Möglichkeit einer Verwendung des *MAPE* als Fehlerfunktion in der Parametrisierung des NN wird dabei nicht berücksichtigt.

Eine über den Anwendungsbereich der Absatzprognose hinausgehende Literaturrecherche belegt, dass auch in anderen Anwendungsbereichen und Disziplinen der Regression fast ausschließlich die quadratische Fehlerfunktion der *SSE* verwendet wird. Von den 4694 Publikationen zur Prognose und Vorhersage aus der Literaturrecherche in Abschnitt 4.4.1.3 berücksichtigen nur drei Publikationen abweichende Fehlerfunktionen. NUNNARI und CANNAVO [2005] untersuchen die Anwendung von fünf unterschiedlichen Fehlerfunktionen zur Vorhersage der täglichen Ozonkonzentration in Italien.³ Sie ermitteln jedoch nur geringe Unterschiede in der Prognosegenauigkeit der Fehlerfunktionen und bieten keine formale statistische oder anwendungsbezogene Begründung für deren Einsatz. ENGLISCH und HIEMSTRA [1994] erweitern die quadratische Fehlerfunktion zur Abbildung eines korrelationsbasierten Lernens. FORTI et al. [1992] entwickeln mehrere robuste, nicht-quadratische Fehlerfunktionen für FIR-Filter in der Signalverarbeitung, was aufgrund der Ähnlichkeit zur Formulierung von AR(p)-Modellen als Anwendung der Regression betrachtet werden kann. REFENES et al. [1997] modifizieren die Zielfunktion durch eine Gewichtung der quadratischen Fehler durch „discounted least squares“ in der Vorhersage von Finanzmarkt-Zeitreihen. Weiter zurückliegende quadratische Fehlerbeträge erhalten dabei ein geringeres Gewicht, um veränderliche Datenstrukturen abbilden zu können. Weitere Publikationen, u. a. [Bode 1998], suggerieren irrtümlich die Nutzung alternativer Fehlerfunktionen, verwenden jedoch ebenfalls den *SSE*.

Die weitgehende Vernachlässigung alternativer, nicht-quadratischer Fehlerfunktionen zur Regression wird auch durch eine Analyse theoretischer Publikationen ohne unmittelbaren Bezug zu einer Anwendung bestätigt. BISHOP leitet unterschiedliche Fehlerfunktionen analytisch in Abhängigkeit von der Wahrscheinlichkeitsverteilung der abhängigen Variablen her [Bishop 1995, S. 195-252]. Er unterscheidet entsprechend zwischen der quadratischen und der linearen Fehlerfunktion zur Regression einerseits und spezifischen Fehlerfunktionen zur Klassifikation andererseits. HANSON und BURR [1988] untersuchen Fehlerfunktionen der Minkowski-*R*-Metriken zur Regression, ebenso BURRASCANO [1991] für ausgewählte Metriken. REED und MARKS [1999] betrachten kurz unterschiedliche Fehlerfunktionen, untersuchen jedoch ausschließlich die Eigenschaften des Trainings mit der quadratischen Fehlerfunktion. Die Erkenntnisse der wenigen Untersuchungen zur Regression werden im

¹ Mit Ausnahme vorangehender Veröffentlichungen von CRONE et al. [Crone 2002; Crone 2004; Crone 2005; Crone et al. 2006a; Crone et al. 2006b; Crone et al. 2006c; Crone et al. 2006d], die den Forschungsgegenstand alternativer Fehlerfunktionen für NN bereits untersuchen.

² Es wurde nach den vielfach synonym verwendeten Begriffen für Fehlerfunktion (engl. *error function*), Kostenfunktion (engl. *cost function*), Verlustfunktion (engl. *loss function*) und Zielfunktion (engl. *objective function*) gesucht. Veröffentlichungen in denen Autoren zwar auf die Minimierung einer Fehlerfunktion hinweisen, diese aber nicht spezifizieren, wurden somit ebenfalls berücksichtigt.

³ Die dieser Arbeit vorangegangenen Publikationen des Autors [Crone 2002; Crone 2003] werden nicht dazu gezählt. Ihr Inhalt wird im nachfolgenden Abschnitt dargestellt und die Untersuchung entsprechend erweitert.

Abschnitt 5.2.2 zusammengetragen und analysiert. Auch zu Anwendungen der Klassifikation¹ oder Optimierung² finden sich nur vereinzelte Untersuchungen. Da diese jedoch nicht den Gegenstand der vorliegenden Arbeit berühren, werden sie nicht explizit diskutiert.

Darüber hinaus existiert umfassende Literatur zur Erweiterung der Zielfunktion um Parameter zur Regularisierung des Lernverhaltens durch gewichtete Strafterme λE_R (engl. *penalty term*),

$$E(W) = \sum_{p=1}^m E(t_p, o_p(w_{p,0}, w_{p,1}, \dots, w_{p,n}; x_p)) + \lambda E_R((w_{p,0}, w_{p,1}, \dots, w_{p,n})) \rightarrow \text{MIN!}, \quad (5.6)$$

die in Abhängigkeit von den Verbindungsgewichten w_{ij} die Generalisierungsfähigkeit im Lernprozess verbessern sollen, jedoch ohne eine Modifikation der Fehlerfunktion [siehe u.a. Moody 1992; Corradi und White 1995; Girosi et al. 1995; Kwok und Yeung 1997; Andras 2002; Osowski und Siwek 2002; Suarez-Romero et al. 2003]. Diese werden nachfolgend jedoch vernachlässigt, da sie lediglich Konvergenz und Generalisierungsfähigkeit des Lernverfahrens beeinflussen, nicht aber die Lernaufgabe selbst definieren.

Die fast ausschließliche Verwendung der quadratischen Fehlerfunktion des *SSE* zur Prognose, entgegen der theoretischen Fähigkeit von NN zur Nutzung beliebiger Zielfunktionen in der Spezifikation der Lernaufgabe, erfordert eine genauere Analyse ihrer Eigenschaften und Vorzüge. BISHOP stellt fest: „There are many other possible choices of error function which can also be considered, depending on the particular application” [Bishop 1995, S. 194]. Diese sind in nachfolgenden Abschnitten auf ihre Fähigkeiten zu untersuchen.

¹ Die Betrachtung nicht-quadratischer Fehlerfunktionen ist in Anwendungen der Klassifikation ebenfalls selten, jedoch aufgrund der Notwendigkeit einer Modellierung nicht-quadratischer Zielfunktionen für Mehrklassenprobleme weiter verbreitet. DORLING et al. [2003] untersuchen eine Reihe von Fehlerfunktionen für Cross-Entropy, homoskedastisch-quadratische und heteroskedastisch-quadratische Verteilungsannahmen zur Vorhersage der Luftverschmutzung als binäre Klassifikation. Darüber hinaus existiert eine Reihe von Veröffentlichungen, die die Eigenschaften statistischer Fehlerfunktionen zur Klassifikation untersuchen [Holt 1992; Lee und Kim 2002; Kline und Berardi 2005; Babri et al. 2003; Theera-Umpon und Gader 2000; Osman und Blostein 1999; Oh und Lee 1999; Huang 1999; Alotto et al. 1999; Weber und Breitenbach 1997; Rhee und Oh 1996; Holt und Semnani 1990; Wang und Brown 2003] oder in der Kontrolltheorie als Klassifikation anwenden [Juang 2006; Juang 2000]. Zudem entstanden in der Klassifikation Ansätze zur Berücksichtigung asymmetrischer Kosten der Fehlklassifikation [Joo et al. 2003], die im Kontext des „cost sensitive learning“ zunehmende Beachtung finden.

² Anwendungen der Optimierung unterscheiden sich von denen zur Prognose und Klassifikation, da durch abweichende Netzwerktopologien keine überwachten Lernalgorithmen wie Backpropagation verwendet werden können und somit die Formulierung der Zielfunktion abweichend erfolgt. Folglich wird auf eine Betrachtung quadratischer Zielfunktionen zur Optimierung unter Nebenbedingungen [Liang und Wang 2000; Wiskott und Sejnowski 1997], linearer Zielfunktionen [Matsuda 1998] oder stückweise quadratischer Zielfunktionen [Thomae und Vandenbout 1990; Lee und Kim 2002; Park et al. 1993] verzichtet.

5.2 Zielfunktionen der statistischen Prognosefehler

5.2.1 Quadratische Fehlerfunktionen

Die weite Verbreitung der quadratischen Fehlerfunktion geht über den Backpropagation-Algorithmus auf ihre Verwendung im LMS-Lernalgorithmus durch WIDROW und HOFF [1960] zurück, die als erste fehlerbasierte lokale Suchheuristik des Gradientenabstiegs in Abschnitt 4.3.2 vorgestellt und hergeleitet wurde. Die sogenannte Delta-Lernregel bestimmt den Netzwerkfehler e_p aus der Ist-Ausgabe o_p und der Soll-Ausgabe t_p und bewertet jeden Fehler mit der Fehlerfunktion des quadratischen Fehlers $SE(e_p)$, durch

$$E_p(e_p) = SE_p(e_p) = (t_p - o_p)^2 . \quad (5.7)$$

Der $SE(e)$ bildet die bekannte Summe der quadratischen Abweichungen. Die Verwendung der quadratischen Fehlerfunktion ist sowohl mathematisch durch die Differenzierbarkeit zur Berechnung lokaler Gradienten begründet, als auch durch den Anwendungsbezug. WIDROW und HOFF [1960] entwickelten den Lernalgorithmus zur Rauschunterdrückung in Signalleitungen, in denen der Einfluss höherer Störungen in der Signalübertragung überproportional bestraft werden sollte.¹ WIDROW wählte die quadratische Fehlerfunktion wie auch WIENER nach dem Ziel der Rauschunterdrückung, und legitimierte die quadratische Fehlerfunktion so aus dem Anwendungsbezug und nicht aus einer formalen Herleitung unter der Annahme einer spezifischen Verteilungsfunktion [Widrow 2003].² Statistische Verfahren der Regression verwenden üblicherweise ebenfalls quadratische Zielfunktionen zur Modellierung von Kleinste-Quadrate-Schätzern. Die Anwendung von NN als adaptive Filter ist eng mit Problemstellungen der Regression verwandt³, so dass eine Anwendung der quadratischen Fehlerfunktion im Sinne von Kleinste-Quadrate-Schätzern zur Regression aus mehreren Gründen sinnvoll erscheint.

WERBOS [1974] und später RUMELHART et al. [1986] erweiterten das Gradientenverfahren der Delta-Lernregel zum Backpropagation-Algorithmus für mehrlagige MLPs mit semilinearen Aktivierungsfunktionen (vgl. Abschnitt 4.3.2).⁴ In Anlehnung an die Berechnungsvorschrift nach WIDROW und HOFF wählten sie in der mathematischen Herleitung des Algorithmus ebenfalls eine quadratische Fehlerfunktion, unter Summierung der Fehler je Datenmuster p über alle Ausgabeneuronen j . Die Fehlerfunktion der Summe der quadratischen Fehler SSE gewichtet die Differenz zwischen Lerneingabe t_{pj} und Netzwerkausgabe o_{pj} für jedes Datenmuster p als quadratische Distanz SSE , skaliert um den Faktor $1/2$:

$$E(e_p) = \frac{1}{2} SSE(e_p) = \frac{1}{2} \sum_j (t_{pj} - o_{pj})^2 . \quad (5.8)$$

¹ Die Entwicklung der Lernregel für ADALINES von WIDROW und HOFF [1960] stellt eine Erweiterung der von WIENER entwickelten Parametrisierung eines adaptiven WIENER-Filters dar.

² Die ursprüngliche Entwicklung der Zielfunktion folgte somit einer spezifischen Anwendung des so genannten adaptiven Filters. Dies besitzt für die nachfolgende Betrachtung der quadratischen Fehlerfunktion im Backpropagation-Algorithmus und die Entwicklung von anwendungsspezifischen Kostenfunktionen in der Warendisposition besondere Relevanz.

³ Die mathematische Gleichung von adaptiven WIENER-Filtern oder adaptiven Finite-Impulse-Response-(FIR-)Filtern entspricht der Formulierung autoregressiver AR(p)-Prozesse, die von Infinite-Impulse-Response-(IIR-)Filtern der von ARMA(p,q)-Prozessen. Die Architektur eines Filterelementes entspricht somit der Struktur eines vorwärtgerichteten MLP ohne versteckte Schichten mit einem einzelnen Neuron und einer beliebigen Anzahl von Eingabeneuronen.

⁴ Ebenso wie der Backpropagation-Algorithmus verwenden auch seine Erweiterungen sowie Lernverfahren höherer Ordnung die Zielfunktion des $SSE(e)$; die nachfolgenden Ausführungen gelten somit auch für sie.

Der zusätzliche Term $\frac{1}{2}$ der *SSE*-Fehlerfunktion wurde zur mathematisch-algorithmischen Vereinfachung eingeführt, um den Faktor 2 nach der partiellen Ableitung der Fehlerfunktion zu unterdrücken [Rumelhart et al. 1994]. Dazu ist anzumerken, dass diese symmetrische Skalierung durch die monotone Transformation der Fehlerfunktion auch die Steigung der Fehlerfläche systematisch verändert und somit zu Veränderungen im Lernprozess führt.¹ Für Netzwerkarchitekturen mit nur einem Ausgabeneuron u_j entspricht der *SSE*(e_p) dem quadratischen Fehler *SE*(e_p) und kann somit nachfolgend synonym zu diesem verwendet werden. Die Fehlerfunktion des quadratischen Fehlers *SE*, des um $\frac{1}{2}$ skalierten *SE* und ihrer Differenz² ist in Abbildung 5.1 dargestellt.

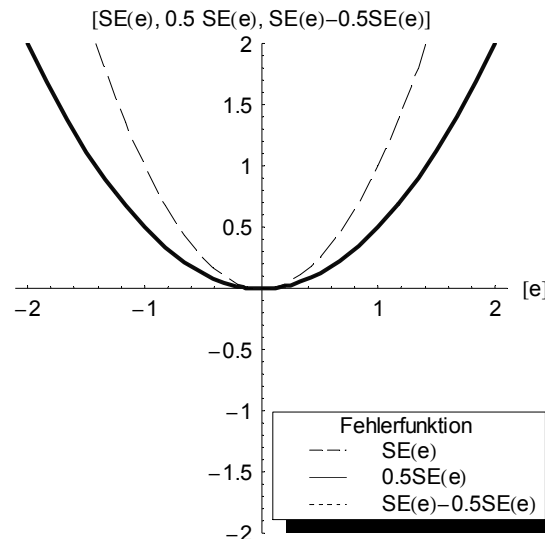


Abbildung 5.1: Quadratische Fehlerfunktionen des $SE(e)$, des skalierten quadratischen Fehlers $0,5 SE(e)$ und der Differenz $SE(e) - 0,5 SE(e)$

Abbildung 5.1 verdeutlicht, dass der quadratische Fehler *SE* höheren Abweichungen mit $e > 1$ überproportional mehr Gewicht zuweist, während geringe Fehler $e < 1$ ein unterproportionales Gewicht erhalten.³ Dies bedingt eine überproportionale Gewichtung höherer Abweichungen im Vergleich zu geringen Abweichungen und impliziert aufgrund der asymmetrischen Bewertung der Abweichungen eine implizite Kostenfunktion der Fehlerabweichungen: Größere Fehlerbeträge sind überproportional nachteiliger als geringere Fehlerbeträge. Jedoch wird nicht nur die Sensitivität für Fehlerbeträge $e > 1$ erhöht, sondern auch für $e < 1$ verringert, was insbesondere im Fall skalierten Eingabe- und Ausgabewerte der NN im Intervall $[0, 1]$ bzw. $[-1, 1]$ für Fehler $0 < e < \infty$ zu asymmetrischen Verzerrungen führt.

Die Verwendung einer quadratischen Zielfunktion wird von RUMELHART et al. nicht formal hergeleitet oder explizit begründet [Rumelhart et al. 1994; Rumelhart et al. 1986]. Sie ist vielmehr auf die positiven mathematischen Eigenschaften der partiellen Ableitungen bei semilinearen Aktivierungsfunktionen sowie die Verwendung des *SE* im Vorläufer des LMS-Algorithmus zurückzuführen. So bezeichneten RUMELHART et al. den Algorithmus auch als erweiterte Delta-Regel, in Anlehnung an seine Ähnlichkeit mit der Berechnungsvorschrift nach WIDROW und HOFF [Rumelhart et al. 1994].

¹ Diese Modifikation der quadratischen Fehlerfunktion durch RUMELHART et al. [1986] im ursprünglichen Lernalgorithmus, begründet aus der mathematischen Vereinfachung der Berechnungsvorschrift, verdeutlicht bereits die Möglichkeit der Bildung alternativer, angepasster Zielfunktionen für unterschiedliche Lernziele.

² Die Differenz $SE(e) - 0.5SE(e)$ entspricht dem um 0.5 skalierten *SE* und überlagert diese im Diagramm.

³ So führt eine Erhöhung des Fehlers e um den Betrag 1, $e+1$, zu einem überproportionalen Anstieg der Fehlerkosten von $(e+1)^2 - e^2$. Dies wird auch an dem Anstieg des Fehlerbetrages von $e=(0, 1, 2, 3, \dots)$ gegenüber den Fehlerkosten des $SE(e)=(0, 1, 4, 9, \dots)$ deutlich.

BISHOP leitet die Anwendung der quadratischen Zielfunktion des *SSE* im Training von NN zur Regression dagegen formal aus dem Maximum-Likelihood-Prinzip her [Bishop 1995, S. 195 ff.].¹ Eine Menge von Trainingsmustern $\{x_p, t_p\}$ besteht aus einem Vektor unabhängiger Eingabedaten x_p und einer zugehörigen Netzwerkausgabe der abhängigen Variablen t_p . Das NN minimiert durch iterative Anpassung der Verbindungsgewichte w_{ij} die quadratische Abweichung zwischen Netzwerkausgabe und tatsächlichen Beobachtungswerten über alle Datenmuster, analog zur Methode kleinster Fehlerquadrate der linearen Regression. Folgt der zugrundeliegende datengenerierende Prozess in den Trainingsmustern einer linearen oder nichtlinearen Funktion mit einem normalverteilten, unkorreliertem, homoskedastischen, additiven Fehlerterm mit einem Mittelwert von 0, so approximiert ein NN durch Training mit der *SSE*-Zielfunktion den bedingten Mittelwert² der Ausgabewerte im Sinne eines nichtlinearen Kleinste-Quadrate-Schätzers [Bishop 1995, S. 197].³ Die Lernaufgabe und Funktionsweise als Kleinste-Quadrate-Schätzer wird dabei nicht durch die Auswahl der Architekturparameter oder Eingabeinformationen beeinflusst. Im Gegensatz zur linearen Regression erlauben NN jedoch auch die Abbildung nichtlinearer funktionaler Abhängigkeiten, so dass ihr Anwendungsbereich über den linearer Prozesse mit normalverteilten Residuen auf die Abbildung nichtlinearer Prozesse mit normalverteilten Fehlertermen hinausgeht. „Neural networks provide a practical framework for approximating arbitrary non-linear multivariate mappings, and can therefore in principle approximate the conditional average to arbitrary accuracy“ [Bishop 1995, S. 203]. Die Anwendung der Zielfunktion des *SSE* definiert somit die Lernaufgabe der Punktprognose zur Bestimmung des Mittelwertes der abhängigen Variablen, jedoch nicht für beliebige Verteilungsannahmen.⁴

Während die Verwendung der quadratischen Fehlerfunktion im LMS-Algorithmus noch durch den Anwendungsbezug der Regression und die Annahme von quadratisch ansteigenden Fehlerkosten begründet war, ist die Anwendung des Backpropagation-Algorithmus nicht auf Anwendungen des adaptiven Filterns zur Rauschunterdrückung beschränkt. Bei unterschiedlichen Anwendungen kann die quadratische Fehlerfunktion somit nicht länger durch den Anwendungsbezug begründet sein. Zudem wurde der Backpropagation-Algorithmus durch zum Training von MLPs in Klassifikationsaufgaben entwickelt [Rumelhart et al. 1986], was zudem von dem Lernziel der Regression durch den LMS-Algorithmus oder die Herleitung durch WERBOS [1974] abweicht. Dies erfordert eine Analyse der Legitimität der quadratischen Zielfunktion in beliebigen Anwendungen der Regression und Klassifikation. Nachfolgend werden daher kurz die Eigenschaften des *SSE* in Anwendungen zur Klassifikation analysiert.

Die Fehlerzuweisung des *SE* in Anwendungen der Klassifikation kann je nach dem Skalenniveau von Netzwerkausgabe o_{pj} und Soll-Ausgabe t_{pj} unterschieden werden. Die Fehlerberechnung für nominal skalierte bzw. binäre Ist- und Sollausgaben $t_{pj}, o_{pj} \in \{-1, 1\}$ ist in Tabelle 5.1 dargestellt, die eine Übersicht der Zustände der Ist- und Soll-Ausgabewerte,

¹ Dies steht im Gegensatz zur ursprünglichen Herleitung des Backpropagation Algorithmus durch WERBOS [1974] oder RUMELHART et al. [1986], erlaubt jedoch die Untersuchung der Eigenschaften von NN mit quadratischen Zielfunktionen als Schätzer der Regression.

² Nachfolgend wird vereinfachend von dem Fall einer univariaten Regression ausgegangen, so dass keine weiteren exogenen Variablen zur Bestimmung der Prognose geschätzt werden müssen. Die nachfolgenden Ausführungen, insbesondere die Herleitungen nach BISHOP [1995], gelten ebenso für die Schätzung bedingter Verteilungsfunktionen und bedingter Erwartungswerte.

³ Es ist dabei anzumerken, dass die quadratische Fehlerfunktion keine Normalverteilung der Ausgabedaten erfordert. Die Anwendung einer quadratischen Fehlerfunktion geht implizit von einer Normalverteilung aus [Bishop 1995, S. 197]. Das Resultat systematisch verzerrter Ergebnisse ist anschließend durch eine Analyse der Residuen zu identifizieren.

⁴ Ferner ist anzumerken, dass die Herleitung für den Fall einer Einschnitt-Prognose Gültigkeit besitzt, nicht jedoch gleichzeitig für Mehrschritt-Prognosen.

ihrer Fehler e_{pj} sowie der quadratischen Fehler $SE(e_p)$ in einer Tabelle der Ein- und Ausgabewerte sowie einer Konfusionsmatrix der Klassifikationsfehler liefert.¹

Tabelle 5.1: Ausgabewerte und Konfusionsmatrix des $SE(e)$ für nominal skalierte Ausgaben

Eingabe- und Ausgabewerte				Konfusionsmatrix		
Ist-Ausgabe	Soll-Ausgabe	Fehler	Fehler	Soll-Ausgabe	Ist-Ausgabe	
o_{pj}	t_{pj}	e_{pj}	$SE(e_p)$		Klasse 0	Klasse 1
-1	-1	0	0	Klasse 0	0	4
-1	+1	+2	+4		4	0
+1	+1	0	0	Klasse 1	4	0
+1	-1	-2	+4			

Die durch die Fehlerfunktion bewerteten Klassifikationsfehler bilden sinnvolle Fehlersignale für nominale Ist- und Soll-Ausgaben, die ein Training von MLPs für Netzwerkarchitekturen mit binären Ausgabe- bzw. Aktivierungsfunktionen im Intervall $[-1, 1]$ ermöglichen. Korrespondierend kann die Zuweisung von Fehlerwerten für Netzwerkarchitekturen mit nominal skalierten Ist- und Sollausgaben $t_{pj}, o_{pj} \in \{0, 1\}$ durch binäre Aktivierungsfunktionen im Intervall $[0, 1]$ betrachtet werden, siehe Tabelle 5.2.²

Tabelle 5.2: Ausgabewerte und Konfusionsmatrix des $SE(e)$ für binäre Ausgaben

Eingabe- und Ausgabewerte				Konfusionsmatrix		
Ist-Ausgabe	Soll-Ausgabe	Fehler	Fehler	Soll-Ausgabe	Ist-Ausgabe	
o_{pj}	t_{pj}	e_{pj}	$SE(e_p)$		Klasse 0	Klasse 1
0	0	0	0	Klasse 0	0	1
0	1	1	1		1	0
1	1	0	0	Klasse 1	1	0
1	0	1	1			

Auch hier zeigt sich, dass die quadratische Fehlerfunktion $SE(e)$ zur Bildung sinnvoller Fehlersignale als Lerneingaben in ein NN führt.

In den mehrlagigen Netzwerkarchitekturen der MLPs werden im Gegensatz zu binären Aktivierungsfunktionen überwiegend semi-lineare Aktivierungsfunktionen wie die logistische Funktion im Intervall $[0, 1]$ oder der TanH im Intervall $[-1, 1]$ (vgl. Abschnitt 4.2.3.2) zur Vorhersage der Klassenzugehörigkeit oder der Wahrscheinlichkeit der Klassenzugehörigkeit eines Musters p verwendet und so eine metrisch skalierte Ist-Ausgabe $o_{pj} \in [0, 1]$ einer nominalen Soll-Ausgabe $t_{pj} \in \{0, 1\}$ gegenübergestellt, wie in Tabelle 5.3 dargestellt.

Tabelle 5.3: Ausgabewerte und Konfusionsmatrix des $SE(e)$ für metrische Ausgaben

Eingabe- und Ausgabewerte				Konfusionsmatrix		
Ist-Ausgabe	Soll-Ausgabe	Fehler	Fehler	Soll-Ausgabe	Ist-Ausgabe	
o_{pj}	t_{pj}	e_{pj}	$SE(e_p)$		Klasse 0	Klasse 1
$o_{pj} = 0$	0	0	0	Klasse 0	0	$(-o_{pj})^2$
$o_{pj} \neq 1$	1	$1 - o_{pj}$	$(1 - o_{pj})^2$		$(1 - o_{pj})^2$	0
$o_{pj} = 1$	1	0	0	Klasse 1	$(1 - o_{pj})^2$	0
$o_{pj} \neq 0$	0	$0 - o_{pj}$	$(-o_{pj})^2$			

Aufgrund der Sättigungsbereiche der nichtlinearen Aktivierungsfunktionen kann die Netzwerkausgabe o_{pj} die Werte 0 oder 1 nur approximativ erreichen. Daher wird die

¹ Die Ausgabewerte entstehen beispielsweise durch binäre Schwellwertlogiken als Aktivierungsfunktionen in historischen NN des ADALINE. Die Quadrierung durch die Fehlerfunktion des SSE führt so zu einer Aufhebung positiver und negativer Fehlerwerte in der Summenbildung über mehrere Muster p . Die Darstellung der Ausgabewerte für eine Klassifikation kann im Gegensatz zur Regression nicht übersichtlich in Grafen der Fehlerfunktionen erfolgen. Die Konfusionsmatrix ist hingegen eine übliche Darstellungsform.

² So kann durch die Anwendung einer quadratischen Fehlerfunktion auf binäre Ausgabewerte eine ebensolche Fehlerkodierung erzielt werden wie durch die Fehlerfunktion in Formel (5.5), die jedoch nicht vollständig differenzierbar ist und somit keine Lösung durch ein einfaches Gradientenverfahren zulässt.

Netzwerkausgabe nach dem Training durch einen anwendungsspezifischen Schwellwert τ zur Trennung der Klassen kodiert, und zwar durch Zuweisung der Klassenzugehörigkeit 0 für alle Muster p mit $o_{pj} < \tau$ und der Klassenzugehörigkeit 1 für alle Muster p mit $o_{pj} \geq \tau$.¹ Auch hier erlaubt die quadratische Fehlerfunktion des SSE die Berechnung zulässiger Fehlersignale.

Die Gegenüberstellung der unterschiedlichen Netzwerkarchitekturen verdeutlicht die Einsatzfähigkeit der quadratischen Zielfunktion in unterschiedlichen Klassifikationsmodellen. BISHOP [1995] führt dagegen an, dass die quadratische Fehlerfunktion des SSE aus dem Maximum-Likelihood-Prinzip unter der Annahme normalverteilter Sollausgaben abgeleitet sei, was nicht mit der binären Kodierung und der entsprechenden Verteilung der abhängigen Variablen in Klassifikationsproblemen übereinstimme. „In fact, the sum of squares function is not the most appropriate for classification problems“ [Bishop 1995]. Er zeigt dementsprechend passende Fehlerfunktionen für die Klassifikation im Sinne der Diskriminanzanalyse sowie der Bestimmung von A-posteriori-Wahrscheinlichkeiten der Klassenzugehörigkeit für Ein- oder Mehrklassen-Probleme und unterschiedliche Aktivierungsfunktionen [Bishop 1995, S. 222-247]. Da die ursprüngliche Entwicklung des Backpropagation-Algorithmus sowie die Mehrzahl seiner Anwendungen aus dem Bereich der Klassifikation stammen, stellt dies die unkritische, universelle Anwendung der quadratischen Zielfunktion des SSE in Frage.

Dagegen ist die Fehlerfunktion des SSE theoretisch für Anwendungen der linearen und der nichtlinearen Regression geeignet, was ja den eigentlichen Untersuchungsgegenstand dieser Arbeit zur Bedarfsvorhersage darstellt. Dabei sind jedoch die Prämissen der unkorrelierten, homoskedastischen Normalverteilung des Fehlerterms einer stationären Zeitreihe von Beobachtungswerten kritisch zu hinterfragen. Diese mögen in der ursprünglichen Anwendung des LMS zur Rauschunterdrückung in Signalleitungen gegeben sein, sind jedoch in der Prognose von empirischen Zeitreihen auf Absatzmärkten nur selten gegeben. Vielmehr dominieren dort instationäre Zeitreihen mit Niveauverschiebungen, Ausreißern und lokalen zeitlichen Interdependenzen und paarweiser Autokorrelation.

Eine der bekannten Eigenschaften der SSE-Fehlerfunktion ist die Überbewertung großer Fehlerwerte, so dass beim Auftreten von Ausreißern die Schätzung des Mittelwertes vom Rand der Wahrscheinlichkeitsverteilung dominiert werden kann [Bishop 1995, S. 209].² Ein NN wird durch die Fehlerfunktion vor allem auf jene Aspekte der Zeitreihe trainiert, die das höchste Fehlersignal beisteuern. Durch Verwendung einer quadratischen Zielfunktion auf empirischen Absatzzeitreihen wird das NN somit nicht auf die Abbildung des zugrunde liegenden, Daten generierenden Prozesses, sondern auf zufällige Extremwerte der Ausreißer trainiert, was der Zielsetzung der Modellbildung entgegensteht. Ähnliche Probleme entstehen bei fehlenden Beobachtungswerten und fehlerhaft erhobenen Daten. Eine statistische Betrachtungsweise auf der Grundlage von Verteilungsannahmen legt somit vielmehr die Verwendung robuster Fehlerfunktionen im Training von NN nahe.

Unzulänglichkeiten in der Verwendung quadratischer Fehlerfunktionen in der Prognose von Zeitreihen wurden vor dem Hintergrund der Bestimmung adäquater Gütemaße für die M-Competitions intensiv diskutiert [Makridakis 1993; Makridakis et al. 1993; Makridakis und Hibon 2000; Fildes 1992; Fildes et al. 1998; Fildes und Makridakis 1988; Armstrong J. S. und Collopy F. 1993; Armstrong 2001; Armstrong und Collopy 1992; Armstrong und Fildes

¹ Dabei wird üblicherweise $\tau=0.5$ gesetzt. Theoretisch sind auch Fehlerfunktionen unter unmittelbarer Abbildung der Klassengrenzen τ im Trainingsprozess denkbar. Diese weichen jedoch von der Modellierung quadratischer Zielfunktionen ab. Durch die Wahl von τ können für den Fall binärer Klassifikation darüber hinaus asymmetrische Fehlerkosten der falschen positiven Fehler c_{fp} und der falschen negativen Fehler c_{fn} berücksichtigt werden, mit $\tau = c_{fp}/(c_{fn}+c_{fp})$, wobei für symmetrische Kosten $c_{fp}=c_{fn}$ entsprechend $\tau=0,5$ eine symmetrische Klassifikation erzielt.

² Dies entspricht der Bedeutung von Ausreißern in der linearen Regression mit Kleinste-Quadrate-Schätzern, die hinreichend in Standardwerken diskutiert ist.

1995; Armstrong et al. 1998]. Demnach sind quadratische Fehlerfunktionen in der Beurteilung der Prognosegüte auf empirischen Zeitreihen aufgrund ihrer restriktiven Prämissen zu vermeiden. BISHOP führt an: „It may prove useful to assess the performance of networks using a different error function from that used to train them“ [Bishop 1995, S. 197]. Dies wird auch in den Anwendungen von NN bei der Absatzprognose bestätigt, in denen die Ex-post-Prognosegüte konkurrierender Verfahren überwiegend nicht anhand quadratischer Fehlermaße bestimmt wird, obwohl das MLP auf einer quadratischen Fehlerfunktion trainiert wurde. Dabei entsteht eine Diskrepanz zwischen Lernaufgabe und nachträglicher Beurteilung des erlernten Systemzusammenhangs. Die Notwendigkeit der Übereinstimmung von Ex-post-Gütekriterium und Ziel- bzw. Fehlerfunktion zur Parametrisierung einer Methode wurde ebenfalls diskutiert [Fildes und Makridakis 1988; Zellner 1986b]. Eine Beurteilung der Prognosegüte, z. B. anhand des *MAE*, impliziert somit die korrespondierende Verwendung einer linearen Zielfunktion der absoluten Fehler (*AE*) unter der Annahme einer spezifischen Wahrscheinlichkeitsverteilung des Fehlerterms.

Eine Vielzahl von Wissenschaftlern stellt dementsprechend fest, dass quadratische Zielfunktionen in der Ex-post-Beurteilung sowie der Parametrisierung von Verfahren weder notwendig noch für alle Anwendungen sinnvoll sind. Nach REED und MARKS [1999] ist die Anwendung quadratischer Fehlerfunktionen des *SSE* oder *RMSE* für NN vor allem in der einfachen mathematischen Handhabung zur Berechnung der Ableitungen der Fehler und ihrer Interpretation im Kontext linearer Regressionsverfahren begründet [Reed und Marks 1999]. ANDERSON [1995, S. ix] stellt fest, dass die Nutzung nicht-quadratischer Zielfunktionen theoretisch zwar nicht aus algorithmischen oder softwaretechnischen Gründen verhindert, aber dennoch nicht wahrgenommen wird. BISHOP fasst in Übereinstimmung mit BURRASCANO [1991, S. 125] zusammen: „The use of the squared error function is motivated primarily by analytical simplicity“ [Bishop 1995, S. 194]. ANDERSON [Anderson 1995] stellt zudem fest, dass die Nutzung nicht-quadratischer Zielfunktionen ebenfalls nicht durch softwaretechnische Gründe verhindert wird. „A wide range of elementary functions can be inserted into a working simulation program with little effort and their effects on the resulting system investigated. ... there is no reason to limit a network to minimizing mean square error, for example, when it is just as easy to let the network minimize absolute value of error, error to the fourth power, or anything else that might be desired. ... The computer program does not care. ... Computing mean square error requires roughly the same number of machine operations as computing absolute value of error“ [Anderson 1995, S. ix].¹

Die weitgehende Dominanz quadratischer Zielfunktionen ist somit vielmehr auf andere Gründe zurückzuführen. Eine plausible Erklärung dafür ist ihre Einbettung in die ursprüngliche Herleitung des weit verbreiteten Backpropagation-Algorithmus. Eine Untersuchung etablierter Standardsoftware zur Simulation von NN belegt die mangelnde Verfügbarkeit alternativer Fehler- und Zielfunktionen, was die nahezu ausschließliche Verwendung von *SSE* in Anwendungen der Praxis und der akademischen Forschung begründet.² Auch Weiterentwicklungen des Backpropagation-Lernalgorithmus bedienen sich der quadratischen

¹ Allerdings bezieht ANDERSON [Anderson 1995] seine Ausführungen auf symmetrische Kostenfunktionen und verzichtet wie auch andere Autoren auf eine Darstellung, Entwicklung oder Anwendung alternativer Zielfunktionen.

² Die Untersuchung zeigt, dass lediglich das Programm NeuralWorks Professional IIPlus symmetrische Fehlerfunktionen dritter und vierter L_3 - und L_4 -Norm zur Regression zulässt, nicht aber der L_1 -Norm, auch wenn durch Zusatz-Programmierung theoretisch beliebige lineare Fehlerfunktionen sowie asymmetrische Verlust- und Kostenfunktionen realisierbar sind. In Anwendungen der Klassifikation erlauben hingegen die Programme NeuralWorks Professional, Alyuda NeuroIntelligence, SAS Enterprise Miner und SPSS Clementine die Verwendung beliebig definierbarer asymmetrischer Klassifikationskosten als erweiterte Fehlerfunktionen. Eine Liste der untersuchten Software-Simulatoren findet sich auf www.neural-forecasting.com.

Fehlerfunktion, in der etablierten Tradition der LMS-Regel. Aus der vorangehenden Analyse wird deutlich, dass die quadratische Fehlerfunktion des SSE sowohl für Aufgaben der Regression als auch der Klassifikation, mit nahezu beliebiger Skalierung und Kodierung der Netzwerkausgabe o_{pj} sowie der Beobachtungswerte t_{pj} und hohen Freiheitsgraden in der Netzwerkarchitektur, die Bildung sinnvoller Fehlerwerte zur Steuerung des Lernprozesses erlaubt. Diese universelle, unveränderte Anwendbarkeit der quadratischen Zielfunktion im Backpropagation-Algorithmus begründet ihre weitgehend unkritische Verbreitung sowohl für Lernaufgaben der Klassifikation als auch der Regression, auch wenn diese nicht immer verlässliche Ergebnisse liefern oder einer statistischen Überprüfung standhalten.

Trotz der klaren Dominanz der quadratischen Zielfunktion durch ihre Integration im Backpropagation-Lernverfahren und diversen Softwaresimulatoren wurde eine Auswahl alternativer Zielfunktionen entwickelt. Auch wenn diese im Bereich der Prognose und insbesondere der Absatzprognose keine Anwendung gefunden haben, werden sie im nachfolgenden Abschnitt 5.2.2 vorgestellt, um Eigenschaften alternativer Fehlerfunktionen zu verdeutlichen und weiterführende Zielfunktionen auf der Basis betriebswirtschaftlicher Entscheidungskosten zu entwickeln. Darüber hinaus erscheint auch eine Begründung der quadratischen Kostenfunktion in betrieblichen Absatzprognosen als Approximation der tatsächlich entstehenden Fehlerkosten der Problemstellung wie bei WIDROW fraglich. Die Verwendung von Fehlerkosten der Entscheidungssituation wird somit im Abschnitt 5.3 untersucht.

5.2.2 Nicht-quadratische Fehlerfunktionen

Als Lösungsansatz für die Probleme quadratischer Fehlerfunktionen bei der Regression postuliert BISHOP [1995] die Ableitung alternativer Fehlerfunktionen aus einer Generalisierung der Normalverteilung, die als Minkowski- R -Metriken oder l_R -Normen¹ bezeichnet werden. Die Metriken bewerten die Distanz zwischen Netzwerkausgabe o_{pj} und Soll-Ausgabe t_{pj} anhand unterschiedlicher Fehlerfunktionen und bestimmen so die Höhe des Lernsignals des NN:

$$E(e_{pj}) = \left(\sum_p \sum_j |t_{pj} - o_{pj}|^R \right)^{\frac{1}{R}}. \quad (5.9)$$

Die Minkowski- R -Metrik beschreibt für $R = 1$ die so genannte City-Block- oder Manhattan-Metrik, für $R = 2$ die Euklidische Distanz und entsprechend für $R > 2$ höherwertige Metriken, mit $1 < R < \infty$. Die Minkowski-Metrik mit $R \rightarrow \infty$ wird auch als Supremum- oder Maximum-Metrik bezeichnet, welche die maximale Distanz zweier Koordinaten angibt. Für Anwendungen der Regression erfolgt vielfach eine Umformung der Metriken zu Minkowski-Fehlerfunktionen, unter Vernachlässigung der Wurzel [Bishop 1995, S. 209; Hanson und Burr 1988, S. 350] durch:

$$E(e_{pj}) = \sum_p \sum_j |t_{pj} - o_{pj}|^R. \quad (5.10)$$

Der funktionale Verlauf von ausgewählten Minkowski-Fehlerfunktionen ist in Abbildung 5.2 für den absoluten Fehler $AE(e)$ mit $R = 1$, den quadratischen Fehler $SE(e)$ mit $R = 2$ und einen Fehler höherer Ordnung mit $R = 4$ im eindimensionalen Fall (a.) sowie anhand der Distanzen im zweidimensionalen Fall (b.) dargestellt.

¹ Sie wird auch als l_p -Norm bezeichnet, mit $p=R$.

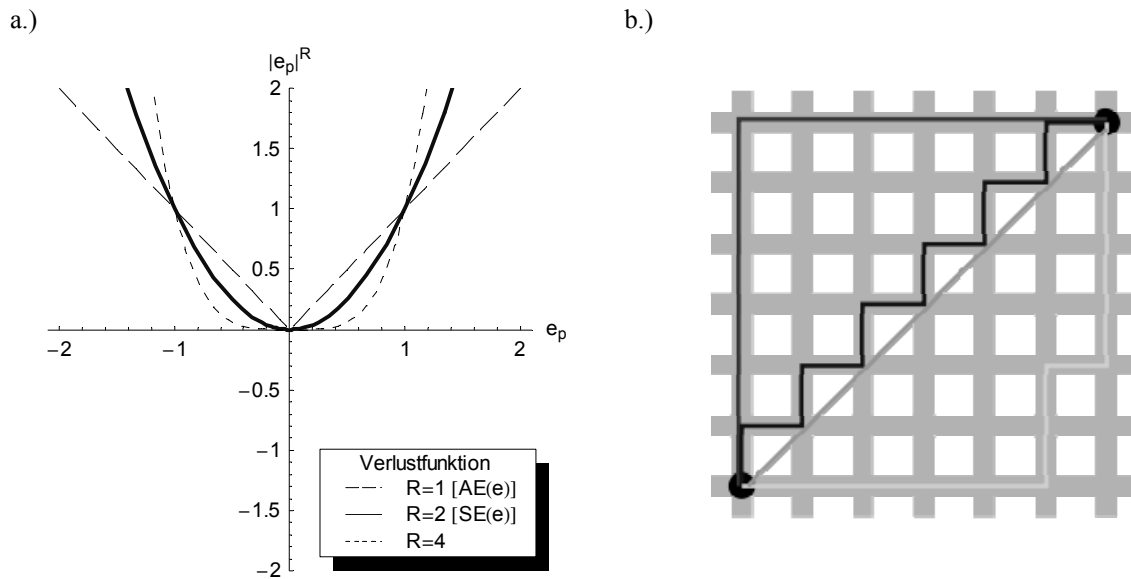


Abbildung 5.2: Minkowski- R -Fehlerfunktionen für $R = 1$, $R = 2$ und $R = 4$ im eindimensionalen Fall (a.) sowie als Distanzen im zweidimensionalen Fall (b.)

BURRASCANO [1991] beweist formal die Abhängigkeit zwischen unterschiedlichen Formen einer generalisierten Wahrscheinlichkeitsverteilung und der l_r -Metrik der Fehlerfunktion für $R = 1, 2, 5, 10$ und ∞ . Danach sollte die Metrik der Fehlerfunktion in Abhängigkeit von der Wahrscheinlichkeitsverteilung der abhängigen Variablen gewählt werden. Für $R = 2$ entspricht die Fehlerfunktion dem quadratischen Fehler $SE(e)$, welcher proportional zum Quadrat der Euklidischen Distanz der Minkowski- R -Metrik bzw. der l_2 -Norm verläuft und eine valide Metrik für Normalverteilte Fehler bietet.¹ Für nicht exakt normalverteilte Daten liefert die quadratische Fehlerfunktion der $R = 2$ -Metrik jedoch keine Maximum-Likelihood-Schätzer [Mosteller und Tukey 1980; Hanson und Burr 1988].

Für $R = 1$ entspricht die Minkowski- R -Metrik dem absoluten Fehler (engl. *absolute error*, vgl. Abschnitt 3.3) $AE(e)$. So korrespondiert die Fehlerfunktion des $AE(e)$ mit einem Maximum-Likelihood-Schätzer für eine Laplace-Wahrscheinlichkeitsverteilung der Ausgabewerte [Mosteller und Tukey 1980].² Sie wird auch als l_1 -Norm des linearen Fehlers (engl. *linear error*) bezeichnet und ist identisch mit der City-Block- oder Manhattan-Metrik der Minkowski- R -Metrik für $R = 1$ [Bishop 1995, S. 208]. Sie ist in Abbildung 5.3 dargestellt.

¹ Im Gegensatz zu einigen Veröffentlichungen [u.a. Bishop 1995; Hanson und Burr 1988] ist hervorzuheben, dass die quadratische Fehlerfunktion nicht der Minkowski- R -Metrik für $R=2$ entspricht, sondern aufgrund der fehlenden Quadratwurzel lediglich proportional zu ihrem Quadrat ist [Neuralware 1991, S. 95].

² BISHOP leitet dies korrespondierend zu der quadratischen Fehlerfunktion als Maximum-Likelihood-Schätzer aus der Normalverteilung her [Bishop 1995].

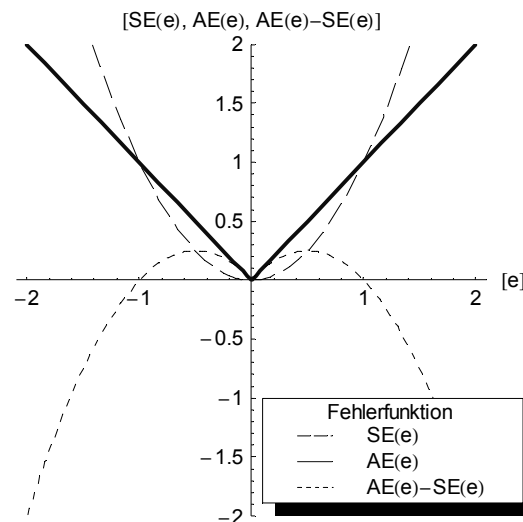


Abbildung 5.3: Fehlerfunktionen des absoluten Fehlers $AE(e)$, des quadratischen Fehlers $SE(e)$ und der Differenz beider Fehlerfunktionen $AE(e)-SE(e)$

Durch die Anwendung der absoluten Fehlerfunktion AE wird ein NN auf die Approximation des Medians der Ausgabedaten trainiert und nicht länger auf die Approximation des Mittelwertes durch den quadratischen SE [Bishop 1995, S. 210]. Dies wird deutlich, wenn die Fehlerfunktion $E_p(e_p)$ des absoluten Fehlers $AE_p(e_p)$ auf einem einzelnen Datenmuster p ,

$$E_p(e_p) = AE_p(e_p) = \sum_j |t_j - o_j| \quad , \quad (5.11)$$

hinsichtlich der Netzwerkausgabe o_j minimiert wird zu

$$\frac{\partial E_p(e_p)}{\partial o_j} = \sum_j \text{sgn}(t_j - o_j) = 0 \quad . \quad (5.12)$$

Die Summe aller Vorzeichen $\text{sgn}(e_p)$ in Formel (5.12) ist dann minimiert, wenn o_j den Median der Ausgabewerte $\{t_p\}$ bezeichnet.¹ Der Betrag einzelner extremer Ausprägungen von t_p hat somit keinen Einfluss auf die Lösung von o_p , weshalb Ansätze des Medians auch als robuste Statistiken bezeichnet werden [Bishop 1995]. Die Fehlerfunktion des $AE(e)$ erscheint vorteilhaft, da er für normalverteilte Prognosewerte identisch ist mit dem Mittelwert, jedoch bei abweichender Verteilung eine genauere Schätzung der zentralen Lageparameter erlaubt.

Die lineare Fehlerfunktion des $AE(e)$ ist im Gegensatz zur quadratischen Fehlerfunktion an ihrer Nullstelle jedoch nicht vollständig differenzierbar und kann daher nicht standardmäßig durch den Backpropagation-Algorithmus minimiert werden. Dies kann jedoch sowohl durch eine Änderung der softwaretechnischen Implementierung als auch durch algorithmische Anpassungen der Ableitung umgangen werden. Die Erweiterung der Backpropagation-Lernregel zum sogenannten Manhattan-Training nach SUTTON [ZELL 2000, S. 118]² entspricht der Verwendung einer linearen Fehlerfunktion des AE . Sie adaptiert die Formel der Ableitung in der konventionellen Backpropagation-Lernregel

¹ In diesem Fall besitzt die gleiche Anzahl von Beobachtungswerten t_p einen höheren Wert als o_j , wie Beobachtungswerte einen niedrigeren Wert als o_j besitzen.

² BISHOP [1995] nimmt hingegen keinen Bezug auf das Manhattan-Training, sondern nur zu zwei anderen Entwicklungen von [Hanson und Burr 1988; Burrascano 1991], die nachfolgend eingehend untersucht werden.

$$\Delta_p w_{ij} = \eta o_{pi} \delta_{pj} \quad , \text{ mit } \delta_{pj} = t_{pj} - o_{pj} \quad (5.13)$$

lediglich durch den Term des Vorzeichens an

$$\Delta_p w_{ij} = \eta o_{pi} \text{sign}(\delta_{pj}) \quad . \quad (5.14)$$

Entsprechend erfolgt die Korrektur der Verbindungsgewichte w_{ij} nicht mehr in Abhängigkeit von den Beträgen der Fehlersignale, sondern lediglich ihrer Vorzeichen. Die Erweiterung erlaubt eine Berechnung der Fehlersignale δ_{pj} wie in der Backpropagation-Lernregel und umgeht so Probleme der mangelnden Differenzierbarkeit der Fehlerfunktion des $AE(e)$ in ihrer Sprungstelle $e = 0$.¹ Nur vereinzelte Erweiterungen des Backpropagation-Algorithmus auf Lernregeln höherer Ordnung berücksichtigen Median-basierte Statistiken in der Ableitung der Fehlerfunktion. So modifiziert die Lernregel Resilient Propagation (Rprop) nach RIEDMILLER und BRAUN in Anlehnung an das Manhattan-Training die Gewichte lediglich nach dem Vorzeichen des Gradienten [Riedmiller und Braun 1993].² Allerdings wurden diese Lernalgorithmen nur vereinzelt systematisch überprüft oder angewendet. Eine Literaturrecherche zeigt nur eine Veröffentlichungen zum Manhattan-Training [Ueda et al. 1993] und nur elf Publikationen zu Rprop in allen Anwendungsbereichen und Disziplinen [Hubert 1993; Vollmer und Strey 1999; Zhou et al. 2002; Wang et al. 2003; Souza et al. 2004; Jiao et al. 2004; Fernandes et al. 2005; Mastorocostas und Rekanos 2005; Treadgold und Gedeon 1997; Igel und Husken 2003; Kantsila et al. 2004; Shaheed 2005].

Als Alternative zur nicht-differenzierbaren $AE(e)$ -Fehlerfunktion schlagen NEUNEIER und ZIMMERMANN [1998] die Verwendung eines skalierten Kosinus Hyperbolicus ($\ln(\cosh)$) als Fehlerfunktion vor. $\ln(\cosh)$ repräsentiert eine Approximation des $AE(e)$, deren Verlauf sich für einen Skalierungsparameter $c = 10$ dem Verlauf des absoluter Fehlers $AE(e)$ annähert, wie in Abbildung 5.4 grafisch dargestellt ist.

¹ SUTTON [1992] bezeichnet seine Weiterentwicklung als eigenständigen Lernalgorithmus, auch wenn es sich vielmehr um die Anwendung einer abweichenden Zielfunktion bzw. Fehlerfunktion in der Berechnungsvorschrift des Backpropagation-Algorithmus handelt. Da die Berechnungsvorschrift der Backpropagation jedoch unverändert verwendet wird, erscheint eine Bezeichnung in Abhängigkeit der Zielfunktion als zutreffender.

² Darüber hinaus berücksichtigt Rprop auch die Steigungen vorangegangener Gewichtsänderungen wie in den Algorithmen SuperSAB oder Quickprop sowie eigene Parameter zur Steuerung der Schrittweite je Verbindungsgewicht wie in der SuperSAB- und der Delta-Bar-Delta-Lernregel [Igel und Husken 2003]. Ihre Funktionalität übersteigt somit die der Manhattan-Regel in einigen Aspekten, sie stellt daher keine rein Median-basierte Lernregel dar.

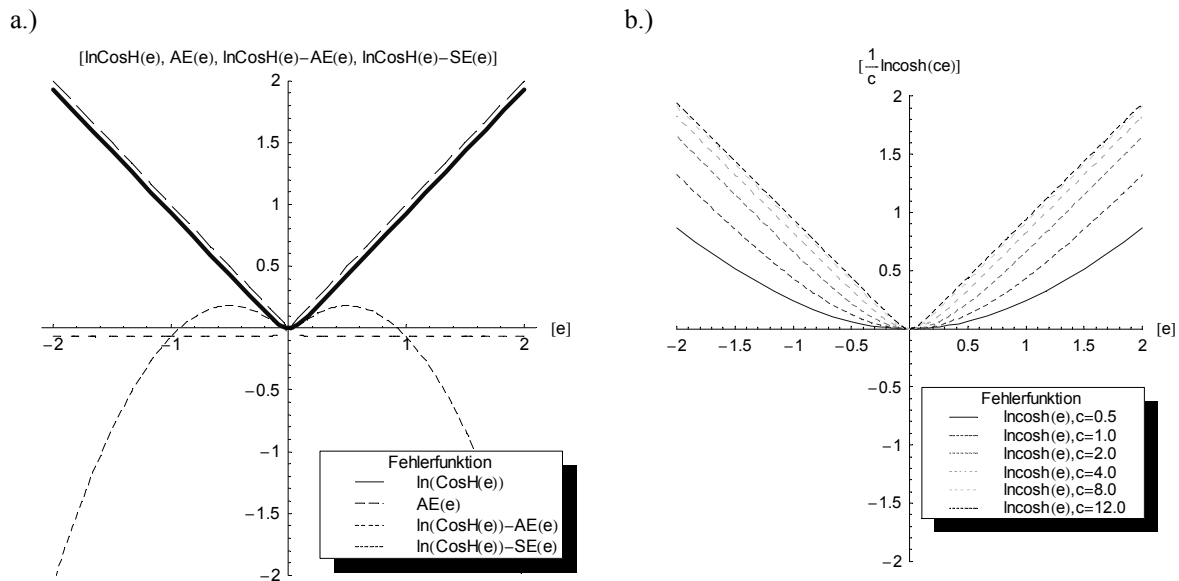


Abbildung 5.4: Fehlerverlauf der $\ln(\cosh)$ -Fehlerfunktion, der absoluten Fehlerfunktion $AE(e)$ und der Differenz von $\ln(\cosh(e)) - AE(e)$ für $c = 10$ (a.) und für unterschiedliche c (b.)

$\ln(\cosh(e_p))$ ist eine kontinuierliche, vollständig differenzierbare Fehlerfunktion, die somit problemlos abgeleitet und mittels Backpropagation oder erweiterter Lernverfahren höherer Ordnung minimiert werden kann. Die Fehlerfunktion des $\ln(\cosh)$ ist von der Form

$$E_p(e_p) = \ln(\cosh(t_p - o_p)) \quad (5.15)$$

und wird durch eine Konstante c an den linearen Fehlerverlauf des $AE(e)$ angenähert, mit

$$E_p(e_p) = \frac{1}{c} \ln(\cosh(c(t_p - o_p))) \quad (5.16)$$

Ihre Ableitung entspricht dem um c skalierten Tangens Hyperbolicus (\tanh), mit

$$\frac{1}{c} \frac{\partial E_p(e_p)}{\partial o_p} = \tanh(c(t_p - o_p)) \quad (5.17)$$

was die Berücksichtigung in den Gleichungen des Backpropagation-Algorithmus und seiner Erweiterungen erleichtert. NEUNEIER und ZIMMERMANN [1998] zeigen neben der theoretischen Herleitung der Fehlerfunktion jedoch keine Experimente oder Anwendungen. Darüber hinaus nimmt einzig THIESING [1998, S. 222] Bezug zur $\ln(\cosh)$ -Fehlerfunktion und verwendet sie in Experimenten zur Absatzprognose von 28 Zeitreihen eines Supermarktes. Neben einer Randbemerkung zu ihrer Überlegenheit in den Experimenten präsentiert oder diskutiert er jedoch keinerlei Ergebnisse, sondern verwendet die quadratische Fehlerfunktion des $SE(e)$.¹

Ziel der Varianten der linearen Fehlerfunktion in Form des absoluten Fehlers, des $\ln(\cosh)$ oder des Manhattan-Trainings als Varianten des $R = 1$ bzw. l_1 -Loss ist eine höhere Robustheit gegenüber Ausreißern. Darüber hinaus kann ihre Anwendung auch zu einer Verbesserung des

¹ Er postuliert auch die Möglichkeit einer linearen, asymmetrischen $\ln(\cosh)$ -Funktion mit unterschiedlichen Skalierungsparametern für positive und negative Fehlerbeträge durch Ausreißer [Thiesing 1998, S. 32]. Er gibt jedoch keine Hintergründe für ihre Verwendung, Ableitung oder Entstehung und wendet sie in den nachfolgenden Experimenten – trotz anderslautender Hinweise im Text – anscheinend nicht an.

Lernverhaltens und der Konvergenz im Trainingsprozess führen. Nach Zell bewirkt die Anwendung linearer Fehlerfunktionen eine Quasi-Normierung der Schritte auf der Fehlerfläche und führt somit zu einer Lösung von Problemen geringer Gradienten auf flachen Plateaus sowie hoher Gradienten in steilen Tälern der Fehlerfläche [Zell 2000, S. 118].

Neben Minkowski-Fehlerfunktionen mit $R < 2$ können auch höherwertige Fehlerfunktionen zum Training von MLPs verwendet werden, die hohe Fehlerbeiträge noch stärker bewerten als der quadratische Fehler $SE(e)$. Die Software NeuralWorks stellt beispielsweise eine kubische Fehlerfunktion der Norm $R = 3$ bereit [Neuralware 1991, S. 95],

$$E_p(e_p) = \frac{1}{3} \sum_j (t_{pj} - o_{pj})^3, \quad (5.18)$$

sowie eine quartische Fehlerfunktion der $R = 4$ -Norm,

$$E_p(e_p) = \frac{1}{4} \sum_j (t_{pj} - o_{pj})^4. \quad (5.19)$$

Die Implementierung in einem Software-Simulator liefert einen empirischen Hinweis auf die Möglichkeit und ggf. praktische Notwendigkeit, die Standard-Fehlerfunktion des quadratischen $SE(e)$ durch eine andere, vollständig differenzierbare Fehlerfunktion ersetzen zu können.¹

BURRASCANO [1991] erweitert den Backpropagation-Lernalgorithmus ferner zur Minimierung der Supremum-Norm mit $R = \infty$ zur Approximation einer Box-Car-Verteilung als extreme Form einer diskontinuierlichen, uniformen Verteilung. Experimente mit logischen Funktionen ohne Rauschen zeigen eine beschleunigte Konvergenz des NN im Lernprozess.

Darüber hinaus erweitern HANSON und BURR [1988] auf Anregung von WHITE² das Training von NN auf Minkowski- R -Fehlerfunktionen mit dynamischer Anpassung der l_R -Norm zwischen $1 \leq R < \infty$ im Training. Sie erweitern den Backpropagation-Algorithmus zur simultanen Berechnung der Gradienten für eine Anpassung der Verbindungsgewichte w_{ij} und eines variablen R im Trainingsprozess. Die Anpassung von R kontrolliert die Konvergenz des Trainingsprozesses durch die veränderte Gewichtung der Ausgabefehler [Hanson und Burr 1988]. Trotz vielversprechender Experimente auf einfachen Problemklassen beschränken sie ihre Analyse komplexerer Problemstellungen auf konstante R .³ Sie leiten dementsprechend typische Eigenschaften der Fehlerfunktionen auf der Basis von Minkowski- R -Metriken ab. Die Verwendung von Fehlerfunktionen mit $R < 2$ reduziert die Sensitivität der Fehlerfunktion für hohe Fehlerbeträge $e > 1$ einschließlich Ausreißern. So modellieren Metriken mit $1 \leq R < 2$ nicht normalverteilte Wahrscheinlichkeitsverteilungen mit einem stärkeren Rand der Verteilung [Hanson und Burr 1988]. TUKEY empfiehlt eine Fehlerfunktion mit $R = 1,5$ für viele allgemeine verrauschte Problemstellungen. HANSON und BURR [1988, S. 353] stellen dagegen fest, dass die jeweilige Fehlerfunktion abhängig von der Problemstellung und der Verteilung der Zufallsschwankungen variiert und erzielen positive Ergebnisse mit $R = 1,5$ und $R = 1,2$. Im Gegenzug wird für Fehlerfunktionen mit $R \geq 2$ die Sensitivität für hohe Fehlerbe-

¹ NeuralWare bezeichnet sie nach der Form ihrer ersten Ableitungen irreführenderweise als „quadratic error function“ und „cubic error function“ [Neuralware 1991]. Allerdings zeigt die Software weder die Eigenschaften oder Anwendungsgebiete der Fehlerfunktionen auf noch wird ihre Herleitung begründet.

² Die Anregung erfolgte in einer Reihe unveröffentlichter, persönlicher Kommunikationen mit WHITE auf der Grundlage einer unveröffentlichten Version des Artikels [White 1989], die jedoch nicht veröffentlicht sind.

³ Sie führen dies auf zunehmende Probleme mit lokalen Minima zurück, welche einer robusten Konvergenz des Trainingsprozesses widersprechen. Die Zielsetzung einer variablen Anpassung von R erscheint daher fragwürdig.

träge $e > 1$ erhöht, indem große Abweichungen ein hohes Gewicht erhalten. Für Problemstellungen ohne oder mit nur geringen Zufallsschwankungen, z. B. zur Approximation von Boole'schen Operatoren sowie physikalischen oder synthetischen Zeitreihen, erlaubt dies eine schnellere Konvergenz und robustere Modellierung und somit eine valide und reliable Anwendung höherwertiger Fehlerfunktionen [Hanson und Burr 1988, S. 354-357]. Ihre Erkenntnisse zur Konvergenz und Geschwindigkeit werden durch die Experimente von BURRASCANO auf dem XOR-Problem und dem 8-3-8-Encoder-Problem mit einer Fehlerfunktion von $R \rightarrow \infty$ und einer Box-Car-Wahrscheinlichkeitsverteilung bestätigt [Burrascano 1991, S. 129]. Auch wenn für quadratische und höherwertige Fehlerfunktionen mit $R \geq 2$ die Sensitivität für Fehlerbeträge $e > 1$ im Vergleich zum linearen Fehler erhöht wird, so wird sie für niedrige Fehlerbeträge $e < 1$ verringert (vgl. Abbildung 5.3). Dies führt im Fall der skalierten Eingabe- und Ausgabewerte von NN im Intervall $[-1, 1]$ oder $[0, 1]$ und somit $|e| < 1$ zu systematischen Verzerrungen, die bislang in der Literatur nicht berücksichtigt wurden.¹

Darüber hinaus existieren vereinzelte Ansätze zu erweiterten Fehlerfunktionen, die nicht auf den Minkowski- R -Metriken basieren. REED und MARKS schlagen eine bereichsweise lineare Fehlerfunktion vor, bei der ein Korridor akzeptabler Fehlerabweichungen um den wahren Wert keine Fehler verursacht und erst Fehlerabweichungen außerhalb des Akzeptanzbereichs zu einer korrigierenden Lerneingabe führen [Reed und Marks 1999]. Sie geben jedoch keine Hinweise auf eine mögliche Anwendung der Fehlerfunktion oder eine formale Begründung ihrer Herleitung.

Die vorgeschlagene Fehlerfunktion ähnelt dem Konzept der Epsilon-insensitiven Verlustfunktion (engl. *epsilon-insensitive loss-function*) aus der statistischen Lerntheorie nach VAPNIK und CHERVONENKIS [Vapnik und Chervonenkis 1979; Vapnik 1995; Vapnik 1998], von der Form

$$E_p(e_p) = \begin{cases} 0 & \text{für } e_p \leq \varepsilon \\ |e_p - \varepsilon| & \text{für } e_p > \varepsilon \end{cases} \quad (5.20)$$

Die Fehlerfunktion stellt einen wichtigen Bestandteil der Parametrisierung der Verfahren der Support Vector Regression (SVR) dar, auf die hier nicht eingegangen werden kann.² Sie weist Fehlern unter dem Schwellwert ε den Fehlerbetrag 0 und Fehlern über dem Parameter ε einen positiven Fehlerbetrag von $e_p - \varepsilon$ zu [Vapnik 1995].³ Sie ist in Abbildung 5.5 dargestellt.

¹ HANSON und BURR untersuchen ebenfalls keine Fehlerfunktionen für $R < 1$. Fehlerfunktionen für $0 < R < 1$ verletzen einige Annahmen von Metriken, könnten jedoch für bestimmte Anwendungsfälle sinnvoll erscheinen. Für $R < 0$ erhalten geringere Fehlerbeträge $e \approx 0$ hohe Fehlerkosten und hohe Fehlerbeträge nur geringes Gewicht, was im Kontext der Bedarfsprognose zur Warendisposition nur wenig sinnvoll erscheint. Sie werden daher nachfolgend vernachlässigt.

² SVR wird auch als Erweiterung von NN unter Verwendung eines Regularisierungsterms in der Zielfunktion interpretiert, wobei die Wahl eines sigmoiden, radialen oder polynomialen Kernels der SVR mit der Wahl der Aktivierungsfunktion in einem MLP mit einer versteckten Schicht korrespondiert. Jedoch bestimmen SVR die freien Parameter als Problem der quadratischen Programmierung mit linearen Nebenbedingungen, während NN ein nicht-konvexes Minimierungsproblem ohne Nebenbedingungen lösen. Für eine Herleitung ε -insensitiver Fehlerfunktionen als robuste Schätzer siehe [Smola und Schölkopf 2004].

³ VAPNIK bezeichnet die Fehlerfunktion (5.44) als Kostenfunktion, die bestimmt, wie die Prognosefehler bestraft werden [Smola und Schölkopf 2004, S. 9]. Dies ordnet die ε -insensitiven Verlustfunktionen vielmehr in den Kontext der nachfolgenden Diskussion allgemeingültiger Kostenfunktionen ein, weniger in den statistischen Fehlerfunktionen. Sie wird hier jedoch aufgrund ihres direkten Bezuges zu einer angenommenen Wahrscheinlichkeitsverteilung unter Maximum-Likelihood-Annahme als Fehlerfunktion dargestellt.

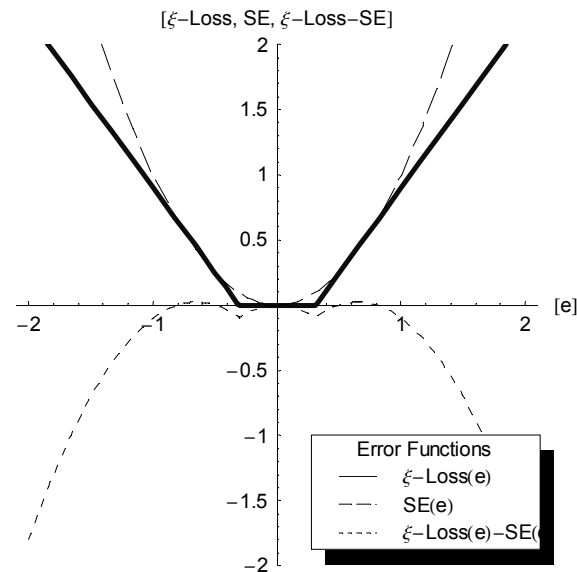


Abbildung 5.5: Epsilon-insensitive Fehlerfunktion der Support Vector Regression¹

Die Epsilon-insensitive Fehlerfunktion zeichnet sich durch einen linearen Anstieg im Intervall $]-\infty, -\varepsilon]$ und $[\varepsilon, +\infty[$ aus, ist symmetrisch um Null, und ihre erste Ableitung hat maximal zwei Diskontinuitäten bei $-\varepsilon, +\varepsilon \geq 0$ und ist Null im Intervall $[-\varepsilon, +\varepsilon]$ [Smola und Schölkopf 2004, S. 3, 9-14]. Die Bestimmung des Parameters ε muss erfolgt in Abhängigkeit von den Zufallschwankungen der Zeitreihe, und ist so zu bestimmen das zufällige Fehlersignale vernachlässigt werden und dadurch die Generalisierungsleistung des Verfahrens gesteigert wird. Der lineare Verlauf der Kostenfunktion vermeidet die Überbewertung von Ausreißern und zielt so auf eine robuste Schätzung. Dies erfordert jedoch die Ex-ante-Bestimmung des ε -insensitiven Bereichs der Fehlerfunktion $[-\varepsilon, +\varepsilon]$, was jedoch der eigentlichen Problemstellung der Prognose ähnelt und somit ein eigenständiges Entscheidungsproblem darstellt.

Trotz der wachsenden Popularität von SVR und ihrer engen Verwandtschaft zu NN existieren nur wenige Untersuchungen zur Anwendung ε -insensitiver Fehlerfunktionen in NN. SUYKENS und VANDEWALLE [2000] erweitern rekurrente NN um ausgewählte Aspekte der SVR einschließlich der ε -insensitiven Kostenfunktion zur Entwicklung eines rekurrenten SVR-Verfahrens.² LESKI und HENZEL [2003] überprüfen ε -insensitive Fehlerfunktionen in Neuro-Fuzzy-Verfahren zur Klassifikation. WANG et al. [2007] untersuchen Verfahren zum fuzzy-clustering. Im Zusammenhang mit der grundsätzlich geringen Beachtung alternativer Ziel- und Fehlerfunktionen für NN erscheint das geringe Interesse an diesen Zielfunktionen jedoch nicht überraschend.

Theoretisch können darüber hinaus sowohl eine Reihe weiterer robuster parametrischer Statistiken der sogenannten R -, L - und M -Schätzer als auch robuste nicht-parametrische Fehlerfunktionen verwendet werden. HANSON und BURR [1988, S. 352] weisen explizit auf die Nutzung robuster Schätzer für stärker verrauschte Daten hin, geben jedoch keine Beispiele. Nachfolgend werden ausgewählte Fehlerfunktionen der robusten Statistik nur kurz erörtert, um ihr Potenzial für das Training von NN zu verdeutlichen und um die Vielfalt etablierter statistischer Zielfunktionen aufzuzeigen, die nicht in der Anwendung von NN verwendet werden. Das Untersuchungsgebiet der robusten Statistik geht zurück auf Untersuchungen von TUKEY [1960, reprinted in 1998] sowie HUBER [1981] und HAMPEL [1986]. Sie untersuchen

¹ Die Epsilon-insensitive Fehlerfunktion wird mit $\varepsilon = 0,3$ berechnet.

² FYFE und MACDONALD [2002] verwenden ε -insensitive Fehlerfunktionen zur PCA in Anwendungen des unüberwachten Lernens von NN, was jedoch nicht mit den Fragestellungen der Prognose verwandt ist.

die Auswirkungen einer Abweichung von den Modellierungsprämissen in Form der Normalverteilung oder der Unabhängigkeit der Residuen zur Entwicklung verbesserter Prozeduren unter Ausreißern und abweichender Verteilungsformen. So werden alternative parametrische Modelle und korrespondierende Fehlerfunktionen entwickelt, welche die quadratische Fehlerfunktion des $SE(e)$ aus der Normalverteilungsannahme durch robuste Fehlerfunktionen ersetzen können.

HUBER [1981] entwickelt eine winsorisierte Fehlerfunktion, die die Eigenschaften des $SE(e)$ und des $AE(e)$ kombiniert und so den Einfluss extremer Beobachtungswerte verringert:

$$E_p(e_p) = \begin{cases} e_p^2 & \text{für } |e_p| \leq c \\ c(2|e_p| - c) & \text{für } e_p > c \end{cases} \quad (5.21)$$

Dies entspricht einer Wahrscheinlichkeitsverteilung mit normalverteiltem Zentrum und doppelt-exponentialverteilten Rändern der Verteilung. Für $c \rightarrow 0$ nähert sie sich der absoluten Fehlerfunktion an und approximiert den Median, für $c \rightarrow \infty$ den Mittelwert als quadratische Fehlerfunktion.¹ Die Fehlerfunktion und ihre erste Ableitung sind in Abbildung 5.6 im Vergleich zu den Fehlerfunktionen $SE(e)$, $AE(e)$ und TUKEYS bivariater Fehlerfunktion dargestellt.

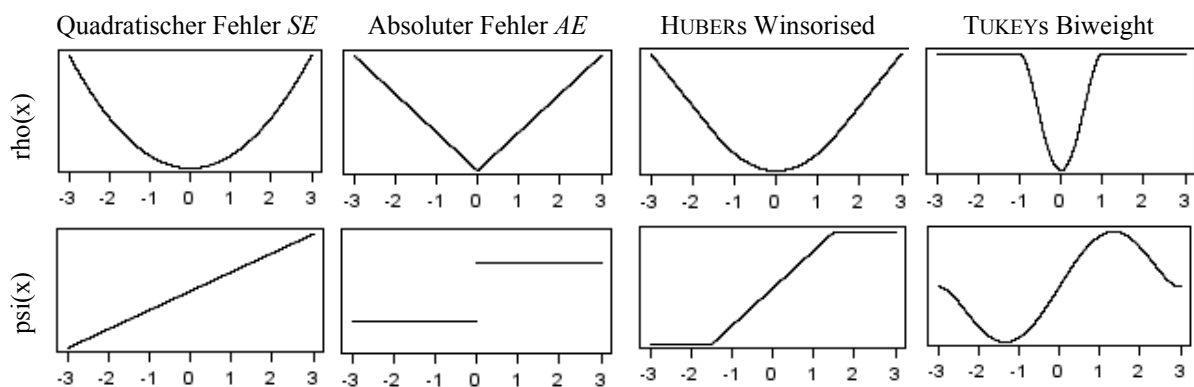


Abbildung 5.6: Fehlerfunktion und erste Ableitungen des quadratischen Fehler $SE(e)$ im Vergleich zu robusten Funktionen des $AE(e)$, HUBERS winsorisierte Fehlerfunktion und TUKEYS Biweight-Fehlerfunktion²

Zur Beschränkung von Ausreißern können gekappte Fehlerfunktionen eingesetzt werden. In Fehlerfunktionen des gekappten Mittelwertes (engl. *trimmed mean*) wird der Einfluss von Ausreißern durch die Fehlerfunktion beschränkt, ohne Korrektur der Daten:

$$E_p(e_p) = \begin{cases} e_p^2 & \text{für } |e_p| \leq c \\ 0 & \text{für } e_p > c \end{cases} \quad (5.22)$$

HAMPEL [1986] erweitert die gekappte Fehlerfunktion in Anlehnung an TUKEYS bivariate Fehlerfunktion; ihre ersten Ableitungen sind in Abbildung 5.7 dargestellt.

¹ Aus der Darstellung der Huber-Fehlerfunktion wird die Nähe zur $\text{Incosh}(e)$ -Fehlerfunktion deutlich, die für niedrige Skalierungsparameter c den Verlauf des Huber-Schätzers approximiert, für hohe c den $AE(e)$.

² Hubers Fehlerfunktion wird mit einem Parameter von $c = 1,5$ gebildet.

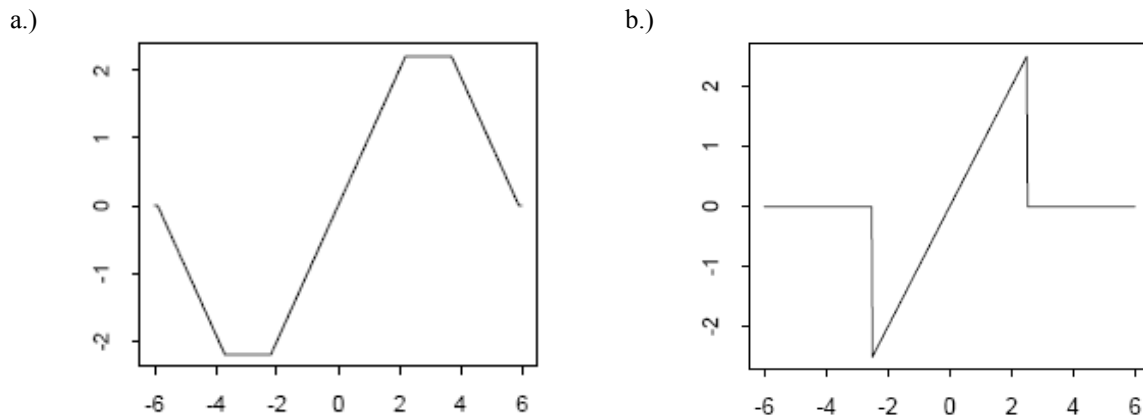


Abbildung 5.7: Erste Ableitungen von HAMPELS Fehlerfunktion (a.) und einer gekappten Fehlerfunktion (b.)

Aus dieser knappen Übersicht wird deutlich, dass vielfache Möglichkeiten zur Anpassung der Fehlerfunktion in NN bestehen, die eine Parametrisierung als robuste Schätzer zulassen.¹ Dennoch zeigt eine Literaturrecherche bislang kaum Untersuchungen von robusten Fehlerfunktionen zum Training von NN.² TAKEUCHI et al. [2002] zeigen einen Ansatz zur robusten Regression für einen asymmetrischen Rand der Verteilung durch mehrere NN, ohne jedoch Details der Verwendung der NN zu diskutieren. ALLENDE et al. [2002] untersuchen einen robusten NN-Schätzer für Zeitreihenprognosen.³

Zusammenfassend ist festzustellen, dass alle dargestellten quadratischen sowie nicht-quadratischen, einschließlich der robusten Fehlerfunktionen symmetrisch um Null verlaufen und aus symmetrischen Wahrscheinlichkeitsfunktionen abgeleitet wurden. Sie berücksichtigen unterschiedliche Abweichungen von der Normalverteilungsannahme und erlauben so eine robuste Schätzung der zentralen Lageparameter des Mittelwertes bzw. Medians als Erwartungswerte der Prognose. Durch die Symmetrie der Gewichtung positiver und negativer Fehlerabweichungen unterscheiden sich statistische Fehlerfunktionen signifikant von vielen empirischen Problemstellungen mit asymmetrischen Fehlerkosten. Sie sind explizit statistisch und nicht empirisch aus der Anwendungsdomäne motiviert.⁴ Darüber hinaus existiert eine Reihe empirischer Problemstellungen mit nicht-symmetrischen Fehlerkosten sowie eine Reihe empirisch motivierter Fehlerfunktionen, die unterschiedliche Aufgabenstellungen abbilden und theoretisch von NN im Training verwendet werden können. Nachfolgend werden diese Fehlerfunktionen im inhaltlichen Zusammenhang statistischer Kostenfunktionen betrachtet und für die Nutzung in der Prognose mit NN hergeleitet, anschließend werden betriebswirtschaftlich motivierte Fehlerfunktionen auf der Basis der Dispositionskosten hergeleitet.

¹ Sowohl die Parametrisierung einzelner NN als auch ihre Kombination in Ensembles homogener oder heterogener robuster Schätzer mit unterschiedlichen Fehlerfunktionen erfordert eine systematische Untersuchung.

² Die Publikation von [Capobianco 2000] untersucht nicht die Eigenschaften der Fehlerfunktion, sondern der Aktivierungsfunktionen auf die Robustheit eines NN-Schätzers. Andere Veröffentlichungen zur robusten Schätzung mit NN entfernen im Vorhinein Ausreißer und trainieren erst im Nachhinein ein NN auf den bereinigten Ergebnissen [Connor et al. 1994; Connor 1996]. Auch wenn dies zu besseren Ergebnissen führt, stellt dieses Verfahren keine Verwendung von NN als robuste Schätzer im eigentlichen Sinne dar.

³ Weitere, unveröffentlichte Arbeiten zu diesem Thema [Chen 2000; McDowall 1997] waren leider nicht zugänglich und konnten daher nicht berücksichtigt werden.

⁴ Darüber hinaus existiert eine Auswahl von Erweiterungen der Zielfunktionen zur Regularisierung, die durch Strafterme die Komplexität der Netzwerkarchitektur in der Fehlerberechnung bestrafen. Da diese jedoch nicht auf eine abweichende Bestimmung der Lernaufgabe zielen, werden sie hier vernachlässigt.

5.3 Zielfunktionen der betriebswirtschaftlichen Entscheidungskosten

5.3.1 Asymmetrische Kostenfunktionen in der Statistik

5.3.1.1 Erweiterung auf asymmetrische Kostenfunktionen

In den vorangehenden Abschnitten wurde hinreichend auf die Bedeutung der Fehlerfunktion im Training von NN und ihre implizite Verwendung durch die Einbettung in den jeweiligen Lernalgorithmus hingewiesen. Begünstigt durch die einfache mathematische Handhabung und die weite Verbreitung des Backpropagation-Algorithmus dominiert die quadratische Fehlerfunktion des $SE(e)$ in nahezu allen Anwendungsbereichen. Somit ist die Mehrzahl aller Publikationen über NN auf die Bildung nichtlinearer Kleinste-Quadrate-Schätzer beschränkt.

WHITE [1988] kritisiert in einer der ersten Veröffentlichungen zur zeitreihenanalytischen Prognose mit NN¹ die unkritische Verwendung quadratischer Fehlerfunktionen:

„Finally, it is extremely important to point out that while the method of least squares (equivalently, back-propagation) is adequate for testing the efficient markets hypothesis, it is not necessarily the method that one should use if interest attaches to building a network for market trading purposes. Such networks should be evaluated and trained using profit and loss in dollars from generated trades, not squared forecast error“ [White 1988, S. 457].

WHITE postuliert somit eine Beurteilung sowohl der Prognosegüte eines NN als auch seines Trainings anhand der tatsächlichen Kosten, die eine Prognose im Anwendungskontext verursacht, und nicht mit quadratischen Fehlern. Er fordert die Übereinstimmung von Lernziel bzw. Lernaufgabe eines NN mit der sie abbildenden Zielfunktion. Diese Überlegungen gelten auch im Kontext der Warendisposition, in der die Beurteilung der Prognosegüte und die Parametrisierung des Prognoseverfahrens anhand von Fehlermaßen losgelöst von der Beurteilung der Entscheidungsqualität durch die verursachten Kosten der daraus bestimmten Waren- und Sicherheitsbestände erfolgt. Vergleichbare Überlegungen zur Inkongruenz von Fehlerfunktionen und dem ökonomischen Entscheidungskontext wurden in der Statistik durch GRANGER initiiert [Granger 1969]:²

„Classical prediction theory limits itself to quadratic cost functions, and hence least-square predictors. However, the cost functions that arise in practice in Economics and management are not likely to be quadratic in form, and frequently will be non-symmetric“ [Granger 1969, S. 199].

GRANGER zeigt, dass Prognosen auf der Basis quadratischer Fehlerfunktionen für generelle Kostenfunktionen der Entscheidung keine optimalen Schätzer liefern. Er postuliert ferner die Schätzung und Beurteilung der Prognose anhand der tatsächlichen Kosten der ökonomischen Entscheidungssituation. Da diese Kosten vielfach auch asymmetrisch und nicht-quadratisch sein können, entwickelt er eine generalisierte, lineare Fehlerfunktion der Entscheidungskosten, die er als Kostenfunktionen (engl. *cost functions*) bezeichnet. Sie sind eine Erweiterung statistischer Fehlerfunktionen und folgen der grundsätzlichen funktionalen Form:

$$C(y_{t+h}, \hat{y}_{t+h,t}) = C(y_{t+h} - \hat{y}_{t+h,t}) = C(e_{t+h,t}) \quad (5.23)$$

Die entstehenden Kosten $C(e)$ werden durch den Wert des Prognosefehlers und die a priori zu bestimmende funktionale Form der symmetrischen oder asymmetrischen Kostenfunktion

¹ Er überprüft NN in der Prognose der täglichen IBM-Aktienkurse [White 1988], vgl. Abschnitt 4.4.1.3.

² Die Entwicklung der Ideen durch GRANGER in der Ökonometrie und WHITE im Bereich der NN muss nicht unabhängig voneinander erfolgt sein, da beide an der University of California San Diego (UCSD) unterrichten und mehrere wissenschaftliche Publikationen gemeinsam verfasst haben, wenn auch nicht zu diesem Thema.

$C(\bullet)$ bestimmt, welche der Distanz zwischen Netzwerkausgabe und tatsächlichem Beobachtungswert einen Kostenbetrag zuweist. Diese Kostenfunktionen werden nachfolgend analysiert und auf das Entscheidungsproblem der Warendisposition mit NN übertragen.

Ökonomische Prognosen erfolgen im Kontext betrieblicher Entscheidungen. „In the real, non academic world, forecasts are made for a purpose and the relevant purpose in economics is to help decision makers improve their decisions“ [Granger und Pesaran 2000a]. DIEBOLD bestätigt die Bedeutung des Entscheidungskontexts einer Prognose: „Forecasts are not made in a vacuum. ... forecasts are made to guide decisions. ... Forecasts are made in a wide variety of situations, but in every case forecasts are made and are of value because they aid in decision making“ [Diebold 2001, S. 32]. Durch ökonomische Entscheidungen auf der Basis einer fehlerhaften Prognose entstehen dem Entscheider Kosten aus der Wahl einer suboptimalen Handlungsalternative. Jeder ökonomischen Prognose liegt grundsätzlich ein Entscheidungsproblem zugrunde, mit dem eine spezifische Verlust- oder Kostenstruktur assoziiert ist. Somit kann jeder Handlungsalternative ein Verlust zugewiesen werden, der eine ordinale Rangreihung der Wahlakte und eine Auswahl zulässt. Diese Kosten nehmen für eine korrekte Entscheidung Null an und für inkorrekte Entscheidungen einen positiven Verlust durch die tatsächlich entstehenden monetären Kosten bzw. durch die Opportunitätskosten der unterlassenen Handlungsalternative [Diebold 2001, S. 34].¹

Die Kostenstruktur der Entscheidungssituation induziert eine ähnliche Kostenstruktur für die Prognose, welche die Information zur Entscheidungsfindung liefert [Granger und Pesaran 2000b; Granger 2001; Diebold 2001, S. 34]. BROCKHOFF stellt dazu fest: „Es ist nicht zu erwarten, dass die nach statistischen Gesichtspunkten ex post beurteilte Güte von Prognoseverfahren systematisch mit der Beurteilung auf Grundlage eines Entscheidungsmodells übereinstimmt“ [Brockhoff 1977, S. 52]. Somit kann die Güte eines Prognoseverfahrens nicht endgültig anhand statistischer Prognosefehler beurteilt werden, sondern nur anhand der resultierenden Fehlerkosten [Granger 1980; Brockhoff 1977; Chambers et al. 1971] oder anhand korrespondierender nutzentheoretischer Überlegungen [Schwarze 1980, S. 318].

Darüber hinaus postuliert GRANGER: „Asymptotically, if we believe that a particular criterion ... should be used to evaluate forecasts then it should also be used at the estimation stage of the modelling process“ [Granger 1993]. Um eine optimale Prognose für ein spezifisches Entscheidungsproblem zu entwickeln, sollte das Prognoseverfahren daher dasselbe Kriterium der Kostenfunktion zur Erstellung der Prognose berücksichtigen wie in der Beurteilung der Prognose- und Entscheidungsgüte. „Zur formalen Beurteilung eines Prognosemodells sollte in erster Linie ein Fehlermaß verwendet werden, das dem Optimierungs- und Anpassungskriterium des Modells entspricht“ [Schwarze 1980, S. 319].² Die Kostenfunktion der Entscheidung kann so einheitlich zur Parametrisierung, zur Beurteilung der Prognosegüte und zur Beurteilung der Entscheidungsgüte verwendet werden. „Despite the rationale behind this, and the role of forecasting in decision-making under uncertainty, the relation between predictions and

¹ Dies erklärt die vielfach synonyme Bezeichnung einer Kostenfunktion als „Verlustfunktion“ (engl. *loss function*) $L(e)$. Nachfolgend wird in Anlehnung an die Originalarbeiten GRANGERS [1969] die etablierte Bezeichnung „Kostenfunktionen“ verwendet, um den direkten Bezug zu den ökonomischen Kosten einer Entscheidung in Anwendungen der Disposition aufzuzeigen, die hier besonders interessiert. Dies erlaubt auch eine Unterscheidung der vielfach synonym verwendeten Begriffe Verlustfunktion, Kostenfunktion, Fehlerfunktion und Zielfunktion. Verlustfunktionen und Kostenfunktionen der Prognose entstammen der Anwendungsdomäne des Entscheidungsproblems. Dagegen bezeichnen Fehlerfunktionen statistische Distanzmaße in der Beurteilung oder Parametrisierung von Prognosen ohne unmittelbaren Entscheidungsbezug, weshalb GRANGER explizit eine Unterscheidung in Fehlerfunktionen versus Kostenfunktionen empfiehlt [Granger 1969]. Zielfunktionen formulieren dagegen das Minimierungsproblem eines Verfahrens wie der NN und können beide Funktionen abbilden.

² Allerdings schlägt SCHWARZE [1980, S. 319] die Parametrisierung von Verfahren nach dem Kriterium der kleinsten Quadrate und die Verwendung der Varianz der Prognosefehler als geeignetes Fehlermaß vor.

decisions has surprisingly escaped the attention of most economists” [Capistrán-Carmona 2005].

Der Verlauf der Kostenfunktion sollte die Struktur der empirischen Entscheidungssituation durch suboptimale Prognosen approximativ repräsentieren und kann von dem symmetrischen Verlauf statistischer Fehlerfunktionen abweichen. Die Kostenfunktion des quadratischen Fehlers $SE(e)$,

$$C(e_{t+h,t}) = e_{t+h,t}^2, \quad (5.24)$$

wird von DIEBOLD als quadratischer Verlust (engl. *quadratic loss* oder *squared error loss*) bezeichnet [Diebold 2001, S. 35]. Seine Anwendung impliziert eine spezifische symmetrische, kontinuierliche und homogene Kostenstruktur der Entscheidungsfehler, in der hohe positive bzw. negative Fehler höhere Kosten verursachen als geringe Fehler und die Zuwachsrate der Kosten steigt. Die Fehlerfunktion des absoluten Fehlers wird korrespondierend als absoluter Verlust (engl. *absolute loss* oder *absolute error loss*) bezeichnet und impliziert eine konstante Zuwachsrate mit dem Betrag des Fehlers.

Auch wenn höhere Beträge der Prognosefehler in vielen ökonomischen Entscheidungssituationen höhere Kosten verursachen, folgen die Fehlerkosten nur in Ausnahmefällen einem symmetrischenquadratischen oder kontinuierlich-linearen Verlauf. „In many important decision environments, a symmetric loss structure closely approximates the true losses of the forecaster. In other decision environments, however, symmetric loss may not be realistic; in general, there is no reason for loss to be symmetric” [Diebold 2001, S. 33]. Im betrieblichen Entscheidungskontext, z. B. bei der Warendisposition, spielt nicht nur die absolute Höhe der Fehler, sondern insbesondere das Vorzeichen des Prognosefehlers durch Über- oder Unterprognose eine entscheidende Rolle [Granger 1980, S. 13]. Sind die Kosten durch Über- bzw. Unterprognosen nicht identisch, bildet sich eine asymmetrische Kostenfunktion. Dies erlaubt eine Erweiterung der Betrachtung von statistischen Fehlerfunktionen $E(e)$ auf generalisierte, asymmetrische Kostenfunktionen der Prognosefehler (engl. *asymmetric cost functions*).¹ GRANGER [1999] sowie DIEBOLD [2001] entwickeln drei grundlegende Bedingungen für generalisierte Kostenfunktionen $C(e)$:

- (1) $C(0) = 0$.
- (2) $C(e) > 0$ für $e \neq 0$, und
- (3) $C(e)$ ist monoton steigend in $|e|$

Die Anforderungen spezifizieren, dass kein Verlust eintritt, wenn der Prognosefehler 0 ist, im Sinne einer fehlerfreien Prognose (1), dass der zugewiesene Verlust positiv ist für alle Prognosefehler (2) und dass der zugewiesene Verlust kontinuierlich ansteigt, wenn der absolute Betrag des Prognosefehlers steigt (3). Große Prognosefehler führen so zu hohen Verlusten, woraus ebenso folgt, dass nahezu identische Prognosefehler ebenfalls ähnliche Verluste verursachen [Diebold 2001]. Darüber hinaus bestehen keine Einschränkungen hinsichtlich der Form der Kostenfunktion. GRANGER [1999] spezifiziert mögliche Eigenschaften von Verlustfunktionen, die jedoch nicht erfüllt sein müssen:

- Symmetrie: $C(-e) = C(e)$
- Homogenität: $C(ae) = h(a) C(e)$
- Konvexität
- Kontinuität

¹ Sie werden auch als nichtsymmetrische Kostenfunktionen (engl. *non-symmetric cost functions*) bezeichnet, oder durch VARIAN [1975] als asymmetrische Verlustfunktionen (engl. *asymmetric loss functions*).

- Differenzierbarkeit

Zulässige Kostenfunktionen umfassen auch die statistischen Fehlerfunktionen der Minkowski- R -Metriken für $0 < R < \infty$ und der statistischen Prognosefehler aus Abschnitt 3.3 als symmetrische Unterklassen¹, gehen jedoch über diese hinaus.² Die theoretischen Anforderungen für generalisierte Kostenfunktionen erlauben die Entwicklung einer Vielzahl unterschiedlicher symmetrischer und asymmetrischer Funktionen und somit die Abbildung unterschiedlicher Lernziele für NN durch ihre Anwendung.³

GRANGER [1969] entwickelt darauf aufbauen eine neuartige, lineare Kostenfunktion zur korrekten Abbildung der asymmetrischen Entscheidungskosten. Die analytische Untersuchung ihrer Eigenschaften zeigt, dass ein optimaler Schätzer zur Prognose unter asymmetrischer Kosten oder unter einer asymmetrischen, nicht normalverteilten Wahrscheinlichkeitsfunktion systematisch durch einen konstanten Term vom Mittelwert abweicht. „It is suggested that a useful, although sub-optimal, manner of taking into account generalized cost functions is to add a constant bias term to the predictor” [Granger 1969, S. 199]. In diesen Fällen erlauben quadratische Zielfunktionen keine optimale direkte Vorhersage, sondern erfordern eine Ex-post-Korrektur um einen systematischen Parameter, der durch die Varianz der Residuen bestimmt wird. Diese Betrachtung asymmetrischer Kosten ist im Entscheidungskontext der Warendisposition äquivalent zu der Berechnung von Sicherheitsbeständen des Newsboy-Problems für ein Produkt und eine Planungsperiode (vgl. Abschnitt 2.3.3.1). Die Kosten der Überbevorratung c_o und der Unterdeckung c_u bestimmen den kostenminimalen Punkt auf der kumulierten Verteilungsfunktion gemäß bekannten Formel (2.24) des kostenminimalen Servicegrades:

$$F_Y(S^*) \geq \frac{c_u}{c_o + c_u} \quad (5.25)$$

Die Berücksichtigung asymmetrischer Kostenfunktionen erlaubt somit die Bestimmung eines kostenminimalen Quantils für unterschiedliche asymmetrische Entscheidungskosten sowie für asymmetrische Wahrscheinlichkeitsfunktionen. Dies legt eine Verwendung der asymmetrischen Kostenfunktionen zur unmittelbaren Berechnung kostenminimaler Warenbestände anstelle einer sequenziellen Ermittlung nahe.

Diese Erkenntnisse wurden trotz regelmäßiger Erinnerungen [Granger 1980, S. 11-16; Granger und Newbold 1986, S. 115-119; Granger 1993] nur langsam zu einer Theorie

¹ Sie werden von CHRISTOFFERSEN und DIEBOLD [1997a] als Prognosefehler-Verlustfunktion (engl. *prediction-error loss function*) bezeichnet und nachfolgend vereinfachend als Prognosefehlerfunktion oder einfach Fehlerfunktion bezeichnet, in Abgrenzung zu Verlust- oder Kostenfunktionen. Nicht alle Distanzfunktionen sind jedoch Metriken im mathematischen Sinne. So entspricht die Distanzfunktion des absoluten Fehlers AE der L_1 bzw. R_1 -Norm in Form der City-Block-Distanz der Minkowski- R -Metrik, der quadratische Fehler SE jedoch nicht der L_2 -Norm der Euklidischen Distanz, die die Wurzel der Summe über alle quadratischen Fehler SE entspricht. Dennoch werden vielfach Distanzfunktionen als Metrik bezeichnet und der SE der L_2 -Norm gleichgesetzt.

² So sind auch die im Abschnitt 5.1 dargestellten Zielfunktionen der Wendepunktprognose und der Klassifikation mit symmetrischen oder asymmetrischen Klassifikationskosten als Kostenfunktionen zulässig [Diebold 2001, S. 36]. DIEBOLD stellt symmetrische und asymmetrische Verlustfunktionen für binäre Klassifikationsaufgaben am Beispiel der Warendisposition mit binären Entscheidungen dar [Diebold 2001]. GRANGER untersucht ebenfalls Fälle von Kostenfunktionen zur nominalen Klassifikation [Granger und Pesaran 2000a]. Für die Anwendung asymmetrischer Klassifikationskosten sei auf die Literatur des *cost sensitive learning* im Data Mining verwiesen.

³ Auch für Klassifikationsaufgaben kann den einzelnen Entscheidungsalternativen direkt ein monetärer Verlust zugeordnet werden; so können generalisierte Kostenfunktionen entwickelt werden. Die nominale Kodierung der Handlungsalternativen erlaubt eine vereinfachte und unmittelbare Zuordnung der Kosten, was die zunehmende Verbreitung kostensensitiver Klassifikationsverfahren begründet. Für die Regression nehmen sowohl der Prognosewert als auch der tatsächliche Beobachtungswert kontinuierliche Werte an, was die Bildung einer stetigen Kosten- bzw. Verlustfunktion erfordert.

generalisierter asymmetrischer Kostenfunktionen erweitert, die eine Schätzung, Beurteilung und Selektion von Prognoseverfahren in Übereinstimmung mit den Kostenfunktionen der Anwendung und der Nutzenfunktionen der statistischen Entscheidungstheorie erlaubt.¹ VARIAN [1975] entwickelte eine asymmetrische Kostenfunktion mit linear-exponentiellem Verlauf, in der zu niedrige Prognosen ein lineares Kostenwachstum verursachen, zu hohe Prognosen hingegen ein exponentielles Wachstum.² CHIRSTOFFERSEN und DIEBOLD [1994] erweitern die Bestimmung optimaler Punktprognosen unter asymmetrischen Kosten für unterschiedliche Annahmen der Varianz, insbesondere zeitlicher Heteroskedastizität, veröffentlichten die Erkenntnisse jedoch erst später [Christoffersen und Diebold 1997a; Christoffersen und Diebold 1997b]. GRANGER erweitert die Gültigkeit seiner Ergebnisse auf Einschnitt- auf Mehrschrittprognosen [Granger 1999]. Weiterführende Forschung konzentriert sich auf die theoretische Bestimmung statistisch optimaler Schätzer für lineare asymmetrische Kostenfunktionen [Cain 1989; Parsian 1990; Parsian und Farsipour 1993; Parsian und Farsipour 1997; Weiss 1996; Parsian und Farsipour 1999; Steinberger und Zinner 1999; Chattopadhyay et al. 2000] sowie für nichtlineare asymmetrische Kostenfunktionen [Zellner 1986a], die Modell-Selektion unter asymmetrischen Kosten [Phillips 1996; Christoffersen und Diebold 1996] und die Beurteilung von Prognosen unter asymmetrischen Kosten [Diebold und Mariano 1995; Diebold und Mariano 2002; Granger und Pesaran 2000a]. Dabei ist die Untersuchung asymmetrischer Kosten zur Beurteilung der Prognosegüte in Lagerhaltungssystemen durch ARMINGER und GÖTZ [1999] aufgrund ihrer Relevanz für die vorliegende Arbeit besonders hervorzuheben. Sie berücksichtigen die asymmetrischen Kosten jedoch nicht in der Parametrisierung von Prognoseverfahren, sondern lediglich für ihre Ex-post-Bewertung.

Die Ausführungen durch Granger sind mit den Entwicklungen zur Quantil-Regression nach KOENKER et al. [Koenker und Bassett 1978; Koenker und Dorey 1987; Koenker 2005] verwandt, die jedoch über die Bestimmung eines einzelnen Quantils der Verteilungsfunktion die Beschreibung der gesamten Wahrscheinlichkeitsfunktion im Sinne einer Verteilungsprognose anstrebt. Ist diese bekannt, können Entscheider mit unterschiedlichen Entscheidungskosten optimale Entscheidungen unmittelbar aus der Verteilungsprognose ableiten, weshalb nach GRANGER und PESARAN [2000a] im Grunde nur eine Prognose der gesamten Wahrscheinlichkeitsfunktion erforderlich ist.³

Darüber hinaus wird die Analyse und Anwendung betrieblicher asymmetrischer oder generalisierter Kostenfunktionen in Theorie und Praxis weitgehend vernachlässigt. „Forecasting theory in economics is just beginning to take this relation into account for the formulation, estimation, evaluation and interpretation of forecasting models“ [Capistrán-Carmona 2002]. Eine Vielzahl von Wissenschaftlern der Prognostik und Ökonometrie erwähnt die grundsätzlich Möglichkeit zur Anwendung von Kostenfunktionen in der Evaluation oder Parametrisierung, geht darauf jedoch nicht näher ein [Brockhoff 1977; Makridakis et al. 1998; Schwarze und Backert 1980; Smola und Schölkopf 2004]. Die Forschung zu asymmetrischen Kostenfunktionen durch GRANGER wird trotz der geringen wissenschaftlichen Aufmerksamkeit fortgesetzt [persönliche Kommunikation mit Granger 2004]. Asymmetrische Kostenfunktionen werden daher nachfolgend sowohl für den linearen als auch für den nichtlinearen Fall hergeleitet und analysiert, und anschließend im Training von NN verwendet.

¹ CAPISTRAN-CARMONA [2005] bietet einen kritischen Überblick über die daraus hervorgehende Forschung. Das erste Kapitel dieser unveröffentlichten Dissertation ist als Review verfügbar [Capistrán-Carmona 2002].

² Sie wird als *LINEX*-Kostenfunktion bezeichnet und ist in Abschnitt 5.3.1.3 dargestellt.

³ Aufgrund der zunehmenden Rechenleistung erstellt beispielsweise die Bank von England derartige Dichteprognosen für Inflation und Wirtschaftswachstum [Capistrán-Carmona 2002].

5.3.1.2 Lineare asymmetrische Kostenfunktionen

Das Konzept der asymmetrischen Kostenfunktion wird in Übereinstimmung mit dem Untersuchungsgegenstand der vorliegenden Arbeit und dem ursprünglichen Beispiel durch GRANGER [1969; 1980]¹ anhand der einperiodigen Warendisposition eines Artikels y exemplifiziert.² Dabei bedingt ein positiver oder negativer Prognosefehler aufgrund des Abweichens der Prognose vom tatsächlichen Wert der abzusetzenden Menge unterschiedliche, asymmetrische Fehlerkosten.

Zum Zeitpunkt t wird für die Planungsperiode $t+h$ ein zukünftiger Bedarf bzw. Absatz von \hat{y}_{t+h} prognostiziert. Daraufhin wird exakt die Menge \hat{y}_{t+h} bestellt und unmittelbar geliefert, in $t+h$ sind drei unterschiedliche Szenarien der Entscheidungssituation möglich. Im Fall einer exakten Prognose wurde genau die bestellte Menge \hat{y}_{t+h} verkauft, mit $\hat{y}_{t+h} = y_{t+h}$ und $e_{t+h} = 0$, und es entstehen keine Kosten im Sinne einer optimalen Entscheidung. War die Prognose zu hoch, $\hat{y}_{t+h} > y_{t+h}$ und $e_{t+h} < 0$, so wurden mehr Waren bestellt als abgesetzt werden konnten. Es verbleibt ein Warenüberbestand am Ende der Entscheidungsperiode, für den Bestandskosten in Höhe von b pro überbevorrateter Verkaufseinheit entstehen. War die Prognose des Warenbedarfes hingegen zu niedrig, $\hat{y}_{t+h} < y_{t+h}$ und $e_{t+h} > 0$, so wurden weniger Waren bestellt als hätten abgesetzt werden können. Es tritt also ein Unterbestand auf, bei dem Fehlmengenkosten von a für jede in der Periode zusätzlich verkaufbare Einheit entstehen.³ [Granger 1980, S. 13]. Die Kostenfunktion einer fehlerhaften Prognose $C(e_{t+h})$, mit $e_{t+h} = (y_{t+h} - \hat{y}_{t+h})$, wird demnach allgemeingültig beschrieben durch

$$C(e_{t+h}) = \begin{cases} a e_{t+h} & \text{für } e_{t+h} > 0 \\ 0 & \text{für } e_{t+h} = 0 \\ b (-e_{t+h}) & \text{für } e_{t+h} < 0 \end{cases} \quad (5.26)$$

Die resultierende Kostenfunktion geht auf die Arbeit von GRANGER [1969] zurück und wird nach CHRISTOFFERSEN und DIEBOLD [1996; 1997a] aufgrund ihrer linearen funktionalen Form als lineare-lineare Kostenfunktion (engl. *linear-linear cost function, LLC*) oder kurz als Kostenfunktion des Typs LINLIN bezeichnet. In Übereinstimmung mit statistischen Fehlern nimmt sie den Wert Null als Maß einer exakten Prognose an und steigt links und rechts des Ursprungs monoton linear an. Für $a \neq b$ steigt die Kostenfunktion nicht symmetrisch um den Ursprung, woher ihre Bezeichnung als lineare asymmetrische Kostenfunktion stammt. Eine alternative Schreibweise der LINLIN-Kostenfunktion verdeutlicht ihre asymmetrische Form:

$$LINLIN(e_{t+h}) = \begin{cases} a |y_{t+h} - \hat{y}_{t+h}| & \text{für } y_{t+h} > \hat{y}_{t+h} \\ b |y_{t+h} - \hat{y}_{t+h}| & \text{für } y_{t+h} \leq \hat{y}_{t+h} \end{cases} \quad (5.27)$$

Abbildung 5.8 bildet exemplarisch eine asymmetrische Kostenfunktion des Typs LINLIN ab:

¹ Ein für das Auftreten asymmetrischer Kosten häufig angeführtes Beispiel, nämlich eine Ankunft am Flughafen zehn Minuten zu früh zum Erreichen eines Flugzeuges im Gegensatz zu zehn Minuten zu spät [Capistrán-Carmona 2005, S. 3] ist irreführend, da es sich beim Abflug um ein binäres Ereignis im Sinne der Klassifikation, weniger der Regression handelt.

² Ziel der Disposition ist gemäß Abschnitt 2.1 die Minimierung der Dispositionskosten durch eine Minimierung der Kosten der Warenbevorratung bei einer optimalen Warenverfügbarkeit des Artikels. Die marginalanalytische Herleitung erfolgt unter Vernachlässigung von Beschaffungszeiten und Lieferfehlern (vgl. Abschnitt 2.3 ff.).

³ Bestandskosten entstehen u. a. durch das in den Überbeständen gebundene Kapital der ex ante finanzierten Ware, Kosten für nicht anderweitig nutzbare Lagerkapazität sowie unnötige Ein-, Um- und Auslagerungen etc. Fehlmengenkosten entstehen als Opportunitätskosten in Höhe des entgangenen Deckungsbeitrages nicht verkaufter Einheiten sowie durch einen mittel- bis langfristigen Goodwill-Verlust der unbefriedigten Kundennachfrage. Für eine detaillierte Analyse siehe Abschnitt 2.4.

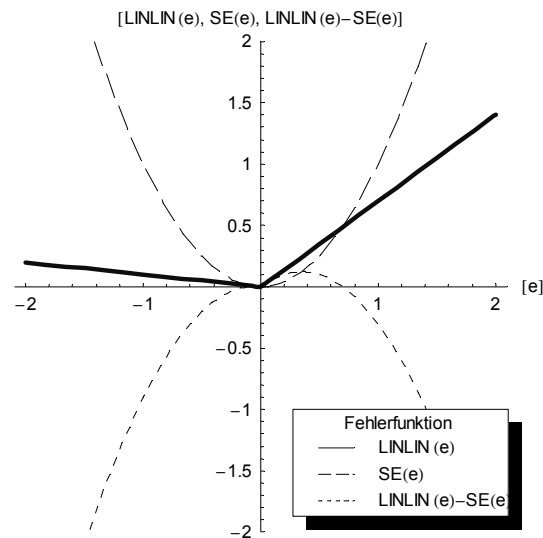


Abbildung 5.8: Asymmetrische $LINLIN(e)$ -Kostenfunktion im Vergleich zum quadratischen Fehler $SE(e)$ und die Differenz beider Funktionen $LINLIN(e)-SE(e)$ ¹

Die Kostenparameter a , b bestimmen für alle $a \neq b$ die Steigung der Kostenfunktion. Sie können als Fehlerkosten pro Mengeneinheit der Differenz zwischen dem wahren Wert y_{t+h} und der Prognose \hat{y}_{t+h} , und somit als Grenzkosten der Entscheidung interpretiert werden. Der Grad der Asymmetrie wird durch das Verhältnis von c_u / c_o determiniert [Christoffersen und Diebold 1997a]. Dies erscheint besonders relevant, da folglich eine exakte monetäre Quantifizierung der Entscheidungsalternativen nicht zwingend erforderlich ist, sondern bereits eine Schätzung des Verhältnisses der Kosten die Ableitung valider Kostenfunktionen erlaubt. Im Fall identischer Parameter $a = b$ liegt eine symmetrische Kostenfunktion vor. Für $a = b = 1$ entspricht die $LINLIN(e)$ -Kostenfunktion der statistischen Fehlerfunktion des $AE(e)$. Symmetrische lineare Kostenfunktionen sowie statistische lineare Fehlermaße bilden somit Sonderfälle linearer asymmetrischer Kostenfunktionen.

Frühe Untersuchungen diskutierten kritisch die empirische Existenz linearer asymmetrischer Kostenfunktionen und die Häufigkeit ihres Auftretens. Die Existenz asymmetrischer Entscheidungskosten in den Wirtschaftswissenschaften kann durch eine marginalanalytische Betrachtung betriebswirtschaftlicher Kosten in der Warendisposition theoretisch verifiziert werden.² Symmetrische Kostenfunktionen unterstellen die Identität der Fehlerkosten $a = b$ für die Überbevorratung bzw. Unterbevorratung jeder zusätzlich verkauften Mengeneinheit. Dies impliziert für Fehlerkosten mit beliebigem symmetrischen Verlauf die Identität der Grenzkosten und der Grenzgewinne einer zusätzlichen verkauften Einheit ab der ersten Verkaufseinheit. Bei einer dichotomen, marginalanalytischen Entscheidung wäre somit die Bevorratung einer zusätzlichen Mengeneinheit selbst unter Vernachlässigung der Fixkosten nicht anreizkompatibel, da keine Gewinne mit ihr erzielt werden könnten. Somit kann argumentiert werden, dass erst asymmetrische Kostenverteilungen den Eintritt und das dauerhafte Engagement in einem Markt unter ökonomischen Prinzipien der Gewinnerzielung rechtfertigen. Daher erscheint die Existenz asymmetrischer Kostenfunktionen theoretisch als auch empirisch plausibel. Die Existenz asymmetrischer Kostenverläufe wird ferner durch die Betrachtung empirischer Entscheidungsprobleme sowie der theoretischen Modelle der Bestellmengenrechnung, z. B. des Newsboy-Problems, bestätigt, die eine Asymmetrie von

¹ Die $LINLIN$ -Funktion besitzt die Parameter $a = 0,7$ und $b = 0,1$.

² Die Darstellung erhebt nicht den Anspruch eines Beweises, insbesondere nicht für unterschiedliche Marktformen oder andere Zustandsparameter der Entscheidungsmodelle der Bestellmengenrechnung.

Fehlmengenkosten $c_{F,L}$ und Bestellkosten c_B , mit $c_{F,L} > c_B$, für die Berechnung stochastischer Modelle voraussetzen. (vgl. Abschnitt 2.3.3.1).

Darüber hinaus erscheint eine Vielzahl nichtlinearer sowie stückweise-linearer Kostenfunktionen mit abnehmenden oder s-förmig verlaufenden Grenzkosten sowie asymmetrischen Grenzgewinnen als wirtschaftlich plausibel, die entsprechend durch asymmetrische Kostenfunktionen approximiert werden können. Die Mehrheit von Untersuchungen zu asymmetrischen Kostenfunktionen konzentriert sich ausschließlich auf lineare Kostenfunktionen des Typs $LINLIN(e)$, da eine Vielzahl empirischer Kostenfunktionen auf die lineare Form zurückgeführt werden kann: „A non-symmetric linear cost function ... will frequently prove to be a useful approximation to any actual non-symmetric cost function“ [Granger 1969, S. 205].

ARMINGER und GÖTZ [1999] erweitern die $LINLIN(e)$ -Kostenfunktion um eine dynamische Komponente zeitabhängiger Kostenparameter zur zeitabhängigen linearen Kostenfunktion (engl. *time dependent linlin asymmetric cost function, TD-LINLIN*). Sie erlaubt dynamische Veränderungen der Einkaufs- und Verkaufspreise, der internen Zinssätze usw. und somit die zeitliche Anpassung der Kostenfunktion „as a basis for deriving a more realistic cost function“ [Arminger und Götz 1999, S. 3]:

$$TDLL(e_{s,t+h}) = \begin{cases} a_{y,t+h} |e_{t+h}| & \text{für } e_{t+h} > 0 \\ b_{y,t+h} |e_{t+h}| & \text{für } e_{t+h} \leq 0 \end{cases} \quad (5.28)$$

Diese Erweiterung erscheint jedoch nur im Fall hochgradig instationärer Parameter mit langem Prognosehorizont plausibel, da die Kostenparameter für zukünftige Planungsperioden ex ante in eigenständigen Prognosen geschätzt werden müssen und somit eine signifikante Instationarität aufweisen müssen. Die zusätzliche Erweiterung der Kostenparameter a, b für unterschiedliche Produkte y [Arminger und Götz 1999, S. 3] erscheint trivial, verdeutlicht jedoch die Notwendigkeit der produkt- und entscheidungsabhängigen Bestimmung der Kostenfunktion.

GRANGER überprüft symmetrische und asymmetrische Kostenfunktionen und ihre Interaktion mit symmetrischen und asymmetrischen Wahrscheinlichkeitsfunktionen, die in der Bestimmung optimaler Schätzer interagieren. In Abbildung 5.9.a ist der Fall einer symmetrischen Kostenfunktion und einer symmetrischen Wahrscheinlichkeitsverteilung nach CAPISTRÁN-CARMONA [2002] dargestellt. Hier ist der optimale Schätzer der Median, der ebenfalls mit dem Mittelwert identisch ist.

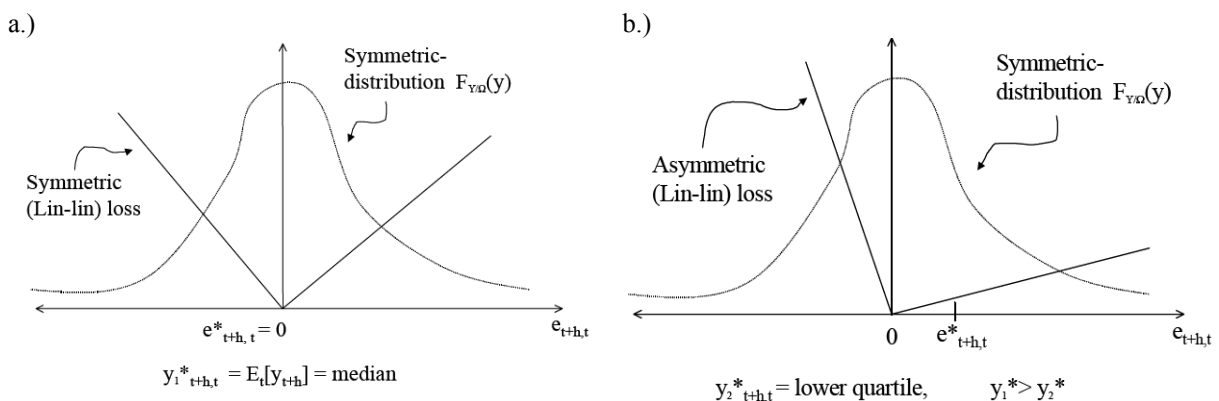


Abbildung 5.9: Optimaler Schätzer einer symmetrischen Verteilung bei symmetrischer Kostenfunktion (a.) und asymmetrischer Kostenfunktion (b.)¹

¹ Abbildung entnommen aus CAPISTRÁN-CARMONA [2002, S. 37].

In Abbildung 5.9.b ist der Fall einer asymmetrischen Kostenfunktion bei symmetrischer Wahrscheinlichkeitsverteilung dargestellt. Der optimale Schätzer wird so unter dem Median liegen, um eine geringe Wahrscheinlichkeit für das Auftreten von Überprognosen zu erzielen, wobei das exakte Quantil durch die Parameter c_u, c_o der Kostenfunktion bestimmt wird.

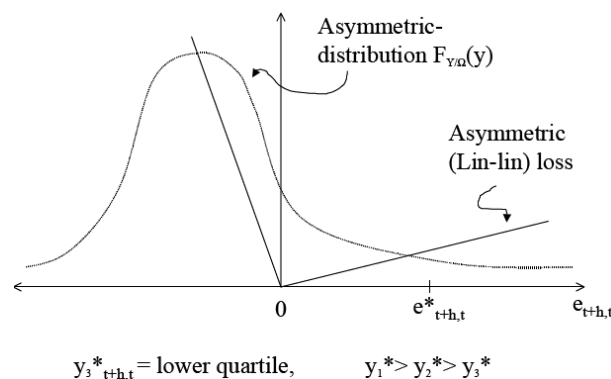


Abbildung 5.10: Optimaler Schätzer einer asymmetrischen Verteilung bei einer asymmetrischen Kostenfunktion¹

Der Fall einer asymmetrischen Kostenfunktion und einer rechtsschiefen Wahrscheinlichkeitsverteilung mit einem Mittelwert größer als der Median ist in Abbildung 5.10 dargestellt. Der optimale Schätzer liegt erneut unter dem Median, jedoch in einem anderen Quantil aufgrund der Veränderung der Verteilungsfunktion bei konstanten Kosten.

5.3.1.3 Nichtlineare asymmetrische Kostenfunktionen

Die lineare Grundform asymmetrischer Kostenfunktionen kann auf beliebige nichtlineare, monoton steigende, nichtstetige, asymmetrische Kostenfunktionen erweitert werden. In Anlehnung an die populäre quadratische Fehlerfunktion $SE(e)$ kann eine asymmetrische, quadratisch-quadratische Kostenfunktion (QQC) gebildet werden:

$$QQC(e_{t+h}) = \begin{cases} ae_{t+h}^2 & \text{für } e_{t+h} > 0 \\ be_{t+h}^2 & \text{für } e_{t+h} \leq 0 \end{cases} \quad (5.29)$$

Sie wird in Anlehnung an die *LINLIN*-Kostenfunktion auch als *QUADQUAD*-Kostenfunktion bezeichnet. Sie ist exemplarisch in Abbildung 5.11 dargestellt. Ebenso wie im linearen Fall erlauben die Kostenparameter die Kontrolle des Grades der Asymmetrie und des Anstiegs der Fehlerkosten für Über- und Unterprognosen.

¹ Abbildung entnommen aus CAPISTRÁN-CARMONA [2002, S. 38].

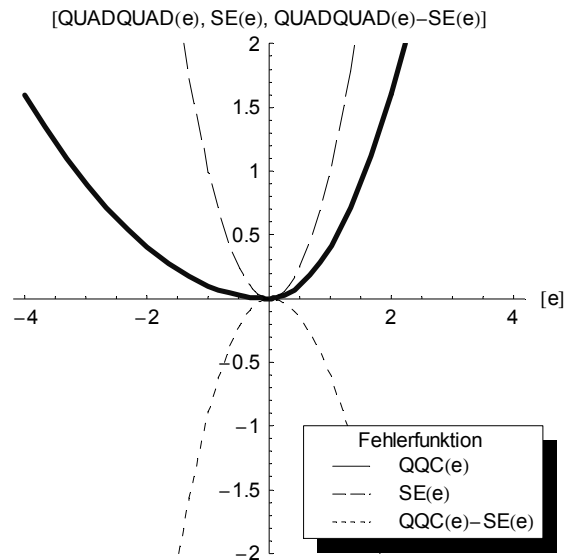


Abbildung 5.11: $QQC(e)$ -asymmetrische Kostenfunktion im Vergleich zur quadratischen Kostenfunktion $SE(e)$ und ihrer Differenz $QQC(e) - SE(e)$ ¹

VARIAN [1975] erweitert die Betrachtung um die *LINEX*-Kostenfunktion, die durch einen linearen und einen exponentiellen Ast gekennzeichnet ist²:

$$LINEX(e_{t+h}) = d \left[\exp(c(e_{t+h})) - c(e_{t+h}) - 1 \right] \quad \text{für } c \neq 0, d \in \mathbb{R} \quad (5.30)$$

Die Skalierungsparameter c und d erlauben die Anpassung der *LINEX*-Kostenfunktion an die empirische Kostensituation. Für $c > 0$ verläuft die Kostenfunktion angenähert linear im Wertebereich links des Ursprungs und angenähert exponentiell im Wertebereich rechts des Ursprungs, vice versa für $c < 0$ [Arminger und Götz 1999, S. 5]. Eine Darstellung der asymmetrischen *LINEX*(e)-Kostenfunktion im Vergleich zum quadratischen Fehler und für unterschiedliche Skalierungsparameter findet sich in Abbildung 5.12.³

Darüber hinaus können beliebig kombinierte lineare oder nichtlineare Kostenfunktionen eingeführt werden, die durch die Kombination von Schwellwerten und spezifischen Kostenverläufen komplexe, auch interdependente empirische Zusammenhänge abbilden können. ARMINGER und GÖTZ [1999] stellen eine Funktion des kombinierten linearen oder nichtlinearen Verlustes vor (engl. *combined linear loss*, *COLL*):

$$COLL(e_{t+h}) = \begin{cases} C_{a1}(e_{t+h}) & \text{für } S_t > e_{t+h} > 0 \\ C_{a2}(e_{t+h}) & \text{für } 0 < e_{t+h} \leq S_t \\ 0 & \text{für } e_{t+h} = 0 \\ C_b(e_{t+h}) & \text{für } e_{t+h} < 0 \end{cases} \quad (5.31)$$

¹ Die $QQC(e)$ -Kostenfunktion verwendet die Skalierungsparameter $a = 0,1$ und $b = 0,4$.

² VARIAN [1975] entwickelt die *LINEX*-Kostenfunktion in einer Untersuchung der Immobilienwirtschaft. Eine Unterbewertung von Immobilien verursacht einen linearen Verlust von Einkünften, während eine Überbewertung kostspielige Rechtsstreitigkeiten und hohe Folgekosten verursacht. Nach ZELLNER [1986a] wird die Kostenfunktion auch zur Kostenprognose im Staudammabau angewendet, bei der eine zu niedrige Vorhersage des maximalen Wasserstandes um ein Vielfaches teurer ist als eine zu hohe Vorhersage.

³ Die Abbildung verdeutlicht die Ähnlichkeit der *LEC*-Kostenfunktion mit der QQC -Kostenfunktion. *LLC* ist jedoch als Funktion im Gegensatz zu QQC vollständig differenzierbar, wodurch Probleme bei der Minimierung im Lernverfahren des Gradientenabstiegs vermieden werden.

Dabei kann $C_b(e_{t+h})$ beispielsweise einen linearen Verlauf annehmen, während die funktionale Form von $C_a(e_{t+h})$ vom Betrag e_{t+h} abhängt, der beim Übersteigen eines Schwellwertes S_t einen veränderten Kostenverlauf bedingt. Im Fall der Warendisposition ist vielfach anzunehmen, dass neben den linear wachsenden Kapitalkosten ab einer Bestandsgrenze zusätzliche Kosten für Warenbewegungen und Lagerung entstehen [Arminger und Götz 1999, S. 4].

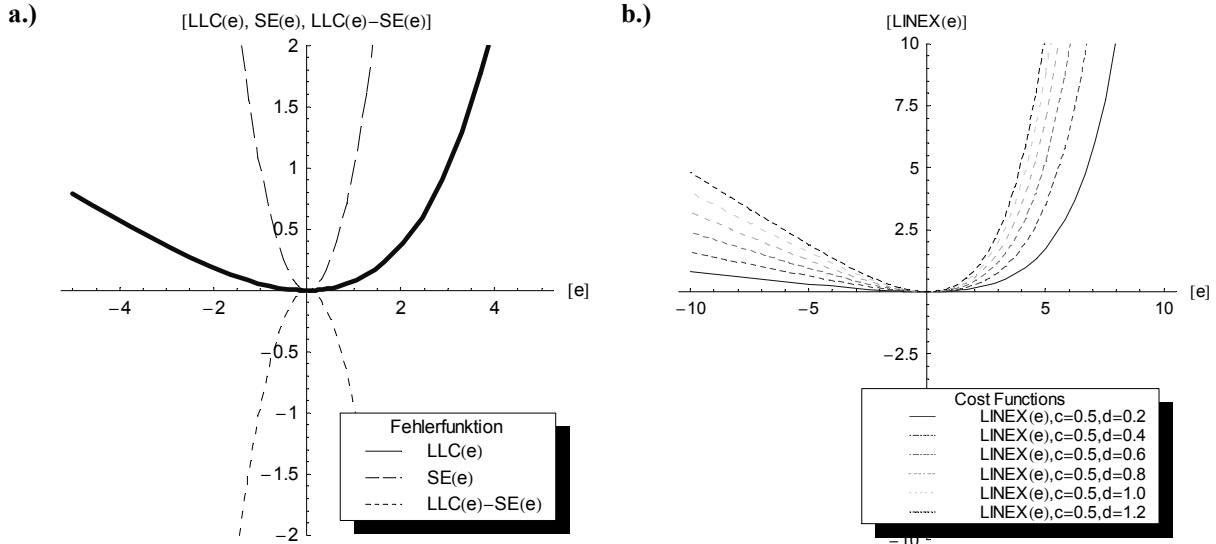


Abbildung 5.12: $LINEX(e)$ -Kostenfunktion im Vergleich zur Kostenfunktion $SE(e)$ und zur Differenz $LINEX(e) - SE(e)$ (a.) sowie unterschiedliche Parametrisierungen von c, d für $LINEX(e)$ (b.)

Für alle ermittelten Kostenwerte der Prognosefehler $C(e)$ können anschließend summierte, mittlere und prozentuale Fehlermaße erstellt werden, die ohne Skaleneffekte eine valide und reliable Bewertung einzelner Verfahren hinsichtlich ihres Beitrages zur Entscheidungsfindung als auch den Vergleich mehrerer Verfahren zulassen. Aufgrund der einheitlichen, metrischen Vergleichsgröße der Kosten ist auch eine beliebige Aggregation über beliebige Hierarchiestufen mehrerer Produkte, Warengruppen oder Absatzstellen möglich. Eine exemplarische Übersicht der üblichen Transformationen wird in Tabelle 5.4 gegeben.

Tabelle 5.4: Asymmetrische Kostenfunktionen zur Beurteilung der Entscheidungskosten

	Lineare Kostenfunktionen	Quadratische Kostenfunktionen	Linear-Exponentielle Kostenfunktionen
Einfache Fehler	$LLC(e_{t+h}) = \begin{cases} a e_{t+h} & \text{for } e_{t+h} > 0 \\ b e_{t+h} & \text{for } e_{t+h} \leq 0 \end{cases}$ linear linear cost function (5.32)	$QQC(e_{t+h}) = \begin{cases} a_t e_t^2 & \text{for } e_t > 0 \\ b_t e_t^2 & \text{for } e_t \leq 0 \end{cases}$ quadratic-quadratic cost function (5.33)	$LEC(e_{t+h}) = d \left[\exp^{ce_{t+h}} - ce_{t+h} - 1 \right]$ for $c \neq 0, d \in \mathbb{R}$ $LINEX$ / linear-exponential cost function (5.34)
Summierte Fehler	$SLLC_l(e_{t+h}) = \sum_{t=l-h}^{T-h} LLC_{t+h}$ Sum of $LINLIN$ cost function (5.35)	$SQQC_l(e_{t+h}) = \sum_{t=l-h}^{T-h} QQC_{t+h}$ Sum of QQC cost function (5.36)	$SLEC_l(e_{t+h}) = \sum_{t=l-h}^{T-h} LEC_{t+h}$ Sum of $LINEX$ cost function (5.37)
Mittlerer Fehler	$MLLC_l(e_{t+h}) = \frac{1}{T-l} SLLC_l$ Mean $LINLIN$ cost function (5.38)	$MQQC_l(e_{t+h}) = \frac{1}{T-l} SQQC_l$ Mean QQC cost function (5.39)	$MLEC_l(e_{t+h}) = \frac{1}{T-l} SLEC_l$ Mean $LINEX$ cost function (5.40)

Obwohl weitere prototypische Beispiele linearer asymmetrischer Kostenfunktionen denkbar sind, werden diese hier nicht explizit entwickelt.¹ Vielmehr liegt es im Charakter der asymmetrischen Kostenfunktion, dass sie jeweils individuell an die Kostensituation des Prognosegegenstandes angepasst werden muss. Anstelle einer Herleitung archetypischer asymmetrischer Kostenfunktionen erscheint eine situationsbedingte Formulierung und Parametrisierung in Anlehnung an die etablierte betriebswirtschaftliche Kostentheorie sinnvoller, solange sie den zugrunde liegenden Entscheidungstatbestand valide und reliabel für das Prognoseproblem charakterisiert.

5.3.1.4 Anwendungsbereiche asymmetrischer Kostenfunktionen

Prognoseverfahren bilden die Informationsgrundlage betrieblicher Entscheidungen wie der Warendisposition. Sie erfordern die Spezifikation einer Fehler- oder Kostenfunktion nicht nur zur Parametrisierung des Verfahrens, sondern auch zur Ex-post-Beurteilung ihrer Prognosegüte im Vergleich zu konkurrierender Verfahrenskandidaten.

Auch in der Modellauswahl werden überwiegend statistische Fehlerfunktionen einschließlich des populären quadratischen Fehlers $SE(e)$ eingesetzt, nicht die empirischen Kosten der jeweiligen Prognoseverfahren. „Actual cost functions are often nonsymmetric. ... Thus, the choice of an error-squared cost function is justifiable only on pragmatic grounds ...” [Granger 1980, S. 14]. GRANGER [1980, S. 13-14] legitimiert die Verwendung statistischer Fehlermaße zur Modellselektion mit dem häufigen Fehlen spezifischer Kosteninformationen. In diesem Fall dienen statistische Fehlerfunktionen als valide Approximation der empirischen Kostenstruktur und haben keinen Einfluss auf die Rangreihung der Prognoseverfahren. In der Praxis kann die Kostenfunktion $C(e_{t+h})$ einer einzelnen suboptimalen Entscheidung im Zeitpunkt $t+h$ häufig nicht mit vertretbarem Aufwand zur Informationsbeschaffung erhoben und quantifiziert werden. Durch die interdependente Verflechtung der Entscheidung mit anderen betrieblichen Entscheidungen, die Probleme bei der Identifikation der direkt und indirekt zurechenbaren Effekte oder durch die Notwendigkeit zur Schätzung bzw. Prognose der zukünftigen Kostenfaktoren zur Entscheidung ist eine Erhebung vielfach nicht möglich oder wirtschaftlich. Somit muss der Entscheidungsträger eine symmetrische Kostenfunktion als hinreichende Approximation der tatsächlichen Entscheidungskosten annehmen.² GRANGER [1980] betrachtet zwei konkurrierende Prognoseverfahren $g(y_t)$ und $h(y_t)$. Durch den Vergleich ihrer Prognosefehler $e_{t,g}$ und $e_{t,h}$ über alle $t = 1, \dots, T$ wird das Verfahren mit dem niedrigsten Prognosefehler zur Anwendung gewählt:

$$\frac{1}{T_1} \sum_{t=1}^{T_1} C(e_{t,g}) < \frac{1}{T_2} \sum_{t=1}^{T_2} C(e_{t,h}) \quad . \quad (5.41)$$

Unter Unkenntnis der tatsächlichen Kosten für Über- und Unterprognosen werden beide Verfahren mit einer symmetrischen Kostenkonstanten A bewertet, mit

$$\frac{1}{T} \sum_{t=1}^T A e_{t,g}^2 < \frac{1}{T} \sum_{t=1}^T A e_{t,h}^2 \quad . \quad (5.42)$$

¹ Im Gegensatz zu der Analyse der asymmetrischen Kostenfunktionen mit linearem, quadratischem und gemischtem Verlauf in der Ökonometrie und Statistik ist auffällig, dass die in der empirischen Praxis und betriebswirtschaftlichen Kostentheorie verbreiteten degressiven Kostenverläufe mit Sättigungseffekten oder s-förmigen Kostenverläufen bislang vernachlässigt werden. Dies ermöglicht eine Erweiterung der Forschung zum Verlauf unterschiedlicher Kostenfunktionen in mikroökonomischen Entscheidungen.

² Die Notation von GRANGER wird zur Übereinstimmung mit den vorigen Abschnitten entsprechend angepasst.

Durch die symmetrische Anwendung von beliebigen Entscheidungskosten A auf beiden Seiten der Gleichung besitzt A für Zeitreihen identischer Länge $T_1 = T_2$ keine Entscheidungsrelevanz für die relative Güte des Verfahrens [Granger 1980, S. 14]. Da das Prognoseverfahren $g(y_t)$ mit den niedrigeren Kosten unabhängig von der verwendeten Kostenkonstante A gewählt wird, wird die Kostenkonstante üblicherweise uniform durch $A = 1$ ersetzt, und somit vernachlässigt:

$$C(e_{t+h}) = Ae_{t+h}^2 = e_{t+h}^2 \quad (5.43)$$

Somit können die gebräuchlichen statistischen Fehler und Fehlermaße als Kostenfunktionen valide und reliable Anwendung finden und der quadratische Fehler zur validen Rangreihung der Wahlakte der Entscheider verwendet werden.

Dieser Beweis verliert jedoch in (5.41) seine Gültigkeit, wenn empirische asymmetrische Kostenfunktionen oder zumindest Informationen über den Grad der Asymmetrie der Entscheidungskosten vorliegen. GRANGERS [1980] Legitimation statistischer Fehlermaße ist zudem nicht theoretisch motiviert, sondern lediglich eine pragmatische Lösung des Problems der mangelnden Informationsversorgung hinsichtlich der asymmetrischen Kosten. BROCKHOFF bezeichnet die Beurteilung von Verfahren durch statistische Prognosefehler daher lediglich „als Vorstufe einer zielbezogenen Bewertung von Prognoseverfahren“ [Brockhoff 1977, S. 52]. Der Grad der Informationsversorgung und der Entscheidungsunterstützung durch die Informationssysteme des Controlling und der Finanzbuchhaltung hat sich jedoch seit den frühen Untersuchungen GRANGERS erheblich verbessert. Der stationäre Handel erhebt beispielsweise Kennzahlen der direkten Produktrentabilität (DPR) in der Warenbevorratung entlang der Distributionsnetzwerke, so dass Unternehmen heute zumindest in der Lage sind, erforderliche Kosten von Prognosen und einzelnen Bestandsentscheidungen zu schätzen. CHRISTOFFERSEN und DIEBOLD begründen die Dominanz quadratischer Fehlerfunktionen zudem durch die mangelnde Berechenbarkeit komplexer, nichtlinearer Funktionen: „In spite of the need for a practical forecasting framework that incorporates realistic loss functions, until recently one was forced to favor mathematical convenience over realism – quite simply, there was no alternative“ [Christoffersen und Diebold 1996]. Die Rechenkapazität moderner Computer erlaubt zudem auch die Untersuchung und Verwendung realistischer Kostenfunktionen. Somit ist die theoretische Verwendung empirischer asymmetrischer Kostenfunktionen nicht nur theoretisch motiviert und legitimiert, sondern nunmehr auch empirisch möglich.

Die von GRANGER [1980] postulierte Indifferenz hinsichtlich der Fehlerfunktion zur Modellselektion gilt ferner nur für ein einzelnes, ex ante bestimmtes Fehlermaß auf einer einzelnen Zeitreihe und nicht über unterschiedliche Zeitreihen und Fehlermaße hinweg. Dies wurde durch die empirische Beurteilung konkurrierender Prognoseverfahren in der M3-Competition und der sich anschließenden Diskussion hinreichend dokumentiert [Fildes und Makridakis 1988; Zellner 1986b; Chatfield 1988; Armstrong 2001, S. 455-456]. Die Auswahl eines Fehlermaßes zur Beurteilung der Prognosequalität beeinflusst die Rangreihung und entsprechend die Auswahl von Prognosemethoden erheblich. ARMSTRONG und COLLOPY stellen fest, dass in der M3-Competition nur eine geringe Korrelation zwischen dem relativen Rang der konkurrierenden Verfahren auf den unterschiedlichen Fehlermaßen *MAPE*, *MdAPE*, *GMRAE*, *MdRAE* und *PB* bestand [Armstrong J. S. und Collopy F. 1993]. Insbesondere quadratische Fehlermaße zeigen aufgrund ihrer Skalierungs- und Ausreißerempfindlichkeit eine beschränkte Validität und sind zur Modellselektion von Prognoseverfahren nicht geeignet [Armstrong 2001, S. 459; Armstrong und Collopy 1992; Armstrong und Fildes 1995].¹ MAKRIDAKIS et al.

¹ Während ihre Relevanz aufgrund der Gewichtung der Prognosefehler in ihrer ursprünglichen Domäne der Signalverarbeitung und Rauschunterdrückung im Sinne einer empirischen Kostenfunktion sinnvoll erscheint,

legitimieren die Nutzung quadratischer Fehlermaße dennoch, da sie explizit wünschen, hohe Fehler aufgrund ihrer disproportional hohen Kosten zu berücksichtigen [Makridakis et al. 1998, S. 48]. ARMSTRONG weist dies als implizite Unterstellung eines Entscheidungskontexts zurück, der einer Beurteilung durch mehrere unterschiedlicher Fehlermaße entgegensteht: „Some analysts use it [Anm.: RMSE] because larger errors may have severe economic consequences. However, economic consequences should be evaluated by planners and decision makers, not by forecasters” [Armstrong 2001, S. 459].¹ Alle statistischen Fehlermaße implizieren eine spezifische, symmetrische Verteilung der Entscheidungskosten und bedingen so Asymmetrien der Rangreihung und der Selektion von Prognoseverfahren. Die Auswahl eines adäquaten Fehlermaßes zur Modellselektion stellt vielmehr ein eigenständiges Entscheidungsproblem dar, das die Überprüfung seiner Validität und Reliabilität im jeweiligen Zusammenhang der wahren Kostenfunktion erfordert. Liegt also ein konkreter Entscheidungskontext vor, sollte dieser nicht durch statistische Fehlermaße approximiert werden, sondern unmittelbar anhand der relevanten Kosten.

Kosten stellen eine universelle metrische Bezugsgröße mit hoher Akzeptanz in der ökonomischen Theorie und insbesondere der Praxis dar. Anders als statistische Fehlerfunktionen erlauben Kostenfunktionen die unmittelbare, explizite Modellierung entscheidungsrelevanter Kosten ohne den Umweg über statistische Fehlermaße, unter Vermeidung der daraus entstehenden Probleme bei der Spezifikation und Auswahl geeigneter Fehlermaße. Sie erlauben so eine unmittelbare Beurteilung der Vorteilhaftigkeit und Bedeutung einer Prognose im Entscheidungskontext sowie eine Selektion optimaler Prognoseverfahren in der relevanten Dimension der Entscheidung für beliebige symmetrische oder asymmetrische, lineare und nichtlineare Kostenverläufe. Die Bewertung einzelner Fehlerkosten erlaubt zudem den direkten Vergleich der Auswirkungen fehlerhafter Prognosen für Produkte mit unterschiedlichem Umsatz, Rentabilität und Umschlagsgeschwindigkeit unter Berücksichtigung von produktindividuellen Kostenverläufen und eine korrekten Aggregation der Kosten über Produktgruppen und -hierarchien hinweg. Darüber hinaus erlaubt die Kostenbewertung eine unmittelbare Bestimmung des Wertzugewinns durch den Einsatz komplexer Verfahren im Vergleich zu herkömmlichen Verfahren sowie die Bestimmung unterschiedlicher Informationsversorgungen durch die Steigerung der Prognosegüte und die Senkung der Fehlerkosten. Somit können alle verfahrensspezifischen Kosten, einschließlich der Kosten zur Integration der Verfahren und Systeme in die betriebliche Anwendung, direkt durch Verfahren der Investitionsrechnung bewertet werden.

Die Verwendung von Kostenfunktionen bietet in vielen Anwendungsbereichen statistischer Fehlermaße Vorteile, insbesondere bei der Ex-post-Beurteilung der Prognose- und Entscheidungsgüte sowie bei der externen Überwachung der Prognosequalität durch Warnsignale, bei der Modellselektion mehrerer konkurrierender Verfahren sowie bei der Bestimmung optimaler Verfahrensparameter, zudem als laufender Indikator in adaptiven Verfahren. Nachfolgend wird die Analyse asymmetrischer Kosten auf die Beurteilung, Auswahl und Parametrisierung von NN mit asymmetrischen Kostenfunktionen erweitert.

ist sie im betrieblichen Kontext einer Anwendung kritisch zu beurteilen und vielfach überhaupt nicht gegeben.

¹ Aufgrund ihrer offensichtlichen Defizite geht ihre Verbreitung in der Modellselektion in theoretischen Vergleichsstudien sowie der Praxis zurück [Armstrong 2001, S. 460; Schwarze 1980, S. 329-330].

5.3.2 Neuronale Netze und asymmetrische Kostenfunktionen

5.3.2.1 Parametrisierung bei asymmetrischen Kostenfunktionen

Die Möglichkeiten zur Anwendung asymmetrischer Kostenfunktionen im Training von NN wurde bislang nur theoretisch untersucht. WHITE [1988] kündigte in einer der ersten Veröffentlichungen zur Prognose mit NN die Anwendung empirischer Kostenfunktionen an:

„Networks should be evaluated and trained using profit and loss in dollars from generated trades, not squared forecast error. Learning methods for this criterion are under development by the author“ [White 1988].¹

Kurz darauf publizierte WHITE [1992a; 1992b] einen formalen Beweis der mathematischen Fähigkeit von NN, ein beliebiges Quantil auf einer bedingten Wahrscheinlichkeitsverteilung der abhängigen Variablen zu einem beliebigen Genauigkeitsgrad zu approximieren. Die formale Herleitung erweitert seinen vorangegangenen Beweis der universellen Approximationsfähigkeit von NN in HORNIK, STINCHCOMBE und WHITE [1989], der sich auf die Schätzung des bedingten Mittelwertes durch quadratische Kostenfunktionen bezog. Die erweiterte formale Herleitung geht insofern darüber hinaus, als sie belegt, dass auch andere Punkte auf einer bedingten Verteilungsfunktion beliebig genau approximiert werden können. WHITE belegt somit formal, das NN neben der Regression des Mittelwertes theoretisch auch zur Quantil-Regression im Sinne von KOENKER [2005] befähigt sind (vgl. Abschnitt 3.1.2.3). Dazu wird eine Modifikation der Zielfunktionen verwendet, die er jedoch nicht spezifiziert.

Über den formalen Beweis der theoretischen Fähigkeit von NN zur Minimierung beliebiger Fehlerfunktionen hinaus wurde die praktische Befähigung von NN zur Quantil-Regression oder Quantil-Autoregression jedoch bislang nicht untersucht. Die durch WHITE [1992a] postulierte mathematische Fähigkeit beweist jedoch noch keine Validität oder Reliabilität in der tatsächlichen Anwendung von NN, die hier interessiert. Der formale Beweis enthält jedoch - ebenso wie seine frühere Herleitung der universellen Approximationsfähigkeit von NN mit einer Netzwerkarchitektur aus einer unendlichen Anzahl von nichtlinearen Neuronen [Hornik et al. 1989] - keine Hinweise auf die tatsächliche Ausgestaltung der Netzwerkarchitektur, des Lernprozesses oder der Zielfunktion. Dem mathematischen Nachweis der universellen Approximationsfähigkeit folgten seit seiner Veröffentlichung daher tausende Untersuchungen, die nach der theoretischen Fähigkeit auch die erfolgreiche Anwendung von NN in der Klassifikation und Regression zu belegen versuchen.

Eine entsprechende Untersuchung asymmetrischer Kostenfunktionen mit NN steht dagegen noch aus. Durch die Veröffentlichung in unzugänglichen Konferenzpublikationen blieb die Forschung WHITES [1992a] bisher weitgehend unbekannt.² Die Andeutungen von HANSON und BURR [1988] deuten zwar auf ihre Kenntnis der Entwicklung beliebiger Kostenfunktionen für NN durch persönlichen Kontakt zu WHITE hin, die genannten Arbeiten wurden allerdings nicht veröffentlicht. WHITE zitierte in einer späteren Publikation zur parametrischen

¹ Siehe auch das einleitende Zitat durch WHITE [1988] in Abschnitt 5.3.1.

² Die geringe Verbreitung der Erkenntnisse WHITES kann neben der mehrfach belegten geringen Popularität der Forschung zu asymmetrischen Kostenfunktionen auch auf die geringe Verbreitung seiner Publikation zurückgeführt werden. Trotz mehrfacher Veröffentlichung in Konferenzbänden und Sammelwerken [White 1992a; White 1992b] erschien sie in keiner Fachzeitschrift und ist daher kaum zugänglich. Dies wird belegt durch lediglich durch eine umfassende Literatur- und Zitationsanalyse. Der ISI Science Citation Index belegt nur ein Zitat durch Taylor [2000]. Ferner existieren sechs Zitate durch Arbeitspapiere in CiteSeer Citation Index, die jedoch nicht auf NN bezogen sind. Die vergleichsweise geringe Anzahl der Referenzen auf diese Forschung wird im direkten Vergleich mit der vorangehenden Veröffentlichung zum Nachweis der universellen Approximationsfähigkeit durch HORNIK, STINCHCOMBE und WHITE [1989] noch auffälliger, die 2253 Zitaten im ISI und 623 im CiteSeer Citation Index belegen. Der Abruf der Informationen erfolgte am 18. Dezember 2005 von den Webseiten des ISI SCI (<http://portal.isiknowledge.com>) und des CiteSeer (<http://citeseer.ist.psu.edu>), die eine Suche nach zitierenden Publikationen erlauben.

Quantil-Regression selbst: „Of course, one can estimate the conditional quantile model without assuming correct specifications using various non-parametric models such as kernel regression [Sheather and Marron (1990)], nearest-neighbor estimation [Bhattacharya and Gangopadhyay (1990)], or using artificial neural networks [White 1992a]” [Kim und White 2003]. Auch eine kürzlich erscheinende Veröffentlichung von WHITE [2006] weist darauf hin, dass in der exklusiven Betrachtung der Kostenfunktionen des MSE alternative, asymmetrische Kostenfunktionen zur Bestimmung der Quantile vernachlässigt werden. Zu den bekannten Kostenfunktionen gebe es jedoch Erweiterungen für asymmetrische Kostenfunktionen durch ihn, auf deren weitere Darstellung er aber erneut verzichtet.

Allein TAYLOR [2000] greift die Idee zur Bestimmung von Quantilen einer Verteilungsfunktion durch NN auf. Er prognostiziert die zeitliche Entwicklung eines Quantile zur Bestimmung des „Value-at-Risk“ aus mehrperiodigen Finanzerträgen durch ein MLP und vergleicht sie mit GARCH-basierten Schätzern. Dabei folgt er jedoch einer Dekomposition und zerlegt die Zeitreihe vergangener Erträge in drei disjunkte Zeitreihen, die er separat prognostiziert und anschließend rekombiniert. Er berücksichtigt dazu (a.) eine separate Prognose der Zeitreihe vergangener Mittelwerte der Erträge zur Schätzung zukünftiger mittlerer Erträge, im Sinne der konventionellen Prognose, (b.) eine separate Prognose durch Extrapolation der Volatilität der Zeitreihe auf der Basis vergangener Varianzen, und (c.) eine Zeitreihe der unterschiedlichen Zeitbezüge der Erträge, die über eine oder mehrere zukünftige Perioden erwirtschaftet werden können [Taylor 2000, S. 304 f.]. Die separate Verwendung mehrere Zeitreihen widerspricht jedoch dem Grundgedanken durch KOENKER [2005] oder durch WHITE [1992a], führt aber auch zu einer approximativen Schätzung veränderlicher Varianzen und somit Quantile im Zeitablauf. Vor allem vernachlässigt TAYLOR [2000] so die asymmetrischen Kosten der Entscheidungssituation durch eine explizite Formulierung von Kostenfunktionen.

Die Arbeiten durch HUSMEIER [1999] folgen einem ähnlichen Ansatz zur Schätzung von Konfidenzintervallen, in Anlehnung an Arbeiten durch WEIGEND et al. [Weigend und Nix 1994; Weigend und Srivastava 1995] die die unterschiedlichen Momente einer Verteilung durch mehrere separate NN schätzen. HUSMEIER [1999] schätzt Mittelwert und Varianz einer Wahrscheinlichkeitsfunktion durch ein RBF NN aus der Zeitreihe der Beobachtungsdaten, und gibt sie an separaten Ausgabeneuronen aus. Er untersucht sowohl stationäre [Husmeier und Taylor 1997] als auch instationäre Zeitreihen [Husmeier 2000], jedoch erneut nicht im Sinne der Bestimmung von Quantilen oder durch eine Verwendung asymmetrischer Zielfunktionen oder der Entscheidungskosten.

Die mangelnde Untersuchung des Trainings von NN mit asymmetrischen Kostenfunktionen spiegelt sich auch in den wenigen Veröffentlichungen zu parametrischen Ansätzen der asymmetrischen Kostenfunktionen wider. WEISS und schließt seine Veröffentlichung zu asymmetrischen Kostenfunktionen der konventionellen statistischen Verfahren mit den Worten: „I have left a number of issues to future research. These include approximating the optimal predictor using the neural network models“ [Weiss 1996, S. 559]. CHRISTOFFERSEN und DIEBOLD [1996] weisen auf das besondere Interesse der Wissenschaft an einem Vergleich der Verfahrensgüte konventioneller statischer Regressionsverfahren mit dem theoretischen Ansatz der NN unter asymmetrischen Kosten nach WHITE [White 1992a] hin. Auch die vorangegangene Literaturanalyse der Ziel- und Fehlerfunktionen in Abschnitt 5.1 belegt, dass trotz der theoretischen Fähigkeiten von NN keine Untersuchungen zum Training von NN mit Zielfunktionen der betriebswirtschaftlichen Entscheidungskosten oder mit asymmetrischen Kostenfunktionen existieren. Das geringe Interesse erscheint jedoch im Kontext der geringen

Aufmerksamkeit der Entwicklung asymmetrischer Kostenfunktionen in der Statistik und alternativer Zielfunktionen im Training von NN nicht überraschend.¹

Wie auch im Fall der konventionelle Regression zur zeitreihenanalytischen oder kausalanalytischen Prognose bieten NN durch ihre Fähigkeiten (vgl. Abschnitt 4.4.1.1) zur semi-parametrischen, datengetriebenen Bestimmung von linearen oder nichtlinearen Erklärungsmodellen ein erhebliches Potential für die Schätzung von Quantilen einer empirischen Verteilungsfunktion, als Erweiterung der linearen statistischen Verfahren durch KOENKER [2005]. Die Verwendung asymmetrischer Kosten der Entscheidungsbildung im Training von NN führt dabei zu einer Übereinstimmung aller Kriterien der Zielfunktion, der Parametrisierung, der Evaluierung und der Modellselektion eines NN bei der Lösung eines Entscheidungsproblems. Die Fähigkeit zur Parametrisierung, Modellselektion und Beurteilung der Prognose- und Entscheidungsgüte mit NN wird nachfolgend für eine Reihe empirischer Kostenfunktionen analysiert.

5.3.2.2 Erweiterung des Backpropagation-Algorithmus

Die Parametrisierung von NN mit einer asymmetrischen Kostenfunktion erfordert zunächst eine Erweiterung des Backpropagation-Lernalgorithmus und eine Analyse seiner veränderten Funktionsweise. NN bestimmen die Verfahrensparameter der Verbindungsgewichte w_{ij} und der Schwellwerte θ_j durch iterative, nichtlineare Optimierungsverfahren des Gradientenabstiegs, die theoretisch nicht an eine spezifische quadratische Zielfunktion gebunden sind. Somit kann zunächst die Zielfunktion eines MLP² aus (5.1) auf eine Minimierung asymmetrischer Kostenfunktionen erweitert werden, und zwar durch

$$O(W) = \sum_{p=1}^m C(e_p(w_{p,0}, w_{p,1}, \dots, w_{p,n}, x_p)) \rightarrow \text{MIN!} \quad (5.44)$$

Um das Lernen in MLPs mit asymmetrischen Kostenfunktionen beliebiger Form zu ermöglichen, wird der ursprüngliche Lernalgorithmus des Backpropagation nach RUMELHART, HINTON und WILLIAMS [1986] formal erweitert. Insbesondere wird die Bildung der partiellen Ableitungen für Kostenfunktionen beliebiger Art generalisiert, um so die Auswirkungen alternativer Zielfunktionen der Entscheidungskosten zu analysieren.

Anstelle der quadratischen Fehlerfunktion wird eine beliebigen Kostenfunktion $C(\bullet)$ der Prognosefehlern unterstellt, die für jedes Neuron j einen Kostenwert aus der Netzwerkausgabe o_{pj} und der Sollausgabe t_{pj} für ein Datenmuster von Ein- und Ausgabedaten p bestimmt:

$$E_p = C(t_{pj}, o_{pj}) \quad (5.45)$$

Die Erweiterung der Formeln erfolgt analog zu den Schritten der ursprünglichen Veröffentlichung durch RUMELHART et al. [1986], welche hinreichend dokumentiert und -kommentiert sind. Die Bestimmung der Gewichtsänderung $\Delta_p w_{ji}$ erfolgt je Gewicht w_{ij} in Abhängigkeit vom Fehlersignal der Kostenfunktion $C(\bullet)$ mit

¹ Dabei ist zu betrachten, das die Forschung auf den unterschiedlichen Gebieten der Statistik, Ökonometrie und der NN vor allem durch GRANGER, DIEBOLD und WHITE initiiert wurde, die gemeinsam an der University of California, San Diego (UCSD) lehren. Außerhalb dieser Forschungsgruppe existieren nur wenige Beiträge.

² Da die Herleitung des Backpropagation-Algorithmus für vorwärtsgerichtete MLPs erfolgte, beschränkt sich die Gültigkeit der nachfolgenden Ausführungen auf diese Klasse von NN. Die Überlegungen zur Nutzung asymmetrischer Kostenfunktionen sollten jedoch auch auf beliebige Formen von NN erweiterbar sein.

$$\Delta_p w_{ji} \propto - \frac{\partial C(t_{pj}, o_{pj})}{\partial w_{ji}} . \quad (5.46)$$

Die partielle Ableitung der Kostenfunktion kann als Produkt zweier Terme aufgefasst werden:

$$\frac{\partial C(t_{pj}, o_{pj})}{\partial w_{ji}} = \frac{\partial C(t_{pj}, o_{pj})}{\partial net_{pj}} \frac{\partial net_{pj}}{\partial w_{ji}} . \quad (5.47)$$

Unter Einführung des Fehlerterms δ_{pj} für den ersten der Terme von

$$\delta_{pj} = - \frac{\partial C(t_{pj}, o_{pj})}{\partial net_{pj}} \quad (5.48)$$

und nach Anwendung der Kettenregel resultiert daraus

$$\delta_{pj} = - \frac{\partial C(t_{pj}, o_{pj})}{\partial net_{pj}} = \frac{\partial C(t_{pj}, o_{pj})}{\partial o_{pj}} \frac{\partial o_{pj}}{\partial net_{pj}} . \quad (5.49)$$

Aus der unveränderten Formel der Netzwerkausgabe, die von der Form der Ziel- und Kostenfunktion unberührt ist (vgl. Abschnitt 4.2.3)¹,

$$o_{pj} = f_j (net_{pj}) , \quad (5.50)$$

und unter der Annahme, dass

$$\frac{\partial o_{pj}}{\partial net_{pj}} = f'_j (net_{pj}) , \quad (5.51)$$

wird eine erste Lösung der Fehlerberechnung für Neuronen j in der Ausgabeschicht gefunden:

$$\delta_{pj} = \frac{\partial C(t_{pj}, o_{pj})}{\partial o_{pj}} f'_j (net_{pj}) . \quad (5.52)$$

Für alle Neuronen j in den vorgelagerten, versteckten Schichten wird der Fehler durch eine erneute Anwendung der Kettenregel ermittelt als

$$\begin{aligned} \sum_k \frac{\partial C(t_{pj}, o_{pj})}{\partial net_{pk}} \frac{\partial net_{pk}}{\partial o_{pj}} &= \sum_k \frac{\partial C(t_{pj}, o_{pj})}{\partial net_{pk}} \frac{\partial \sum_i w_{ki} o_{pi}}{\partial o_{pj}} \\ &= \sum_k \frac{\partial C(t_{pj}, o_{pj})}{\partial net_{pk}} w_{kj} = - \sum_k \delta_{pj} w_{kj} \end{aligned} \quad (5.53)$$

Unter Einsetzen der ersten beiden Terme in (5.49) wird die Formel gefunden, die durch eine allgemeine Kostenfunktion die Fehlerberechnung für alle versteckten Neuronen j beschreibt:

¹ Die Notation der Aktivierungsfunktion f_j sowie f'_j korrespondiert mit der Notation in der ursprünglichen Herleitung der Backpropagation und entspricht der Notation als $f_{act,j}$ bzw. $f'_{act,j}$ in Abschnitt 4.2.

$$\delta_{pj} = f'_j(\text{net}_{pj}) \sum_k \delta_{pk} w_{kj} \quad (5.54)$$

Dies erlaubt die Berechnung von Lernsignalen eines MLP aus einer allgemeingültigen Kostenfunktion durch eine einfache Generalisierung der Backpropagation in einer unveränderten Lernregel

$$\Delta w_{ij} = \eta o_i \delta_j \quad (5.55)$$

unter einer generalisierten Fehler- bzw. Kostenberechnung durch

$$\text{mit } \delta_j = \begin{cases} \frac{\partial C(t_{pj}, o_{pj})}{\partial o_{pj}} f'_j(\text{net}_{pj}) & \forall \text{ Ausgabezellen } j \\ f'_j(\text{net}_j) \sum_k (\delta_k w_{jk}) & \forall \text{ verdeckte Zellen } j \end{cases} \quad (5.56)$$

Die Gleichung (5.55) zur Bestimmung der Gewichtsänderungen Δw_{ij} aus beliebigen Kostenfunktionen bleibt im Vergleich zur ursprünglichen Formel der Backpropagation-Lernregel (vgl. Formel (4.34) in Abschnitt 4.3.2.2) unverändert. Die Erweiterung des Algorithmus auf asymmetrische Kostenfunktionen hat lediglich Einfluss auf die Berechnung der Fehlerwerte in der Ausgabeschicht in (5.56). Die Generalisierung im Hinblick auf asymmetrische Kostenfunktionen bewirkt hingegen für die Neuronen j der versteckten Schichten keine Änderung der Fehlerberechnung. Somit reicht es aus, allein die Fehlerberechnung der Ausgabeneuronen an die Kosten des individuellen Entscheidungsproblems anzupassen. Alle nachfolgenden Veränderungen des Netzwerkverhaltens werden durch die Rückwärtspropagierung in vorliegende Schichten wie bei konventionellen statistischen Fehlermaßen erreicht.

Aus der Herleitung wird deutlich, dass die Wahl der semilinearen Aktivierungsfunktion f_j unabhängig von der Wahl der asymmetrischen Kostenfunktion ist. Somit können wie zuvor unterschiedliche Aktivierungsfunktionen je Schicht oder Neuron in der Ausgabeschicht sowie in den versteckten Schichten zur Unterstützung des Lernverhaltens ausgewählt werden, losgelöst von der Asymmetrie der Zielfunktion. Da wie in Abschnitt 5.3.1.2 gezeigt alle symmetrischen, statistischen Fehlermaße und Zielfunktionen Spezialfälle einer generellen, asymmetrischen Kostenfunktion repräsentieren, erscheint eine universelle Nutzung der generalisierten Backpropagation-Lernregel möglich.¹ Wie unter 5.2.2 dargestellt können nichtquadratische Zielfunktionen einen funktionalen Verlauf annehmen, der nicht vollständig differenzierbar ist und somit nicht mit dem Backpropagation-Verfahren minimiert werden kann. Der Lösungsansatz von NEUNEIER und ZIMMERMANN [1998] sowie von THIESING [1998] zur Approximation durch eine kontinuierliche Fehlerfunktion kann nicht auf asymmetrische Kostenfunktionen übertragen werden. Die dargestellten Ansätze nach BISHOP [1995] oder SUTTON [1998, zitiert nach Zell 2000] erlauben ferner nur eine Lösung der Fehlerberechnung für symmetrische, kontinuierlich lineare Kostenfunktionen mit einer Sprungstelle.² Anwendungen von asymmetrischen Kostenfunktionen mit empirisch abgeleiteten Entscheidungskosten bilden jedoch vielfach mehrere Bereiche von Kostenverläufen und somit mehrere Sprungstellen der Zielfunktion. Da die vollständige Differenzierbarkeit der Zielfunk-

¹ So kann in Software-Simulatoren eine generalisierte Kostenfunktion für Ausgabeneuronen implementiert werden und durch die jeweilige Bestimmung des Grades der Symmetrie oder Asymmetrie die Form des Fehlerverlaufs angepasst werden, einschließlich der Nutzung quadratischer, statistischer Fehlerfunktionen.

² Aus den Veröffentlichungen zur Verwendung nicht-quadratischer, symmetrischer Fehlerfunktionen im Training von NN durch HANSON und BURR [1988] sowie BURRASCANO [1991] geht hingegen nicht hervor, wie diese implementiert wurden.

tion für gradientenbasierte Lernalgorithmen jedoch erforderlich ist, wird in Anlehnung an ANDERSON [1995] eine geringfügige Erweiterung der Software zur Simulation der NN postuliert, die nicht definierte Sprungstellen identifiziert, mögliche Fehler abfängt und ihre Ableitung gleich Null setzt.

5.3.2.3 Auswirkungen von asymmetrischen Kostenfunktionen

Die Anwendung einer asymmetrischen Kostenfunktion $C(\bullet)$ verändert durch die entscheidungskonforme Gewichtung jedes Fehlers e_i die Form des Fehlergebirges, auf dem der Backpropagation-Algorithmus versucht, durch Gradientenabstieg und Anpassung der Verbindungsgewichte w_{ij} und Schwellwerte θ_j ein Minimum zu finden.

Die Veränderung des Fehlergebirges für mehrschichtige Netzwerktopologien erfordert für j Verbindungsgewichte w_{ij} die Darstellung im j -dimensionalen Raum. Zur Visualisierung wird daher die Auswirkung einer asymmetrischen Kostenfunktion am Beispiel einer vereinfachten Netzwerkarchitektur mit zwei Eingabeneuronen, ohne versteckte Neuronen und mit einem Ausgabeneuron zur Approximation eines nichtlinearen autoregressiven Prozess NAR(p) mit $p = 2$ dargestellt.¹ Anschließend wird der Ausgabefehler des MLP für alle Kombinationen von w_1, w_2 mit den Fehlerfunktionen $SSE(e)$, $\frac{1}{2} SSE(e)$ und einer asymmetrischen, linearen Kostenfunktion $LINLIN(e)$ bewertet und die Ergebnisse grafisch dargestellt. Abbildung 5.13.a zeigt die Fehlerfläche für die Kombination der beiden Verbindungsgewichte w_1 und w_2 unter Anwendung der quadratischen Fehlerfunktion des $SSE(e)$ über alle Eingabemuster. In Abbildung 5.13.b wird dagegen die modifizierte, jedoch üblicherweise im Backpropagation-Algorithmus verwendete Zielfunktion des $\frac{1}{2} SSE(e)$ nach der Originalpublikation durch RUMELHART, HINTON und WILLIAMS [1986] dargestellt.

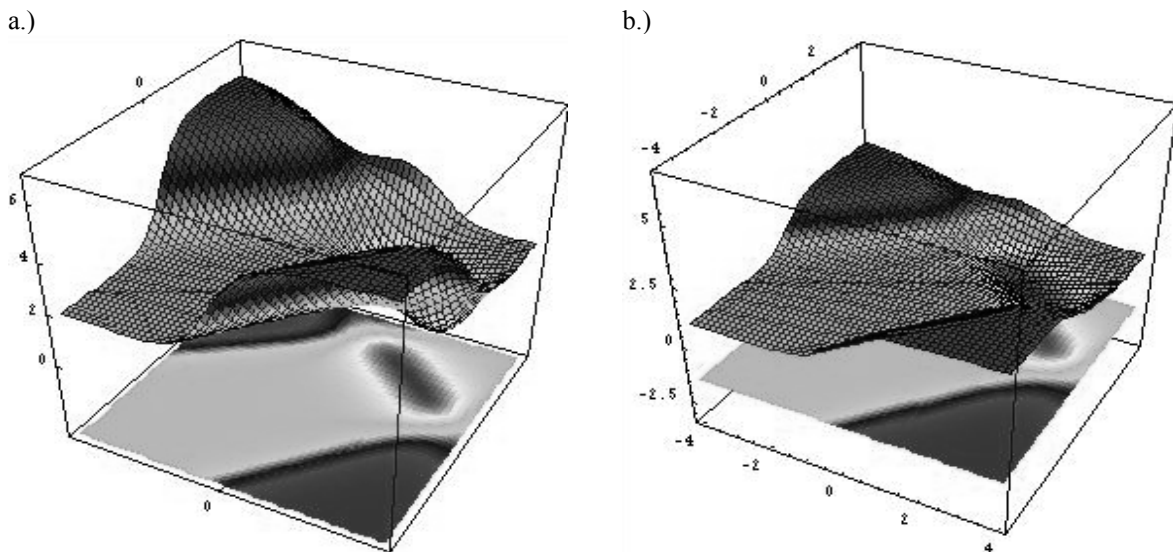


Abbildung 5.13: Bivariate Fehlerflächen der Fehlerkosten für den quadratischen Fehler $SSE(e)$ (a.) und den modifizierten quadratischen Fehler $\frac{1}{2} SSE(e)$ (b.)

Aus dem Vergleich von Abbildung 5.13.a und b wird deutlich, dass die Skalierung der Zielfunktion zu einer Abflachung des Fehlergebirges führt, was unmittelbare Auswirkungen auf die Stärke der Gradienten und somit die Lerngeschwindigkeit hat, nicht aber auf die Richtung des Gradienten. Da es sich um eine monotone Transformation der Fehlerfunktion handelt, bleiben Form und Lage der lokalen und globalen Fehlerminima unverändert.

Im Gegensatz zu der monotonen Transformation von Zielfunktion und Fehlerfläche durch die skalierte $\frac{1}{2} SSE(e)$ -Fehlerfunktion verändert die asymmetrische Kostenfunktion einzelne

¹ Das MLP besitzt zur Vereinfachung keine Schicht mit versteckten Neuronen und keinen Schwellenwert im Ausgabeneuron, so dass lediglich zwei Verbindungsgewichte als freie Parameter zu bestimmen sind.

Bereiche der Fehleroberfläche systematisch. So wird in Abbildung 5.14 a.) die vergleichbare Fehlerfläche für w_1 und w_2 unter Anwendung der asymmetrischen Kostenfunktion des Typs $LINLIN(e)$ visualisiert. Es wird deutlich, dass sowohl Steigung und Struktur der Fehlerfläche als auch lokale und globale Minima signifikant in ihrer Lage und Ausdehnung verändert werden. Zur Veranschaulichung wird in Abbildung 5.14 b.) die Differenz der Fehlerflächen von $LINLIN(e)$ und $SSE(e)$ als $LINLIN(e) - SSE(e)$ explizit dargestellt.

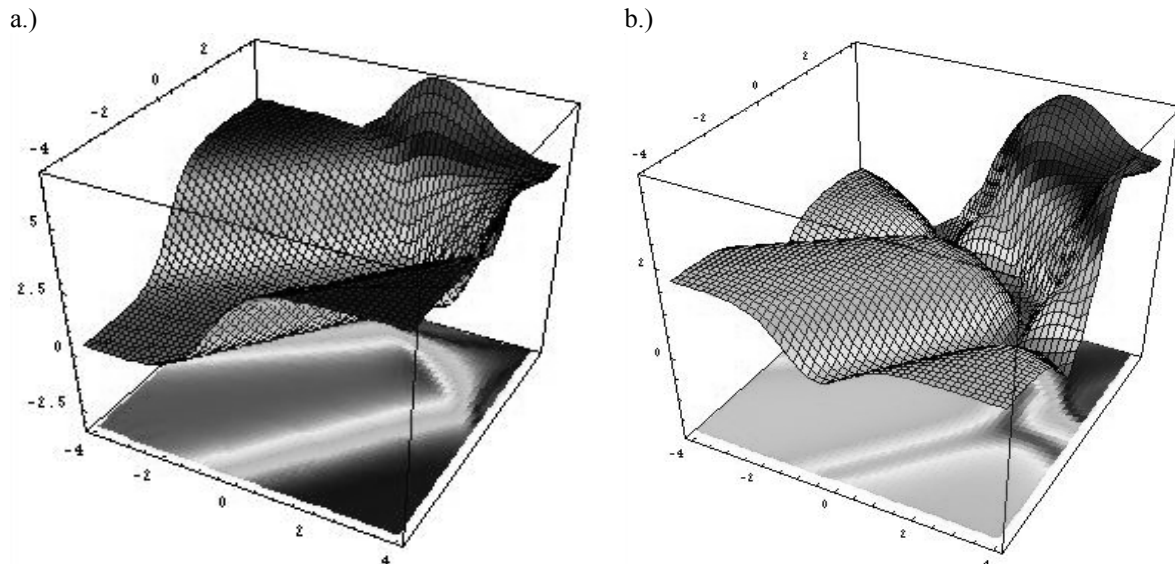


Abbildung 5.14: Bivariate Fehlerfläche der Fehlerkosten für eine asymmetrische Kostenfunktion $LINLIN(e)$ (a.) und der Differenz der Fehlergebirge von $SE(e)$ und $LINLIN(e)$ (b.)

In diesem Fall vereinfacht sich das Lernproblem des NN, da durch die asymmetrische Kostenfunktion die Größe des globalen Optimums vergrößert und somit die Wahrscheinlichkeit einer Konvergenz in dieses Minimum durch lokale Suchheuristiken erhöht wird.¹ Diese Veränderung von Fehlerfläche und Form des globalen Optimums verdeutlicht, dass eine Verwendung asymmetrischer Kostenfunktionen die Lernaufgabe des NN nicht zwingend erschwert. Vielmehr kann durch die Berücksichtigung zusätzlicher Informationen in Form der tatsächlichen Entscheidungskosten, z. B. hinsichtlich besonders kostenintensiver oder insensitiver Entscheidungsrichtungen, auch eine robuste Lernleistung, eine verbesserte Generalisierungsleistung des NN und somit eine effizientere Entscheidung mit verringerten Kosten erzielt werden.²

Nachfolgend sind die Auswirkungen einer asymmetrischen Kostenfunktion auf die Prognose zeitlich geordneter Beobachtungswerte von Zeitreihendaten zu untersuchen. Dies erfolgt zunächst anhand einer Studie zur Einsatzfähigkeit von NN mit unterschiedlichen Kostenfunktionen auf synthetischen Zeitreihen, so als die Experimentumgebung kontrolliert werden kann und generalisierbare Erkenntnisse abgeleitet werden können. Die Auswirkungen asymmetrischer Kostenfunktionen auf empirischen Zeitreihen werden danach in Kapitel 6 eingehen untersucht.

¹ Dies ist jedoch nicht als grundlegende Eigenschaft von ACF zu verstehen, sondern lediglich auf diesen Einzelfall bezogen.

² Eine Verbesserung der Generalisierungs- und Lernfähigkeit eines NN durch asymmetrische Zielfunktionen ist jedoch nicht Teil dieser Arbeit, sondern erfordert eine systematische, weiterführende Analyse.

5.4 Studie zur Einsatzfähigkeit von asymmetrischen Kostenfunktionen

5.4.1 Gestaltung des Versuchsaufbaus

5.4.1.1 Datenbasis der Studie

Im Anschluss an die theoretische Herleitung betrieblicher Ziel und Kostenfunktionen wird im Folgenden die Fähigkeit von MLPs zur Minimierung von asymmetrischen Kostenfunktionen bei der Zeitreihenprognose überprüft. Dazu wird eine Reihe von Experimenten simuliert und mögliche Veränderungen in der Funktionsweise der NN analysiert.¹ Die Experimente erfolgen auf der Grundlage unterschiedlicher synthetischer Zeitreihen und Entscheidungskosten, die bereits an die Problemstellung der Warendisposition im Handel angelehnt sind.

Um verlässliche Aussagen über die Einsatzfähigkeit von asymmetrischen Kostenfunktionen im Training von MLPs treffen zu können, erfolgt die Beurteilung in einem kontrollierten Experiment auf einer repräsentativen Auswahl synthetischer Zeitreihen mit unterschiedlichen Zeitreihenmustern, Verteilungsfunktionen und Nachfrageschwankungen.² Zur Abbildung unterschiedlicher Wahrscheinlichkeitsfunktionen der Bedarfe werden Zeitreihen mit unterschiedlichen Verteilungen der Zufallszahlen generiert: der Normalverteilung (NV), der Poissonverteilung (PV) und der negativen Binomialverteilung (NB).³ Zur Abbildung unterschiedlicher Nachfrageschwankungen werden für jede Wahrscheinlichkeitsfunktion drei Zeitreihen mit wachsender Streuung gebildet. Die Zeitreihen der Zufallszahlen sind stationär mit unkorrelierter, additiver Streuung⁴ um einen konstanten Mittelwert c (vgl. Formel 3.33). Ziel des MLP im Trainingsprozess ist das Erlernen der Modellstruktur zur Prognose unabhängig von der Intensität oder der Verteilung der Zufallsschwankungen. Im Fall stationärer Zeitreihen ohne Saisonalität entspricht dies der korrekten Trennung der erkennbaren Struktur des Mittelwertes der Zeitreihe $c = \mu$ von den Residuen e_t .⁵ Diese einfache, stationäre Modellform der synthetischen Zeitreihen begünstigt eine direkte Beurteilung der Fähigkeiten von MLPs mit unterschiedlichen Zielfunktionen zunächst ohne deterministische Zeitreihenmuster z. B. einer Saisonalität. In Abbildung 5.15 werden die neun Bedarfsverläufe der drei Wahrscheinlichkeitsfunktionen und der drei Streuungsintensitäten anhand ihrer Histogramme mit überlagerter Wahrscheinlichkeitsdichtefunktion der Normalverteilung visualisiert.

¹ Erste Erkenntnisse aus prototypischen Experimenten wurden vom Autor auf wissenschaftlichen Konferenzen zur Diskussion gestellt, dort wurde die Neuartigkeit der methodischen Ansätze bestätigt. Insbesondere wurde die Funktionsweise asymmetrischer Kostenfunktionen auf einer stationären Zeitreihe eines White-noise-Prozesses [Crone 2002; Crone 2003] sowie auf einer saisonalen Zeitreihe untersucht [Crone et al. 2005]. Die Experimente zu asymmetrischen Kostenfunktionen werden hier erstmalig umfassend aufgearbeitet, und anhand einer repräsentativen Auswahl von Zeitreihen und Zielfunktionen in einem validen und reliablen Versuchsaufbau analysiert und einheitlich präsentiert.

² Dies erlaubt jedoch noch keine Rückschlüsse auf die empirische Einsatzfähigkeit und die Güte von NN mit asymmetrischen Kostenfunktionen in empirischen Anwendungen, die anschließend in Abschnitt 6 auf realen Zeitreihen untersucht werden müssen.

³ Alle Zufallszahlen wurden in SPSS 13.0 unter Verwendung der in Abbildung 5.15 angegebenen Funktionen RV.NORM, RV.NEGBIN und RV.POISSON mit den angegebenen Verteilungsparametern erstellt.

⁴ Die Ergebnisse des Experimentes sind somit nicht beliebig auf Zeitreihenverläufe mit einer multiplikativen Verknüpfung der Streuung übertragbar. Dies kann erst in weiterführenden Experimenten überprüft werden.

⁵ Da keine Ausreißer in synthetischen Zeitreihen vorkommen, entspricht der Mittelwert ebenfalls dem Median, so dass MLPs mit linearen Kostenfunktionen sowohl Mittelwert als auch Median approximieren können.

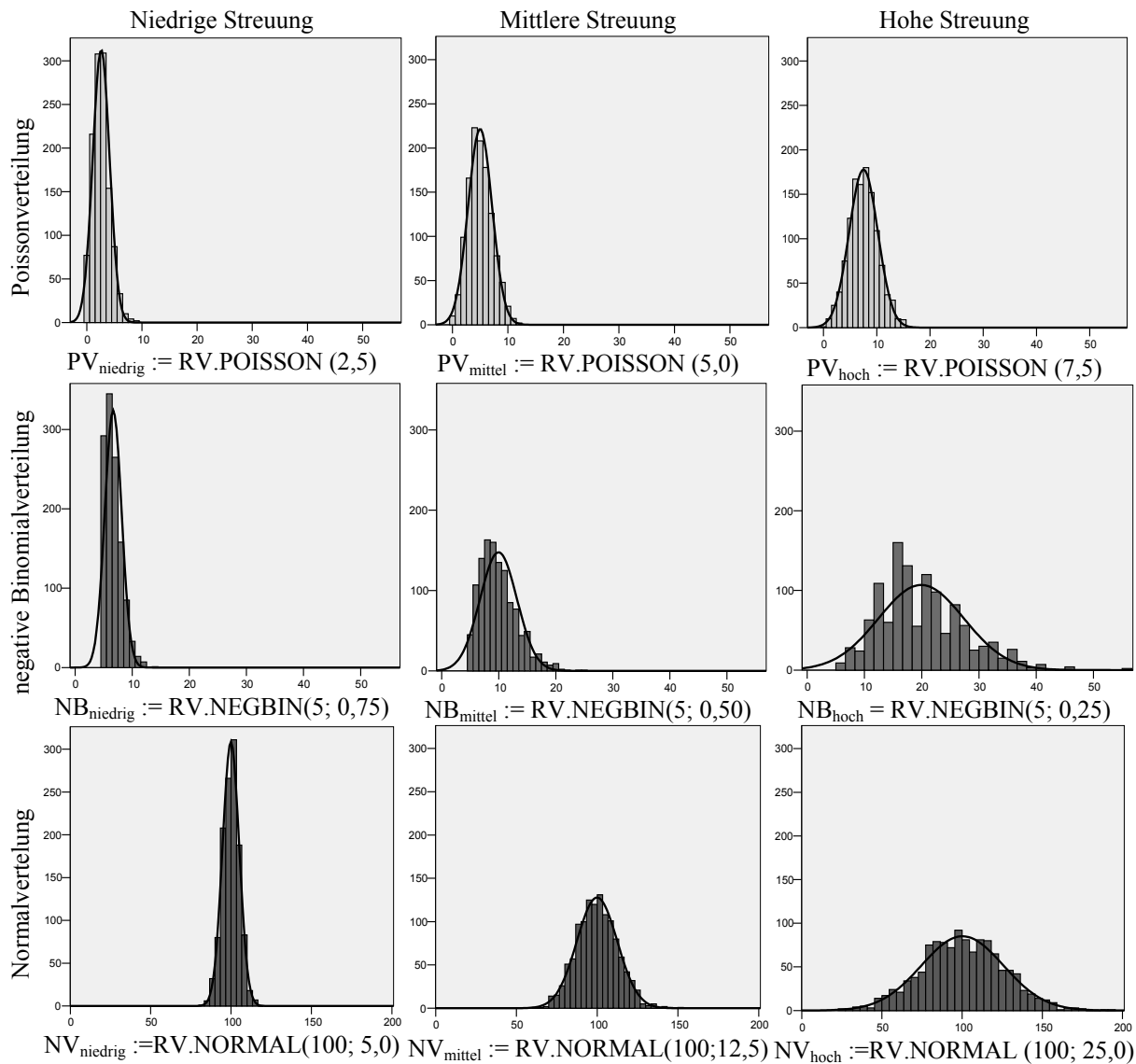


Abbildung 5.15: Histogramme der Wahrscheinlichkeitsverteilungen stationärer Zufallsvariablen mit unterschiedlicher Streuung und überlagerter Darstellung der Dichtefunktion der Normalverteilung

Die Zeitreihen der Normalverteilung $Y \sim N(\mu, \sigma)$ werden für einen konstanten Mittelwert $\mu = 100$ und wachsende Standardabweichungen $\sigma = \{5, 0; 12, 5; 25, 0\}$ erstellt. In Anlehnung an die zunehmende Streuung der Zeitreihen werden sie als NV_{niedrig} , NV_{mittel} und NV_{hoch} bezeichnet. Die Zeitreihen für Artikel mit geringem Absatzvolumen bzw. langsamer Umschlaggeschwindigkeit werden durch die negative Binomialverteilung $Y \sim NB(\mu, \phi)$ mit den Parametern $\mu = 5$ und $\phi = \{0, 25; 0, 50; 0, 75\}$ generiert und als NB_{niedrig} , NB_{mittel} und NB_{hoch} bezeichnet. Die Zeitreihen der Poissonverteilung $Y \sim P(\mu)$ werden durch den Parameter des Mittelwertes $\mu = \{2, 5; 5, 0; 7, 5\}$ erstellt und als PV_{niedrig} , PV_{mittel} und PV_{hoch} bezeichnet. Für jede Zeitreihe werden 1200 Beobachtungswerte als Zufallszahlen generiert, um Probleme der Generalisierbarkeit der Simulations-Ergebnisse auf kurzen Zeitreihen zu umgehen.¹ Der Verlauf der Zeitreihen wird anhand der ersten 120 Beobachtungswerte in Abbildung 5.16 exemplarisch dargestellt.

¹ Ziel des Experimentaufbaus ist somit nicht die Sicherstellung der Repräsentativität dieses Versuchsaufbaus für Zeitreihen des Handels, sondern die grundlegende Überprüfung der Funktionsweise und Güte des Verfahrens auf unterschiedlichen Zeitreihen. Auch bei längeren Testdatensätzen kann immer nur eine Approximation und keine exakte Bestimmung der Kosten und der Servicegrade erzielt werden.

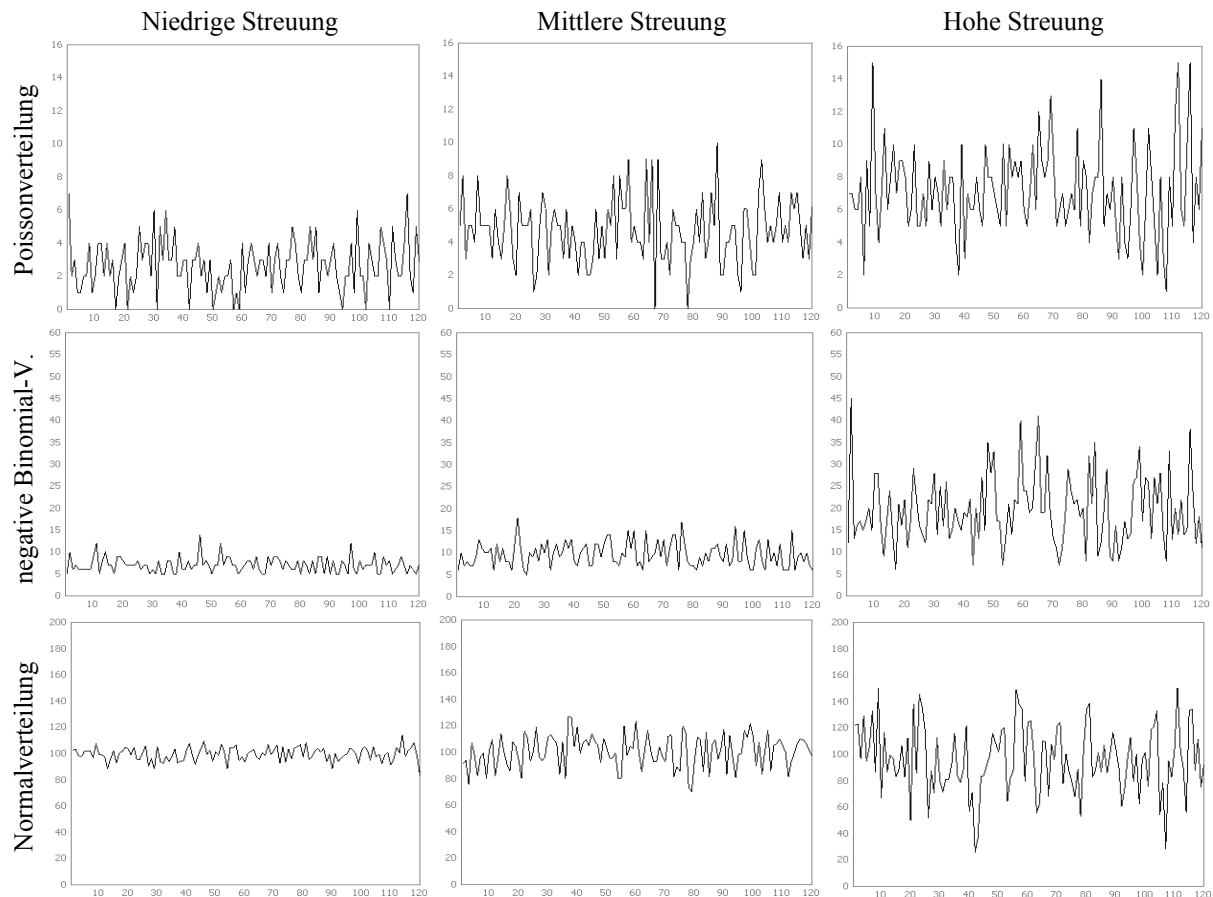


Abbildung 5.16: Ausschnitte der ersten 120 Werte stationärer Zeitreihen mit niedriger, mittlerer und hoher Nachfrageschwankung der Poissonverteilung, negativen Binomialverteilung und der Normalverteilung

Zusätzlich werden diese stationären Zeitreihen mit einem monatlichen saisonalen Index SI_t für $t = 1, 2, \dots, 12$ multipliziert und so neun weitere jährliche Zeitreihen mit monatlichem saisonalem Verlauf erstellt. Diese Zeitreihen sind von der Form

$$y_t = c SI_t + e_t \quad , \quad (5.57)$$

mit den Werten des monatlichen Index von Januar bis Dezember

$$SI_t = \{0,75; 0,80; 0,82; 0,90; 0,94; 0,92; 0,91; 0,99; 0,95; 1,02; 1,20; 1,80\} \quad . \quad (5.58)$$

Der Index ist dem saisonalen Verlauf des aggregierten monatlichen Handelsumsatzes der USA nach ZHANG und QI [2005] nachempfunden und wurde mehrfach in Untersuchungen der Leistungsfähigkeit von MLP zum Zweck der Absatzprognose verwendet [Crone et al. 2006a; Crone et al. 2006b; Crone et al. 2006c; Crone et al. 2006d].¹ Die neun Zeitreihen werden analog zu den stationären Zeitreihen bezeichnet und durch das Präfix ‚S‘ als $S.NV_{\text{niedrig}}$, $S.NB_{\text{niedrig}}$, $S.PV_{\text{hoch}}$ usw. identifiziert. In Abbildung 5.17 sind Ausschnitte der saisonalen Zeitreihen in einer Übersicht dargestellt.

¹ Das saisonale Muster besitzt auch für den aggregierten monatlichen Umsatz im deutschen Einzelhandel Gültigkeit [Statistisches Bundesamt Deutschland 2006]. Der Index zeichnet sich insbesondere durch einen nicht vollständig differenzierbaren Verlauf an der Sprungstelle des Weihnachtsgeschäfts aus. Da auch andere deterministische oder stochastische Verläufe saisonaler Indizes existieren, sind die nachfolgenden Aussagen nicht auf alle Muster saisonaler Zeitreihen verallgemeinerbar.

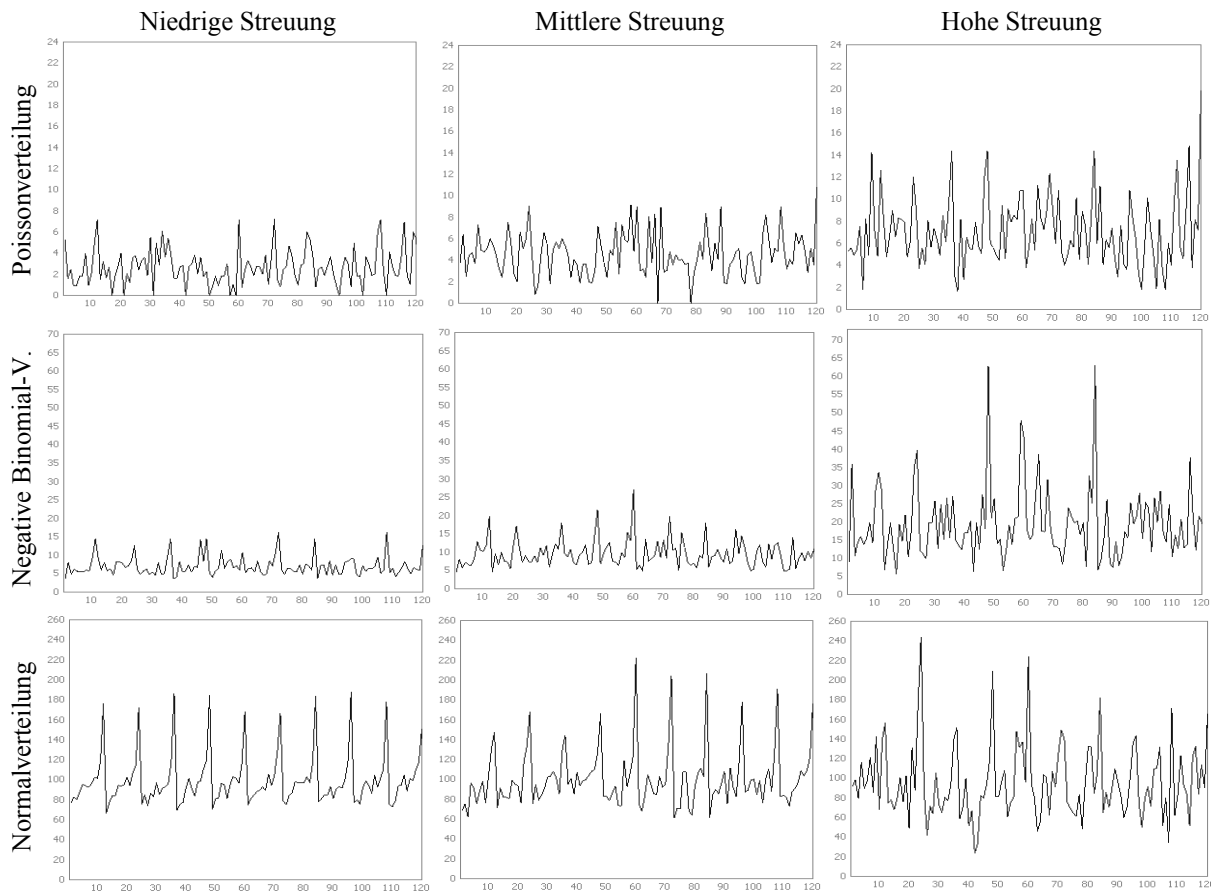


Abbildung 5.17: Ausschnitte der ersten 120 Werte saisonaler Zeitreihen mit niedrigen, mittleren und hohen Nachfrageschwankungen der Poissonverteilung, negativen Binomialverteilung und der Normalverteilung

Wie in vorhergehenden Experimenten gezeigt wurde kann eine derartige, saisonale oder stationäre Zeitreihenstruktur von einem MLP mit quadratischer Zielfunktion erlernt und zur Prognose auf unbekanntem Daten generalisiert werden.¹ Diese Fähigkeit gilt es nachfolgend auch für die Disposition mit asymmetrischen Kostenfunktionen zu beurteilen.

5.4.1.2 Ziel- und Kostenfunktionen

Um die unterschiedlichen Auswirkungen von asymmetrischen Zielfunktionen bestimmen zu können, werden vier betriebswirtschaftliche Kostenfunktionen untersucht, die jeweils mit einem unterschiedlichen Servicelevel der Warendisposition korrespondieren. Die Kostenfunktionen sind von der Form der linearen-linearen Kostenfunktion aus Formel (5.26) bzw. (5.27) mit

$$LLC_{t+h}(e) = \begin{cases} c_u |y_{t+h} - \hat{y}_{t+h}| & \text{für } y_{t+h} > \hat{y}_{t+h} \\ c_o |y_{t+h} - \hat{y}_{t+h}| & \text{für } y_{t+h} \leq \hat{y}_{t+h} \end{cases} \quad (5.59)$$

Unter Anwendung des Newsboy-Modells der einperiodigen Bestellmengenrechnung bestimmt der Grad der Asymmetrie der Entscheidungskosten aus zu niedrigen Prognosen c_u und zu hohen Prognosen c_o den kostenoptimalen Servicelevel. Um die Minimierung unterschiedlicher Zielfunktionen durch das MLP beurteilen zu können, werden vier Zielfunktionen mit abweichenden Kosten c_u , c_o für vier α -Servicegrade gebildet. Die Zielfunktionen umfassen die

¹ Auch wenn die Saisonalität aller Zeitreihen nicht ohne Weiteres aus den Zeitreihengraphen erkennbar ist, bestätigt eine Analyse der absoluten und der partiellen Autokorrelationsfunktionen einen ARIMA(0,0,0)(1,0,0)₁₂- bzw. stationären AR(1)₁₂-Prozess mit signifikanter monatlicher Saisonalität ohne MA-Komponenten.

symmetrische Kostenfunktion des quadratischen Fehlers $SE(e)$ mit $c_u = c_o = 1,00$, die dem Sonderfall einer symmetrischen quadratischen-quadratischen Kostenfunktion entspricht, sowie drei Varianten asymmetrischer Kostenfunktionen des Typs $LINLIN(e)$, die als ACF_1 , ACF_2 und ACF_3 bezeichnet werden und mit $c_u \geq c_o$ einen zunehmenden Grad an Asymmetrie aufweisen.

$$SE : \quad QQC_{t+h}(e) = \begin{cases} 1,00(y_{t+h} - \hat{y}_{t+h})^2 & \text{für } y_{t+h} > \hat{y}_{t+h} \\ 1,00(y_{t+h} - \hat{y}_{t+h})^2 & \text{für } y_{t+h} \leq \hat{y}_{t+h} \end{cases} \quad (5.60)$$

$$F_Y(S^*) \geq \frac{c_u}{c_o + c_u} = \frac{1,00}{1,00 + 1,00} = 0,50$$

$$ACF_1 : \quad LLC_{t+h}(e) = \begin{cases} 1,00|y_{t+h} - \hat{y}_{t+h}| & \text{für } y_{t+h} > \hat{y}_{t+h} \\ 0,17647|y_{t+h} - \hat{y}_{t+h}| & \text{für } y_{t+h} \leq \hat{y}_{t+h} \end{cases} \quad (5.61)$$

$$F_Y(S^*) \geq \frac{c_u}{c_o + c_u} = \frac{1,00}{0,17647 + 1,00} = 0,85$$

$$ACF_2 : \quad LLC_{t+h}(e) = \begin{cases} 1,00|y_{t+h} - \hat{y}_{t+h}| & \text{für } y_{t+h} > \hat{y}_{t+h} \\ 0,052631|y_{t+h} - \hat{y}_{t+h}| & \text{für } y_{t+h} \leq \hat{y}_{t+h} \end{cases} \quad (5.62)$$

$$F_Y(S^*) \geq \frac{c_u}{c_o + c_u} = \frac{1,00}{0,052631 + 1,00} = 0,95$$

$$ACF_3 : \quad LLC_{t+h}(e) = \begin{cases} 1,00|y_{t+h} - \hat{y}_{t+h}| & \text{für } y_{t+h} > \hat{y}_{t+h} \\ 0,01010|y_{t+h} - \hat{y}_{t+h}| & \text{für } y_{t+h} \leq \hat{y}_{t+h} \end{cases} \quad (5.63)$$

$$F_Y(S^*) \geq \frac{c_u}{c_o + c_u} = \frac{1,00}{0,01010 + 1,00} = 0,99$$

Diese Servicegrade gelten als repräsentativ für Problemstellungen der Warendisposition im Handel. Alle Parameter, zugehörige Servicelevel und die Anzahl der k Standardabweichungen zur Berechnung der Sicherheitsbestände im Newsboy-Modell sind in Tabelle 5.5 zusammenfassend dargestellt.

Tabelle 5.5: Parameter der asymmetrischen Zielfunktionen im Experiment

Zielfunktion	Entscheidungskosten	α -Servicelevel	Standardabweichungen
SE	$c_u = 1,00; c_o = 1,000000$	$\alpha_1 = 50,0\%$	$k_1 = 0,000$
ACF_1	$c_u = 1,00; c_o = 0,176470$	$\alpha_2 = 85,0\%$	$k_2 = 1,036$
ACF_2	$c_u = 1,00; c_o = 0,052631$	$\alpha_3 = 95,0\%$	$k_3 = 1,645$
ACF_3	$c_u = 1,00; c_o = 0,010100$	$\alpha_4 = 99,0\%$	$k_4 = 2,327$

5.4.1.3 Modellierung der Verfahren

Die Beurteilung von MLP zur Disposition erfordert wie in jeder Anwendung von NN zunächst die Bestimmung einer problem- und datenspezifischen Netzwerkarchitektur. Um die Komplexität im Versuchsaufbau möglichst gering zu halten, wird eine einheitliche, effiziente

Netzwerkarchitektur mit möglichst geringen Freiheitsgraden in einem Vorexperiment auf den sechs Zeitreihen mittlerer Streuung bestimmt.

Die Modellierung eines MLP erfordert die Bestimmung der Architekturparameter $A = [[N^I - N^{H1}(F^I, F^A, F^O) - N^{H2}(F^I, F^A, F^O) - \dots - N^O(F^I, F^A, F^O)], W, K]$ sowie der Parameter des Lernalgorithmus $L = [[F^N, F^P], G, P_{T,L}, B, S, O]$ einschließlich der Zielfunktion O . Hierzu werden unterschiedliche Modell-Kandidaten erstellt und eine angemessene Architektur schrittweise aus den überlegenen Architekturentscheidungen ausgewählt [Crone 2002]. Die Bestimmung der Anzahl von Eingabeneuronen berücksichtigt zunächst 12 unterschiedliche Eingabevektoren aus zeitlich verzögerten der Beobachtungswerte $N^I = \{t, t-1, t-2, \dots, t-36\}$ in Schritten von drei Neuronen. Neben dieser zeitreihenanalytischen Modellierung werden auch zwei kausalanalytische Varianten unter Bildung von Erklärungszeitreihen der Kalendermonate in Form von zwölf binären Dummy-Zeitreihen $x_{D,1}, x_{D,2}, x_{D,3}, \dots, x_{D,12}$ mit $x_{D,t} = \{0;1\}$ je Monat t sowie eine intervallskalierte Zeitreihe $x_{I,t} = \{1=\text{Januar}, 2=\text{Februar}, \dots, 12=\text{Dezember}\}$ mit korrespondierender Lagstruktur erstellt. Je Struktur der Eingabeschicht werden 11 Varianten von Neuronen der versteckten Schicht $N^{H1} = \{0, 2, 4, \dots, 20\}$ evaluiert. Zur Berücksichtigung unterschiedlicher Formen der Signalverarbeitung werden für jede Topologie zwei Aktivierungsfunktionen der versteckten Neuronen $F^{H1} = \{\text{TanH}, \text{Log}\}$ und drei Aktivierungsfunktionen der Ausgabeneuronen $F^A = \{\text{TanH}, \text{Log}, \text{Id}\}$ evaluiert. Die Netzwerkausgabe erfolgt durch ein einzelnes Ausgabeneuron $N^O = \{1\}$ zur $t+1$ Einschnitt-Prognose für $h=1$. Somit werden insgesamt 2160 Netzwerktopologien von MLPs untersucht.

Jede der 18 Zeitreihen wird gleichmäßig in 400 Beobachtungen zur Parametrisierung der MLP auf dem Trainingsdatensatz, 400 Beobachtungen der Validierungsdaten zum early-stopping des Trainingsprozesses sowie der Selektion der besten Initialisierung und 400 Beobachtungen der Testdaten zur Ex-ante-Beurteilung der Prognosegüte aufgeteilt, um mögliche Ungenauigkeiten und Trainingsprobleme durch das early-stopping und Overfitting durch die Aufteilung der Datenbasis zu vermeiden. Alle Zeitreihen wurden vor dem Training linear in das Intervall $[-0,6; 0,6]$ skaliert, um Sättigungseffekte der nichtlinearen Aktivierungsfunktion zu vermeiden. Anschließend wird jede der Netzwerktopologien 10-fach mit zufälligen Startgewichten initialisiert und mit dem generalisierten Backpropagation-Algorithmus aus Abschnitt 5.3.2.1 für 1.000 Epochen trainiert. Als Zielfunktion des Vorexperimentes dient die Minimierung des quadratischen Fehlers $SE(e)$ im Sinne einer konventionellen Einschnitt-Prognose für $t+1$. Die Lernrate von $\eta = 0,6$ wird bei konstantem Momentum-Term $\alpha = 0,4$ nach jeder Epoche schrittweise um 1% zur Erhöhung des Konvergenzverhaltens abgesenkt. Sofern sich der Zielfunktionswert auf den Validierungsdaten in 100 Epochen des Trainingsprozesses nicht um 0,1% verbessert hat, wird der Lernprozess durch early-stopping abgebrochen und die Gewichte mit dem geringsten Fehler auf den Validierungsdaten gespeichert. Insgesamt werden auf diese Weise je 129000 MLPs in einem balancierten Experimentaufbau von sechs Zeitreihen mittlerer Streuung trainiert.¹

Anschließend wird die Verteilung der mittleren Fehlerwerte auf den Validierungsdaten über alle Zeitreihen für alle Initialisierungen einer Architektur durch eine multifaktorielle Varianzanalyse (ANOVA) analysiert und Architekturvarianten mit statistisch signifikant höheren Prognosefehlern durch eine Untersuchung homogener Subgruppen der „marginal means“ als

¹ Dieses und alle weiteren MLP-Experimente wurden unter Verwendung eines prototypischen Simulators zur Prognose mit NN berechnet, der unter der Leitung des Autors am Institut für Wirtschaftsinformatik der Universität Hamburg entwickelt wurde und der eine Minimierung beliebiger Kosten- und Zielfunktionen erlaubt. Die gesamte Rechenzeit betrug 72 Stunden auf einem Intel-PC mit 3,6-GHz-Prozessor, 4 GB Hauptspeicher und 1000 GB Festplattenspeicher.

suboptimal verworfen.¹ Während Topologien ohne oder mit nur einem Neuron in der versteckten Schicht N^{HI} eine verringerte Prognosegenauigkeit bieten, zeigt die Untersuchung für Topologien $N^{HI} > 2$ keine signifikanten Unterschiede in der Prognosegenauigkeit. Eine Analyse der Länge des Inputvektors identifiziert keine signifikanten Fehlerverbesserungen für mehr als 13 Eingabeneuronen der verzögerten Beobachtungen $t, t-1, \dots, t-12$. Für die drei Varianten der zeitreihenanalytischen versus kausalanalytischen Modellierung mit Erklärungszeitreihen zeigt die univariate Modellierung signifikant höhere Fehlerwerte als die beiden kausalanalytischen Ansätze, zwischen denen keine signifikanten Unterschiede festgestellt werden konnten. Da die Nutzung der intervallskalierten Erklärungszeitreihe x_I bei einer Lagstruktur der Länge 13, mit $x_{I,t}, x_{I,t-1}, x_{I,t-2}, \dots, x_{I,t-12}$, lediglich 13 weitere Eingabeneuronen erfordert, im Gegensatz zu 156 zusätzlichen Eingabeneuronen bei binärer Modellierung, wird die sparsamere Form der Modellierung gewählt.² Eine Analyse der „marginal means“ der Aktivierungsfunktionen zeigt durchschnittlich höhere Fehler der logistischen Funktion in der Ausgabeschicht sowie einen negativen Interaktionseffekt zwischen der TanH-Funktion in der Ausgabeschicht und anderen Aktivierungsfunktionen in der versteckten Schicht. Dagegen bestehen keine signifikanten Unterschiede der Prognosegüte zwischen Topologien mit TanH in der verdeckten bzw. Ausgabeschicht sowie der logistischen Funktion in der versteckten Schicht und der Identitätsfunktion in der Ausgabeschicht.

Aufgrund dieser Ergebnisse wird eine sparsame Topologie aus einem Ausgabeneuron, zwei versteckten Neuronen und 26 Eingabeneuronen für 13 verzögerte Realisationen der Beobachtungswerte $y_t, y_{t-1}, y_{t-2}, \dots, y_{t-12}$ und 13 Variablen der intervallskalierten Erklärungszeitreihe $x_{I,t}, x_{I,t-1}, x_{I,t-2}, \dots, x_{I,t-12}$ ermittelt. Alle Neuronen der versteckten Schicht verwenden die Summe als Eingabefunktion, die logistische Aktivierungsfunktion und die Identität als Ausgabefunktion. Das Ausgabeneuron verwendet die Identität als Aktivierungs- und Ausgabefunktion. Zusätzlich modelliert ein Bias-Neuron die Schwellwerte aller Neuronen in der versteckten und der Ausgabeschicht. Das MLP ist vollständig verbunden ohne Shortcut-Verbindungen, mit $A=[[26-2(\text{Sum}, \text{Log}, \text{Id})-1(\text{Sum}, \text{Id}, \text{Id})], \text{FF}, \text{F}]$. Somit besitzt das MLP insgesamt 57 freie Parameter aus Verbindungsgewichten und Schwellwerten, welche im Vergleich zu den 387 Trainingsdaten und 400 Validierungsdaten die Gefahr des Überlernens als unbedenklich erscheinen lässt.

Diese einheitliche Netzwerkarchitektur wird anschließend mit den vier unterschiedlichen Zielfunktionen SE, ACF_1, ACF_2 und ACF_3 trainiert. Die entstehenden Gruppen von MLPs werden in Anlehnung an die verwendete Zielfunktion als $\text{MLP}.SE, \text{MLP}.ACF_1, \text{MLP}.ACF_2$ und $\text{MLP}.ACF_3$ bezeichnet. Der balancierte Versuchsaufbau betrachtet somit die Auswirkungen von vier asymmetrischen Kostenfunktionen auf 18 Zeitreihen, davon je 9 stationäre und je 9 saisonale Zeitreihen sowie je 6 mit normalverteiltem, poissonverteiltem und negativ-binomialverteiltem Bedarf. Das MLP wird zum Training 10-fach mit zufälligen Startgewichten auf jeder der 18 skalierten Zeitreihen initialisiert und mit Hilfe des generalisierten Backpropagation-Algorithmus zur Minimierung der vier Kostenfunktionen zur Einschnitt-Prognose trainiert. Zur Vergleichbarkeit werden dieselbe Datenaufteilung, dieselbe Datenska-

¹ Aufgrund der Platzbeschränkung und der Ähnlichkeit zu vorangegangenen Darstellungen wird hier auf eine detaillierte Darstellung der Experimente und Ergebnisse verzichtet.

² Hierbei ist anzumerken, dass durch das schrittweise Vorgehen eine Lagstruktur von $t-1, \dots, t-13$ identifiziert wurde, obwohl die synthetische Struktur der Zeitreihen und die Analyse der ACF- und PACF-Funktionen der saisonalen Zeitreihen einen deterministischen AR(12)-Prozess identifiziert. Darüber hinaus ist die Verringerung der Prognosefehler durch die intervallskalierte Zeitreihe der Saisonindizes auf erhöhte Schwankungen zurückzuführen, die bei einer kurzen Lagstruktur von nur einer Saison hohe Prognosefehler verursachen könnten. Alternativ könnte auch eine deutliche Verlängerung der Lagstruktur unter Fortfall der erklärenden Zeitreihe vergleichbare Resultate für saisonale Zeitreihen erzielen, jedoch erhöhte Schwankungen bei stationären Zeitreihen hervorrufen. Eine solche Untersuchung ist jedoch nicht Ziel dieser Arbeit, weshalb dieser Aspekt hier nicht vertieft werden soll.

lierung und dieselben Lernparameter wie beim Vorexperiment verwendet. Nach dem Training aller 1.800 MLP-Varianten wird jeweils das MLP mit dem niedrigsten Zielfunktionswert auf der Validierungsmenge ausgewählt. Somit wird das MLP.*SE* mit dem niedrigsten *SE* auf der Validierungsmenge ausgewählt, das MLP.*ACF*₁ mit dem niedrigsten Kosten nach *ACF*₁ auf der Validierungsmenge usw.¹ Dies entspricht einem Quasi-ex-ante-Verfahrensaufbau mit einer einfachen Kreuzvalidierung (engl. *single cross-validation*) auf einer Hold-out-Menge. Anschließend werden die ausgewählten Verfahrens-Kandidaten anhand von statistischen Kennziffern auf den Daten der Trainings-, Validierungs- und Testmenge beurteilt.

Abschließend werden die vier Gruppen der MLP hinsichtlich ihres Systemverhaltens untersucht. Für eine Bestimmung der grundlegenden Fähigkeit von MLPs zur erfolgreichen Minimierung asymmetrischer Kostenfunktionen der Warendisposition werden die In-sample-Gütemaße der Trainings- und Validierungsmenge verwendet. Für die Beurteilung der Fähigkeit zur Generalisierung auf unbekanntem Daten, die im Hinblick auf die Prognose und Disposition für zukünftige Perioden von eigentlichem Interesse ist, werden ausschließlich Gütemaße auf der Testmenge verwendet, die zu keinem Zeitpunkt zum Training oder zur Auswahl einzelner Modelle verwendet wurden.² Der Versuchsaufbau garantiert so valide und reliable Ergebnisse. Da die Zielfunktion bei identischen Startgewichten und Trainingsparametern über alle Netzwerkvarianten die einzige Veränderung der Netzwerkarchitektur darstellt, ist eine Veränderung der Netzwerkausgabe und des Netzwerkverhaltens der MLP-Varianten MLP.*SE*, MLP.*ACF*₁, MLP.*ACF*₂ und MLP.*ACF*₃ allein auf die Asymmetrie der Kostenfunktionen zurückzuführen.

5.4.2 Ergebnisse

5.4.2.1 Grafische Analyse und Interpretation

Zur Interpretation des abweichenden Systemverhaltens von MLPs mit asymmetrischen Zielfunktionen und zum weiteren Verständnis der Ergebnisse erscheint eine grafische Abbildung als zeitreihengraphen vorteilhaft. Aus Platzgründen wird auf eine Darstellung aller 72 Kombinationen von Zeitreihen und Zielfunktionen verzichtet. Die Verhaltensweise und Robustheit der Vorhersage von MLPs zur Disposition wird exemplarisch anhand der stationären und der saisonalen Zeitreihen der Normalverteilung dargestellt.³

Zunächst werden die Prognosen eines MLP.*SE* mit der konventionellen quadratischen Zielfunktion *SE* grafisch dargestellt. In Abbildung 5.18 wird ein Ausschnitt der drei normalverteilten Zeitreihen NV_{Niedrig} , NV_{Mittel} und NV_{Hoch} mit niedriger, mittlerer und hoher Streuung einschließlich der Prognosewerte des besten MLP.*SE* der jeweiligen Zeitreihe dargestellt.⁴

¹ Dabei ist anzumerken, dass die MLPs keine gerundeten Werte ausgeben und so insbesondere für die Zeitreihen von Zähldaten der Poissonverteilung Rundungsfehler entstehen können. In ersten Experimenten konnten verbesserte Ergebnisse durch Rundung der Netzwerkausgabe im Trainingsprozess bei der Fehlerberechnung und in der Rückwärtspropagierung erzielt werden. Sie erfordern eine eigenständige Untersuchung, die nicht zu den Erkenntnissen asymmetrischer Zielfunktionen beiträgt und somit unterbleibt.

² Die Ergebnisse der Testmenge wurden auch nicht in das Vorexperiment eingebunden. Alle vorangegangenen Entscheidungen beruhen einzig auf Gütekriterien der Trainings- und der Validierungsmenge.

³ Die Güte aller Kombinationen von Zeitreihen und Zielfunktionen wird im nachfolgenden Abschnitt anhand der numerischen Kosten- und Fehlerwerte diskutiert.

⁴ Eine Abbildung der gesamten Zeitreihe wäre aufgrund ihrer Länge zu unübersichtlich. Der Ausschnitt der Zeitreihe zeigt links die Daten der Validierungsmenge von Beobachtung 740 bis 800 mit hellblauem Hintergrund und rechts die Daten der Testmenge von 801 bis 860 mit hellgrünem Hintergrund, um eine Beurteilung der Approximationsgüte und der Generalisierungsfähigkeit der NN zuzulassen. Die Zeitreihe der Beobachtungswerte ist in grau gehalten, die zugehörige Einschnitt-Prognose des MLP auf den Validierungsdaten in blau und jene auf den Testdaten in grün.

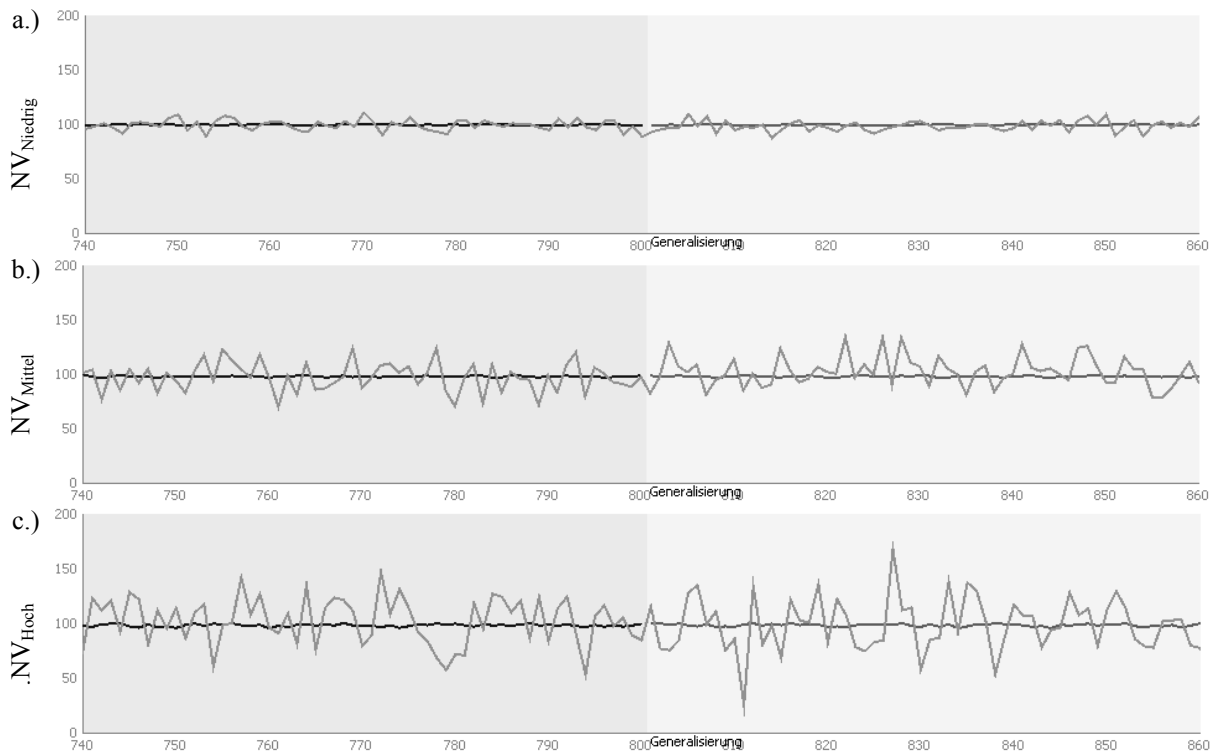


Abbildung 5.18: Originalzeitreihe und Prognosen des MLP.SE auf den stationären Zeitreihen NV_{niedrig} (a.), NV_{mittel} (b.) und NV_{hoch} (c.)

Aufgrund der identischen Verteilung der Kosten von Überprognose und Unterprognose $c_u = c_o = 1,00$ sagt das MLP.SE valide und reliabel den tatsächlichen Mittelwert der Normalverteilung für unterschiedliche Niveaus der Streuung voraus. Da die stationäre Zeitreihe lediglich aus zufälligen Schwankungen um einen konstanten Mittelwert besteht, trennt das MLP die einzige erkennbare Struktur der Zeitreihe $b_0 = \mu = 100$ korrekt von den Residuen e_t mit $e_t \sim N(0, \sigma)$ und $\sigma = \{5, 0; 12, 5; 25, 0\}$.

Trotz einer validen Approximation des zugrunde liegenden Modells zeigt das MLP.SE leichte Tendenzen zum Überlernen wie an der Schwankung der Prognosewerte auf der Validierungs- und Testmenge der Zeitreihe hoher Bedarfsschwankungen NV_{hoch} zu erkennen ist. Dies bestätigt sich bei Betrachtung der Streudiagramme in Abbildung 5.19, in denen die Wertepaare der Prognosewerte des MLP auf Trainings-, Validierungs- und Testmenge an der Abszisse allen Beobachtungswerten an der Ordinate graphisch gegenübergestellt werden.¹ Fehlerfreie Prognosen liegen aufgrund der Übereinstimmung von Beobachtungswert y_t und Prognosewert \hat{y}_t entlang der 45°-Linie, mit $y_t = \hat{y}_t$. Die Streuung der Punktwolke in vertikaler Richtung reflektiert die Varianz der Zeitreihe, während die Variation in horizontaler Richtung die Abweichung der Netzwerkausgabe vom wahren Mittelwert im Sinne eines Prognosefehlers zeigt.² Die Diagramme in Abbildung 5.19 zeigen, dass das MLP für alle Beobachtungswerte auf Trainings-, Validierungs- und Testmenge robuste Werte um den wahren Mittelwert von $\hat{y}_t \approx \mu = 100,0$ prognostiziert und nicht etwa die unkorrelierten Zufallsschwankungen auf der Trainings- und Validierungsmenge durch Überlernen auswen-

¹ Die Streudiagramme geben somit Aufschluss über die Prognosegüte über alle 1.200 Beobachtungswerte der Zeitreihen, während die Darstellung der Zeitreihe aus Gründen der Übersichtlichkeit nur einen Ausschnitt wiedergibt. Sie dienen somit zur Verifizierung der Erkenntnisse für alle Punkte der Zeitreihe.

² Bei einer fehlerfreien Prognose stimmen Beobachtungswert und Prognosewert überein, und es erfolgt ihre Abbildung entlang der 45°-Linie. Dies ist hier jedoch so nicht zu erwarten, da über den Mittelwert hinaus keine prognostizierbaren Informationen vorliegen, sondern lediglich zufälliges Rauschen zunehmender Intensität. Somit stellt die Abbildung des MLP eine gute Prognose dar. Korrespondierend mit der Darstellung der Zeitreihen zeigen blaue Punkte Trainings- und Validierungsdaten und grüne Punkte Testdaten.

dig gelernt hat. Das MLP findet somit eine valide und reliable Approximation und Prognose über alle Bereiche der Zeitreihen.

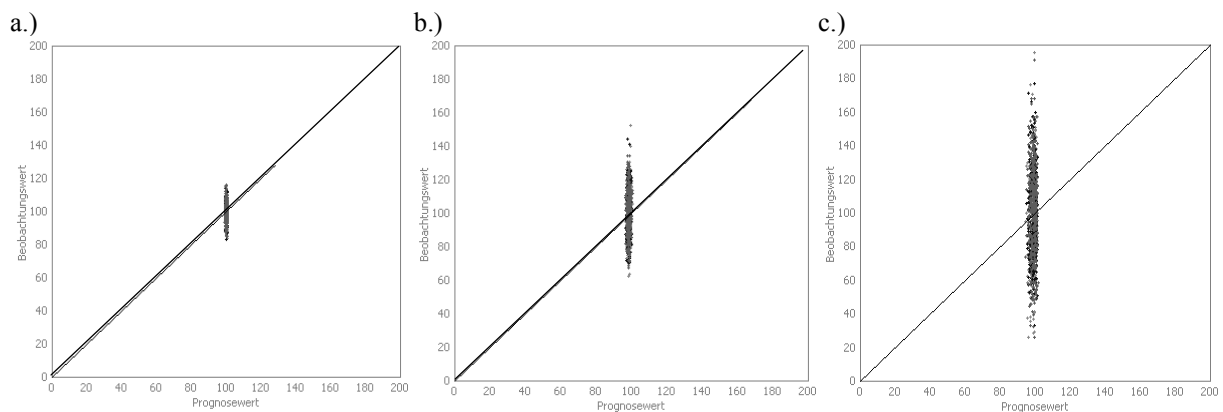


Abbildung 5.19: Streudiagramme der Prognosewerte auf Trainings-, Validierungs- und Testmenge durch MLP.SE auf den Zeitreihen NV_{niedrig} (a.), NV_{mittel} (b.) und NV_{hoch} (c.)

Ein ähnliches Bild zeigt sich für die Analyse des MLP.SE auf den drei saisonalen Zeitreihen mit normalverteilter Streuung $S.NV_{\text{niedrig}}$, $S.NV_{\text{mittel}}$, $S.NV_{\text{hoch}}$ in Abbildung 5.20.

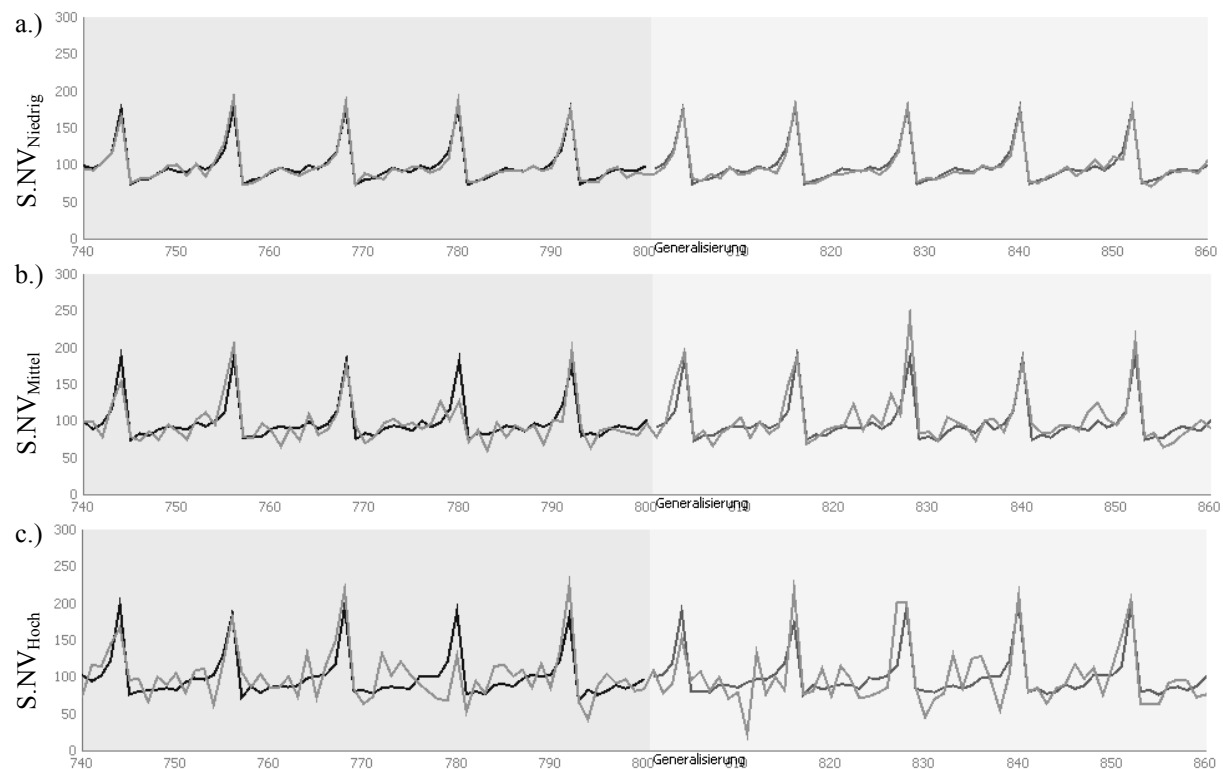


Abbildung 5.20: Originalzeitreihe und Prognosen des MLP.SE auf den saisonalen Zeitreihen $S.NV_{\text{niedrig}}$ (a.), $S.NV_{\text{mittel}}$ (b.) und $S.NV_{\text{hoch}}$ (c.)

Das MLP.SE erkennt Niveau und saisonales Muster der Zeitreihe auf den Trainingsdaten und prognostiziert es valide und reliabel auf den Testdaten für alle drei Intensitätsgrade der Streuung. Auch hier trennt das MLP die Struktur der Zeitreihe von den Zufallsschwankungen. Die Genauigkeit der Vorhersage in Relation zur Streuung der Zeitreihe wird in den Streudiagrammen in Abbildung 5.21 durch Punktwolken entlang der 45° -Linie visualisiert.

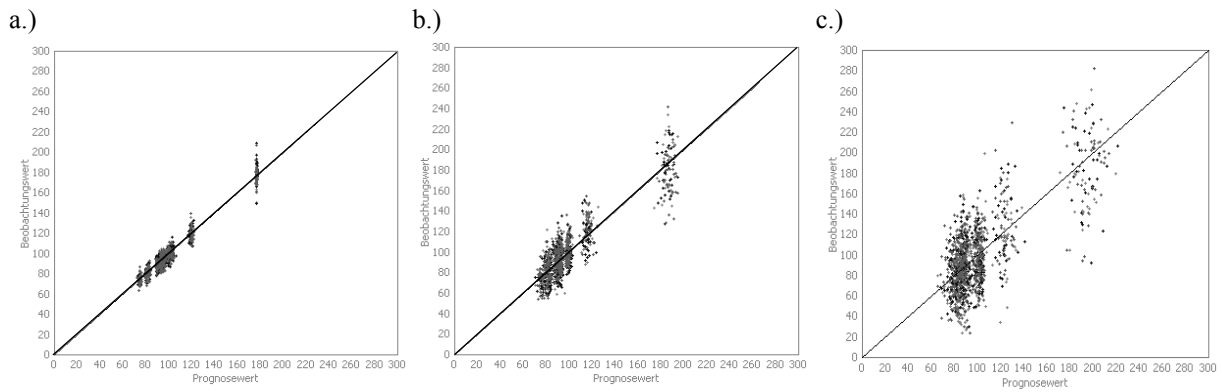


Abbildung 5.21: Streudiagramme der Prognosewerte auf Trainings-, Validierungs- und Testmenge durch MLP.SE auf den Zeitreihen S.NV_{niedrig} (a.), S.NV_{mittel} (b.) und S.NV_{hoch} (c.)

Die Streuung der Punktwolke in vertikaler Richtung reflektiert erneut die inhärente Varianz der Zufallswerte der Zeitreihe, während die Breite der Punktwolke in horizontaler Richtung den Grad der Ungenauigkeit der Prognosen identifiziert. Die Diagramme bestätigen, dass das MLP für alle Beobachtungswerte auf Trainings-, Validierungs- und Testmenge die saisonale Struktur der Zeitreihe erlernt und korrekt in die Zukunft prognostiziert.

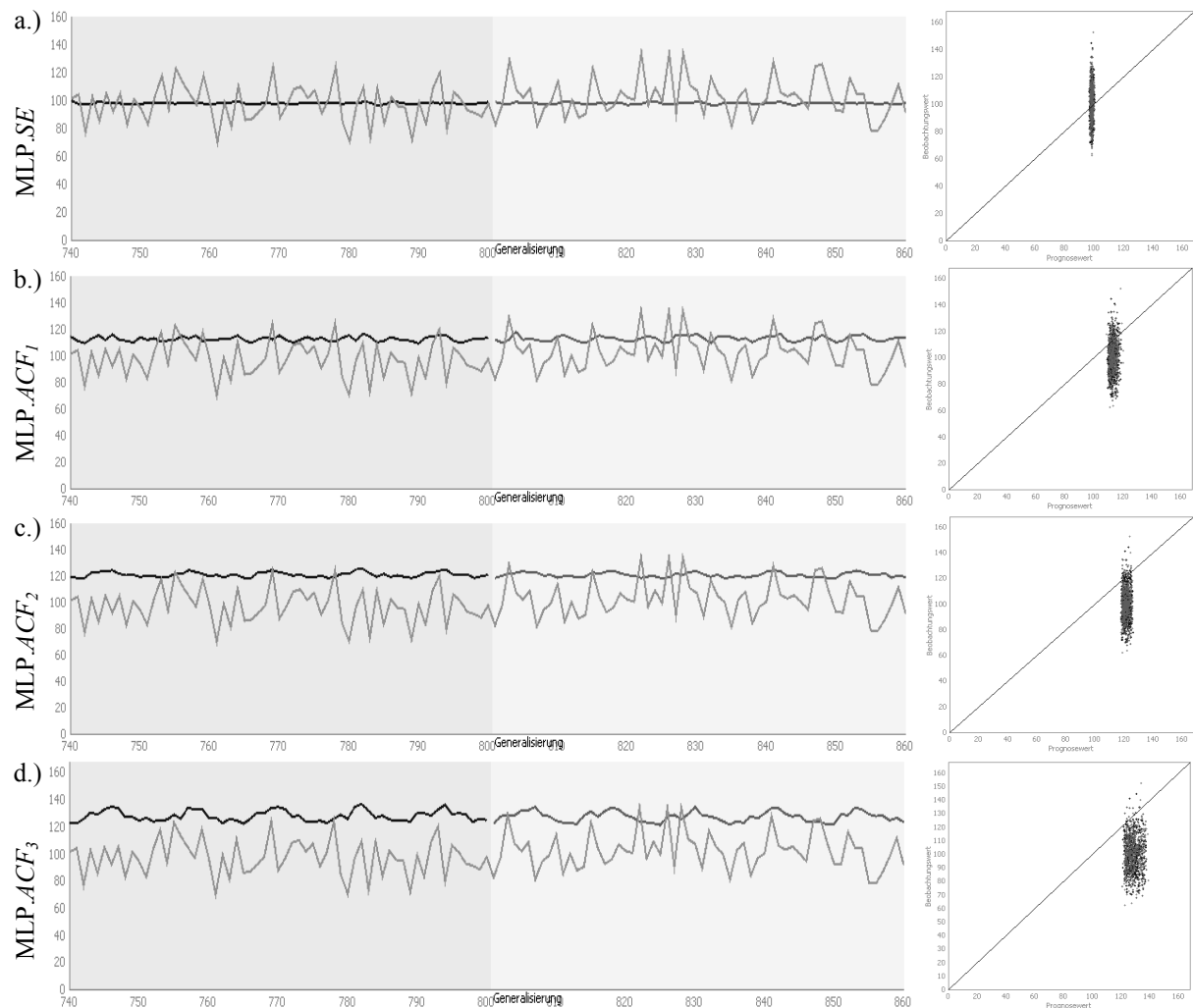


Abbildung 5.22: Originalzeitreihe und Prognosen des MLP.SE (a.), des MLP.ACF₁ (b.), des MLP.ACF₂ (c.) und des MLP.ACF₃ (d.) auf der stationären Zeitreihe NV_{mittel}

Im Gegensatz zur Betrachtung unterschiedlicher Zeitreihen und einer Zielfunktion wird nachfolgend die Auswirkung unterschiedlicher Zielfunktionen bei einer konstanten Zeitreihe

betrachtet. In Abbildung 2.7 wird die Ausgabe von $MLP.SE$, $MLP.ACF_1$, $MLP.ACF_2$ und $MLP.ACF_3$ auf der normalverteilten Zeitreihe mittlerer Streuung NV_{mittel} visualisiert.

Die Abbildung 5.22 verdeutlicht, dass sich die Netzwerkausgabe des MLP in Abhängigkeit von der Zielfunktion SE in Abbildung 2.7.a über die Zielfunktionen ACF_1 und ACF_2 bis auf das Niveau von ACF_3 Abbildung 2.7.d verschiebt. Das MLP erhöht mit steigender Asymmetrie der Zielfunktion sukzessive das Niveau seiner Prognose und sagt so höhere Werte voraus. Dieses veränderte Verhalten wird in den Streudiagrammen bestätigt, die bei gleichen Beobachtungswerten eine Erhöhung des mittleren Prognosewertes von $\hat{y}_t \approx 100$ durch eine Verschiebung der Punktwolken nach rechts zeigen.¹ Während des Lernprozesses kann dementsprechend beobachtet werden, dass ein MLP seine Ausgabe im Training dynamisch erhöht, bis ein konstantes Ausgabeniveau gefunden ist. Diese Ausgabe entspricht nicht länger dem Mittelwert der Wahrscheinlichkeitsfunktion, sondern einem Quantil der Verteilungsfunktion, das durch die Dispositionskosten der Zielfunktion spezifiziert wurde.² Somit verändert das MLP seine Ausgabe von einer Prognose des Erwartungswertes hin zu einer Vorhersage des kostenminimalen Warenbestandes für die zugehörige Bedarfsstruktur als spezifischer Punkt auf der Verteilungsfunktion. Das abweichende Verhalten zeigt sich konsistent über Trainings-, Validierungs- und Testdaten, so dass neben der Fähigkeit der NN zum Lernen aus asymmetrischen Fehlerkosten auch die Fähigkeit der NN zur Generalisierung des gelernten Modells auf den unbekanntem Daten der Testmenge gezeigt wird.³

Im inhaltlichen Zusammenhang der Warendisposition kann das Verhalten des MLP sachlogisch erläutert werden. Das MLP versucht in Abhängigkeit von den Dispositionskosten das Auftreten teurerer Fehlbedarfe durch die Prognose höherer Mengen zu vermeiden und passt so die Netzwerkausgabe an die asymmetrische Kostenfunktion an. Es kompensiert die steigenden Kosten c_u des Risikos einer Unterprognose durch steigende Netzwerkausgaben, korrespondierend mit einer Berücksichtigung zunehmender Sicherheitsbestände. Dies entspricht einer signifikanten Abweichung des konventionellen Verhaltens von MLPs in der Bedarfsprognose hin zu einer Berechnung kostenminimaler Bestandsmengen einer zukünftigen Dispositionsperiode. In diesem Sinne entspricht das Verhalten des MLPs einer simultanen Bedarfsrechnung und Bestellmengenrechnung.

Die Betrachtung der unterschiedlichen Zielfunktionen SE , ACF_1 , ACF_2 und ACF_3 auf einer Zeitreihe mit saisonalem Verlauf und normalverteilten Zufallschwankungen mittlerer Intensität $S.NV_{\text{mittel}}$ in Abbildung 5.23. zeigt ein vergleichbares Verhalten.

¹ Die Streudiagramme zeigen auch eine Zunahme der horizontalen Streuung der Netzwerkausgaben bei konstanter Varianz der Zeitreihe von der Zielfunktion SE auf das Niveau von ACF_1 und ACF_2 sowie weiter zu ACF_3 . Dies ist auf die abweichende Lernaufgabe zurückzuführen, da asymmetrische Kostenfunktionen nicht auf eine Übereinstimmung von Beobachtungswert und Prognosewert der Netzwerkausgabe abzielen, sondern auf ein abweichendes Verhalten des Verfahrens. Dies verbietet die herkömmliche Interpretation der Prognose- bzw. Dispositionsgüte anhand der Streudiagramme für asymmetrische Kostenfunktionen.

² Die Genauigkeit der Vorhersage eines spezifischen Quantils wird im nachfolgenden Abschnitt anhand der Servicegrade betrachtet. Sie kann nicht anhand der grafischen Abbildung verifiziert werden.

³ Weitere Inhomogenitäten und Ungenauigkeiten durch die Verwendung einer einzelnen, für einzelne Muster unter Umständen suboptimalen Netzwerkarchitektur können nicht ausgeschlossen werden. Dies erfordert jedoch weiterführende Untersuchungen, die hier nicht unternommen werden können.

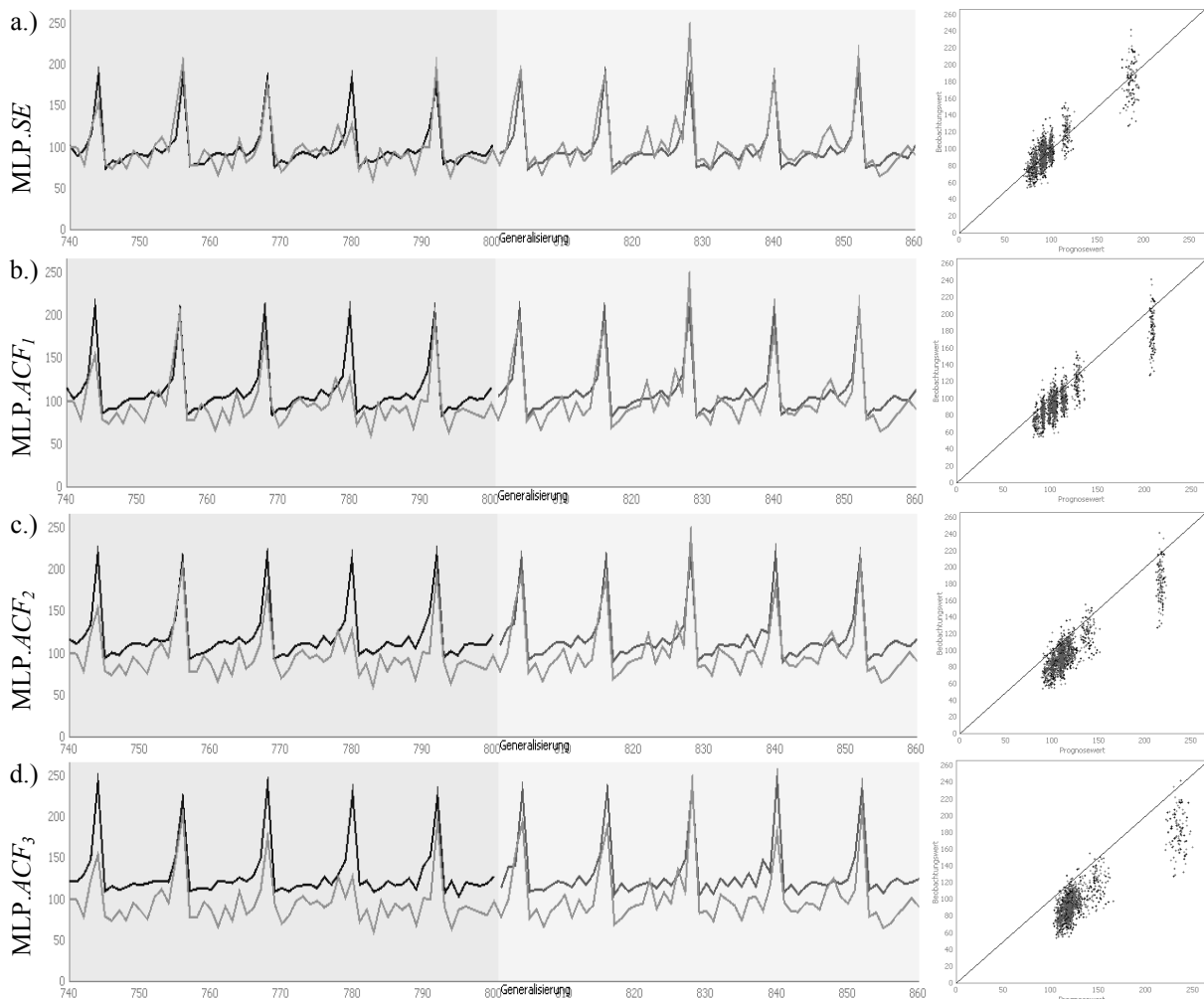


Abbildung 5.23: Originalzeitreihe und Prognosen des MLP.*SE* (a.), des MLP.*ACF*₁ (b.), des MLP.*ACF*₂ (c.) und des MLP.*ACF*₃ (d.) auf der saisonalen Zeitreihe $S.NV_{\text{mittel}}$

Das MLP erfasst das saisonale Muster der Zeitreihe und erhöht in Abhängigkeit von der Zielfunktion mit steigender Asymmetrie der Entscheidungskosten sukzessive das Niveau seiner Prognoseausgabe von *SE* über *ACF*₁ und *ACF*₂ bis auf *ACF*₃. Es demonstriert somit ein konsistentes Verhalten hinsichtlich der Minimierung und Generalisierung unterschiedlicher asymmetrischer Zielfunktionen auch bei Vorliegen eines autoregressiven Musters der Saisonalität. Die Konsistenz dieses Verhaltens wird ebenfalls in den Streudiagrammen aller Beobachtungs- und Prognosewerte in Abbildung 5.23 bestätigt. Erneut zeigt das MLP eine signifikante Abweichung der Netzwerkausgabe von der einer konventionellen Prognose hin zu einer Bestimmung zukünftiger kostenminimaler Warenbestände.

Zur Untersuchung der Interaktion einer asymmetrischen Kostenfunktion mit den Eigenschaften der Zeitreihen wird nachfolgend die Reaktion der Netzwerkausgabe eines MLP mit konstanter, asymmetrischer Kostenfunktion auf unterschiedlichen Zeitreihen zunehmender Streuung analysiert. In Abbildung 5.24 ist die Zeitreihe der Beobachtungswerte und der Prognosen des MLP.*ACF*₃ dargestellt, welches auf drei Zeitreihen normalverteilter Zufallsschwankungen NV_{niedrig} (a.), NV_{mittel} (b.) und NV_{hoch} (c.) trainiert wurde.

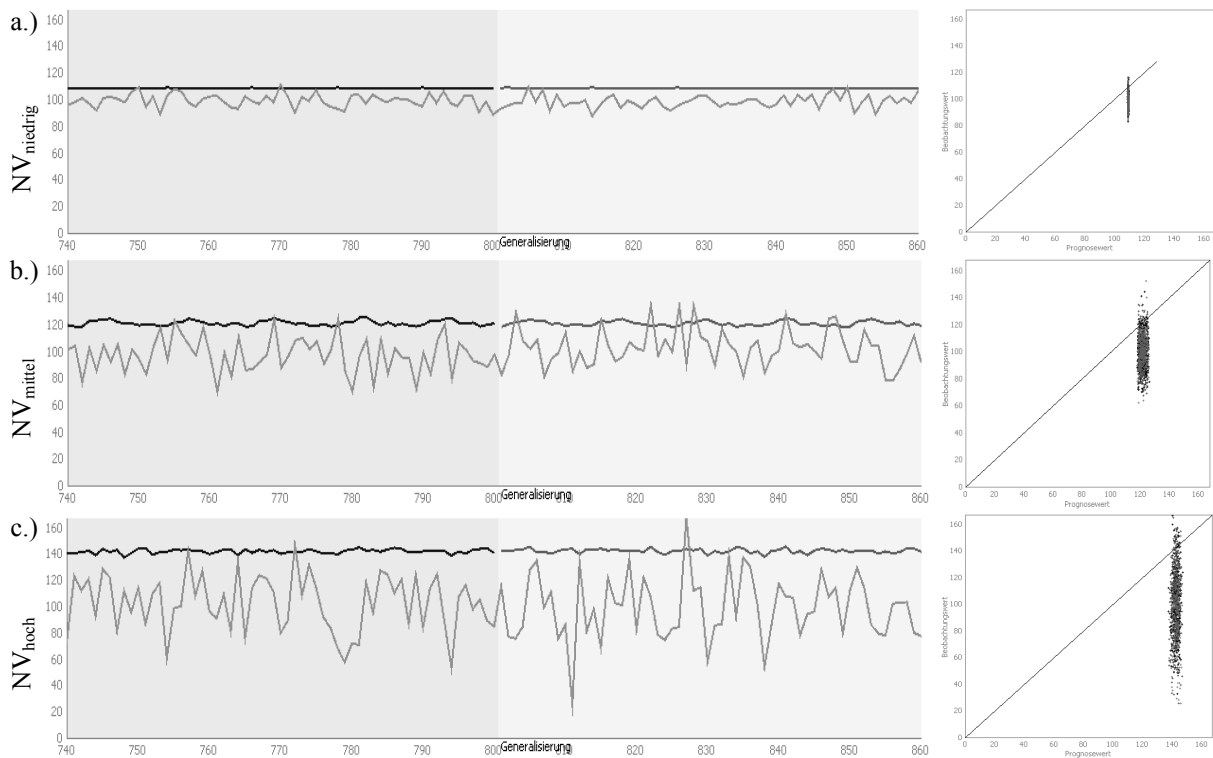


Abbildung 5.24: Originalzeitreihe und Prognose des $MLP.ACF_3$ auf den stationären Zeitreihen NV_{niedrig} (a.), NV_{mittel} (b.) und NV_{hoch} (c.)

Für eine gegebene asymmetrische Kostenfunktion ACF_3 hebt das $MLP.ACF_3$ das Niveau seiner Vorhersage schrittweise mit der zunehmenden Streuung der Zeitreihen an und erzielt so auf unterschiedlichen Bedarfsverteilungen einen konstanten Servicegrad. Ein vergleichbares Bild zeigt sich für ein $MLP.ACF_3$ auf saisonalen Zeitreihen mit zunehmenden Zufallsschwankungen $S.NV_{\text{niedrig}}$ (a.), $S.NV_{\text{mittel}}$ (b.) und $S.NV_{\text{hoch}}$ in Abbildung 5.25.

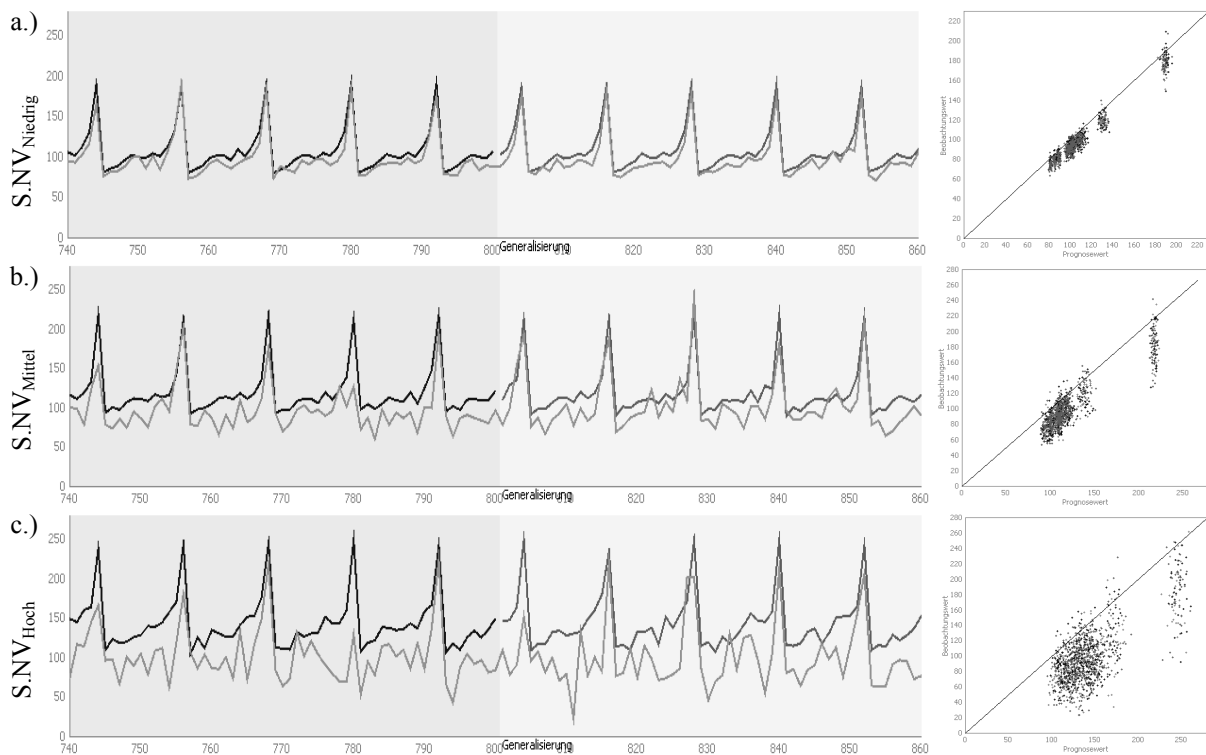


Abbildung 5.25: Originalzeitreihe und Prognosen des $MLP.ACF_3$ auf den saisonalen Zeitreihen $S.NV_{\text{Niedrig}}$ (a.), $S.NV_{\text{Mittel}}$ (b.) und $S.NV_{\text{Hoch}}$ (c.)

Erneut erkennt das MLP die saisonale Zeitreihenstruktur und hebt seine Netzwerkausgabe mit zunehmender Streuung der Zeitreihen von $S.NV_{\text{niedrig}}$ auf $S.NV_{\text{mittel}}$ und auf $S.NV_{\text{hoch}}$ an. Es zeigt sich, dass ein MLP mit einer asymmetrischen Zielfunktion einen konstanten, kostenminimalen Servicelevel unabhängig von dem Grad der Zufallsschwankungen in den Beobachtungswerten approximiert. Als Reaktion auf eine steigende Streuung erhöht das MLP in Abbildung 5.25 das Niveau seiner Vorhersage und vermeidet so teurere Fehlbestände.

Eine grafische Analyse aller stationären und saisonalen Zeitreihen der Normalverteilung anhand der Streudiagramme in Abbildung 5.26 bestätigt die oben gewonnenen Erkenntnisse. Die Analyse der Zeitreihen der Negativen Binomialverteilung und der Poissonverteilung bestätigt die Erkenntnisse der Normalverteilung verteilungsübergreifend. Da ihre Analyse nur einen geringen zusätzlichen Informationsgewinn zum weiteren Verständnis des abweichenden Verhaltens von MLPs beiträgt, findet sich die Darstellung im Anhang. Die Konsistenz aller Ergebnisse wird durch die Analyse der numerischen Fehler- und Kostenwerte in Abschnitt 5.4.2 belegt.

Zusammenfassend kann festgestellt werden, dass ein MLP seine Netzwerkausgabe mit dem Grad der Asymmetrie der Kostenfunktion und der Intensität der Streuung einer Zeitreihe verändert, um einen kostenminimalen Warenbestand vorherzusagen. Dies entspricht der Vorgehensweise konventioneller Modelle zur Berechnung von Sicherheitsbeständen in Abhängigkeit von der Varianz der Prognoseressiduen, z. B. im einperiodigen Newsboy-Modell. Somit kann belegt werden, dass ein MLP mit unveränderter Architektur in der Lage ist, allein durch Anwendung unterschiedlicher asymmetrischer Kostenfunktionen ein abweichendes Systemverhalten hinsichtlich der Warendisposition zu erlernen.

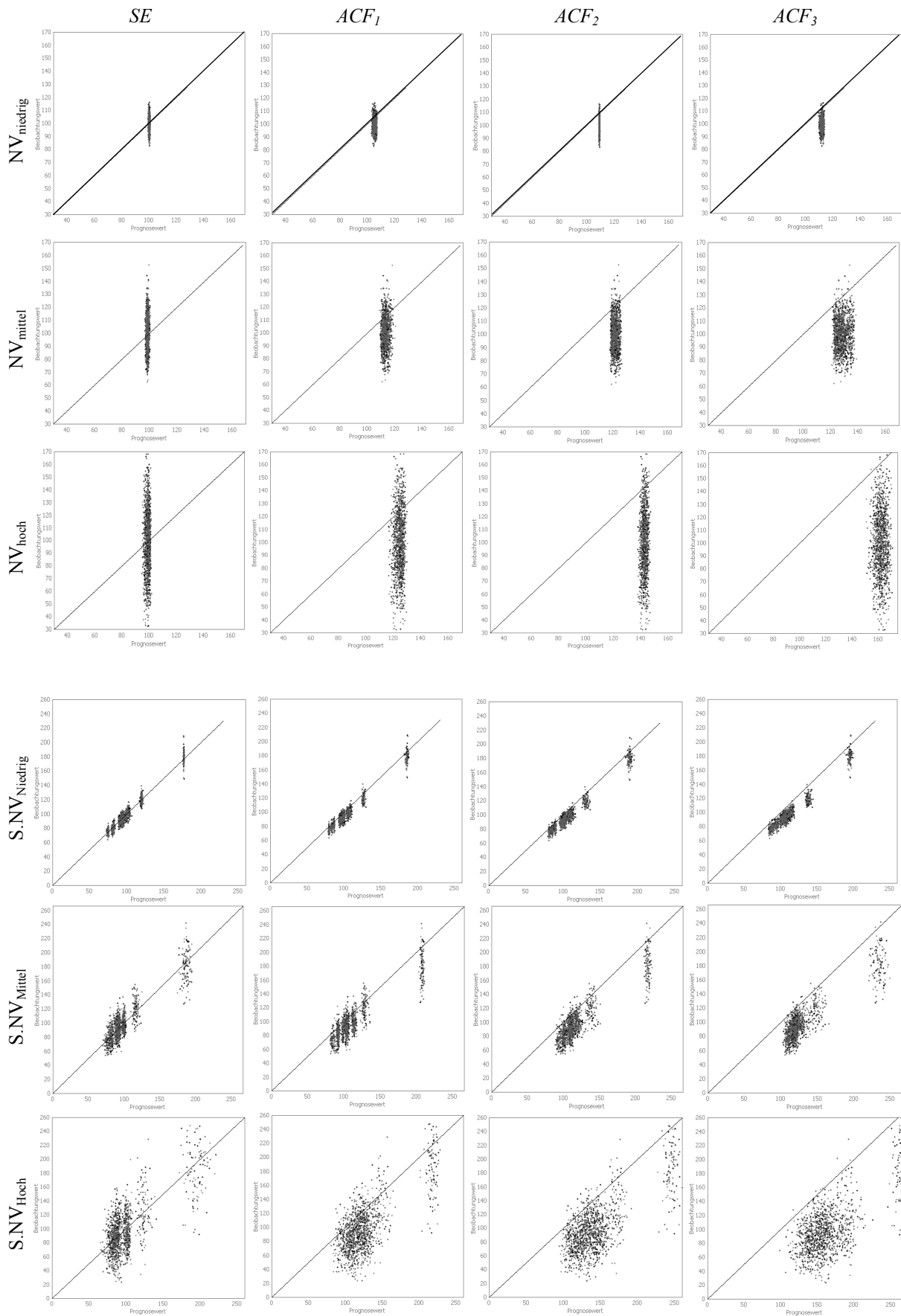


Abbildung 5.26: Streudiagramme der Prognosewerte auf Trainings-, Validierungs- und Testmenge für alle Zielfunktionen und alle Zeitreihen mit normalverteiltem Bedarfsverlauf

5.4.2.2 Beurteilung der Einsatzfähigkeit

Für Anwendungen der Warendisposition wird die Güte von Bestellmengen und Bestellzeitpunkten anhand der erzielten Servicegrade und der Dispositionskosten bestimmt. Entsprechend wird auch die Fähigkeit von NN zur Vorhersage kostenminimaler Warenbestände anhand dieser Beurteilungskriterien gemessen. Zunächst wird anhand statistischer Kennzahlen und Tests die Fähigkeit von MLPs zur validen und reliablen Minimierung unterschiedlicher Zielfunktionen aus der grafischen Analyse bestätigt. Die Güte der Ergebnisse wird anschließend in Abschnitt 5.4.2.3 diskutiert.

Zur Bestimmung der α -Servicegrade werden nach beendetem Training die Ausgabewerte aller MLPs auf Trainings-, Validierungs- und Testmenge untersucht, indem die Anzahl der Perioden in denen ein Fehlbedarf auftritt, gezählt und diese als prozentualer Anteil aller Perioden der Datenmenge ausgedrückt. In Tabelle 5.6 sind die erzielten α -Servicegrade des jeweils besten Kandidaten des MLP.*SE* trainiert mit der quadratischen Zielfunktion des *SE* und der MLP.*ACF*₁, MLP.*ACF*₂ und MLP.*ACF*₃ trainiert mit den jeweiligen asymmetrischen Kostenfunktionen *ACF*₁, *ACF*₂ und *ACF*₃ abgebildet. Die Ergebnisse der 18 Zeitreihen sind nach stationären und saisonalen Verlaufsformen geordnet und zeigen den Mittelwert (Mw) und die Standardabweichung (StAb) über alle Ergebnisse.

Tabelle 5.6: α -Servicegrade aller MLP je Datenmenge und Zeitreihe

	MLP. <i>SE</i> ($\alpha = 50\%$)			MLP. <i>ACF</i> ₁ ($\alpha = 85\%$)			MLP. <i>ACF</i> ₂ ($\alpha = 95\%$)			MLP. <i>ACF</i> ₃ ($\alpha = 99\%$)		
	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
NV _{niedrig}	48,1%**	48,8%**	52,5%**	85,3%**	84,5%**	86,3%**	96,6%**	95,3%**	96,5%**	99,5%**	98,8%**	98,8%**
NV _{mittel}	44,2%**	49,8%**	45,0%**	84,0%**	86,0%**	86,0%**	96,9%**	95,5%**	95,3%**	99,2%**	99,0%**	97,5%**
NV _{hoch}	49,4%**	50,3%**	46,5%**	85,0%**	83,0%**	80,5%**	95,1%**	95,5%**	95,0%**	99,0%**	99,3%**	99,0%**
NB _{niedrig}	53,0%**	52,5%**	50,5%**	79,1%**	81,8%**	80,0%**	95,9%**	94,8%**	97,3%**	98,4%**	98,5%**	99,0%**
NB _{mittel}	54,0%**	51,0%**	47,8%**	85,3%**	85,8%**	80,3%**	96,1%**	94,8%**	94,3%**	98,7%**	99,3%**	98,8%**
NB _{hoch}	47,3%**	48,8%**	49,3%**	82,4%**	83,8%**	83,8%**	93,5%**	95,0%**	94,0%**	99,5%**	99,0%**	99,3%**
PV _{niedrig}	48,6%**	53,0%**	51,8%**	85,8%**	86,3%**	89,5%**	96,4%**	95,5%**	95,3%**	99,2%**	98,5%**	97,8%**
PV _{mittel}	48,1%**	47,0%**	48,3%**	88,1%**	85,3%**	85,5%**	95,6%**	93,5%**	93,5%**	99,7%**	99,3%**	99,0%**
PV _{hoch}	49,1%**	52,0%**	48,8%**	85,3%**	83,5%**	82,5%**	94,1%**	94,5%**	93,0%**	98,7%**	99,0%**	99,3%**
S.NV _{niedrig}	49,4%**	49,0%**	53,3%**	85,0%**	83,3%**	86,5%**	94,8%**	94,0%**	95,5%**	100%**	98,8%**	99,3%**
S.NV _{mittel}	43,9%**	52,8%**	47,0%**	81,1%**	83,8%**	82,8%**	94,8%**	94,3%**	94,8%**	99,5%**	98,8%**	97,8%**
S.NV _{hoch}	54,0%**	51,8%**	50,5%**	87,9%**	84,0%**	83,5%**	96,4%**	94,8%**	94,3%**	99,0%**	100%**	98,8%**
S.NB _{niedrig}	53,2%**	52,5%**	49,8%**	82,4%**	84,5%**	84,5%**	95,9%**	94,8%**	95,3%**	99,0%**	99,3%**	99,0%**
S.NB _{mittel}	51,2%**	48,8%**	47,0%**	84,0%**	84,5%**	79,5%**	96,1%**	94,8%**	94,3%**	99,0%**	99,0%**	98,8%**
S.NB _{hoch}	49,9%**	51,5%**	53,0%**	84,5%**	84,3%**	84,5%**	94,1%**	95,0%**	94,5%**	99,7%**	98,8%**	99,0%**
S.PV _{niedrig}	41,1%**	47,3%**	45,8%**	85,5%**	84,5%**	89,0%**	96,1%**	94,0%**	94,8%**	99,7%**	98,3%**	97,5%**
S.PV _{mittel}	53,7%**	49,3%**	49,0%**	87,9%**	86,0%**	85,0%**	94,3%**	95,5%**	92,5%**	98,4%**	98,5%**	99,0%**
S.PV _{hoch}	50,9%**	51,3%**	48,5%**	87,3%**	85,8%**	86,8%**	96,4%**	95,5%**	95,5%**	100%**	98,8%**	98,3%**
Mw gesamt	49,4%	50,4%	49,1%	84,8%	84,5%	84,2%	95,5%	94,8%	94,7%	99,2%	98,9%	98,6%
StAb gesamt	3,70%	1,90%	2,40%	2,40%	1,20%	2,90%	1,00%	0,60%	1,10%	0,50%	0,40%	0,60%

*Ergebnis ist signifikant bei einem Testniveau von 0.05; **Ergebnis ist hoch signifikant bei einem Testniveau von 0.01

Ein Vergleich der Spalten in Tabelle 5.6 verdeutlicht, dass die unterschiedlichen MLPs auf Trainings-, Validierungs- und Testmenge je nach verwendeter Zielfunktion deutlich abweichende Servicegrade erzielen. So erzielen die mit der quadratischen Zielfunktion trainierten MLP.*SE* im Mittel über alle 18 Zeitreihenmuster einen α -Servicegrad von (49,3% / 50,4% / 49,1%), dargestellt als Tupel der Ergebnisse auf den drei Datensätzen (Trainingsmenge / Validierungsmenge / Testmenge). Unter Berücksichtigung der z. T. erheblichen Zufallsschwankungen der Zeitreihen entspricht dies näherungsweise der Zielsetzung eines Servicegrades von $\alpha_1 = 50,0\%$ der Zielfunktion *SE* und somit einer validen und reliablen Zielerreichung. Die Ergebnisse auf (Trainingsmenge / Validierungsmenge / -) verdeutlichen dabei die Fähigkeit des MLP zur Approximation der Quantile auf den Daten der Vergangenheit, während die Ergebnisse auf den Daten der (- / - / Testmenge) die Fähigkeit

der NN zur Generalisierung des Erlernen auf ungesehenen Daten bestätigen. Wie bereits zuvor grafisch bestätigt erlaubt das $MLP.SE$ eine robuste Schätzung des Mittelwertes durch die konventionelle Verwendung von MLPs als nichtlinearer Kleinste-Quadrate-Schätzer.

Im Gegensatz dazu erzielen die mit der asymmetrischen Kostenfunktion ACF_1 trainierten $MLP.ACF_1$ im Durchschnitt über alle Zeitreihen einen abweichenden α -Servicegrad von (84,8% / 84,5% / 84,2%). Dieser entspricht einem anhand der asymmetrischen Kosten $c_u = 1,00$ und $c_o = 0,176470$ spezifizierten Ziel-Servicegrad von $\alpha_2 = 85,0\%$ und unterscheidet sich damit deutlich von dem Servicegrad von $\alpha_1 = 50,0\%$ und der konventionellen Verhaltensweise von MLPs, die mit der quadratischen Zielfunktionen trainiert wurden. Eine ähnliche Veränderung der Servicegrade zeigt sich für $MLP.ACF_2$ und $MLP.ACF_3$, die ebenfalls den als Ziel bestimmten α -Servicegrad von $\alpha_3 = 95,0\%$ bzw. $\alpha_4 = 99,0\%$ über alle Zeitreihen realisieren. Somit zeigen alle MLPs die Fähigkeit zur robusten Approximation und Generalisierung unterschiedlicher Servicegrade über alle Teildatenmengen der Trainings-, Validierungs- und Testdaten. Darüber hinaus erzielen sie valide und reliabel jeweils genau den Servicegrad, der durch die Entscheidungskosten c_u und c_o des Newsboy-Modells festgelegt wird.

Die Standardabweichung der Servicegrade identifiziert dabei eine robuste Erzielung des Servicegrades durch die geringe Variation um den Mittelwert. Diese liegt durchgängig unter der Distanz zum Servicegrad einer anderen Zielfunktion. So besteht zwischen dem durchschnittlichen Servicelevel von $MLP.SE$ und $MLP.ACF_1$ eine Differenz von (35,4% / 34,1% / 35,1%) auf den drei Datensätzen, die über dem 9-fachen der Standardabweichung liegt. Entsprechend kann eine signifikant abweichende Verhaltensweise der MLPs in Abhängigkeit von ihrer Zielfunktion festgestellt werden, die über mögliche Ungenauigkeiten aus dem Experimentaufbau hinausgeht. Dies wird bei einer grafischen Betrachtung der Box-Whisker-Diagramme in Abbildung 5.27 deutlich, die Minima, Maxima, Median sowie 25%- und 75%-Quantile der Verteilung der realisierten Servicegrade darstellen.

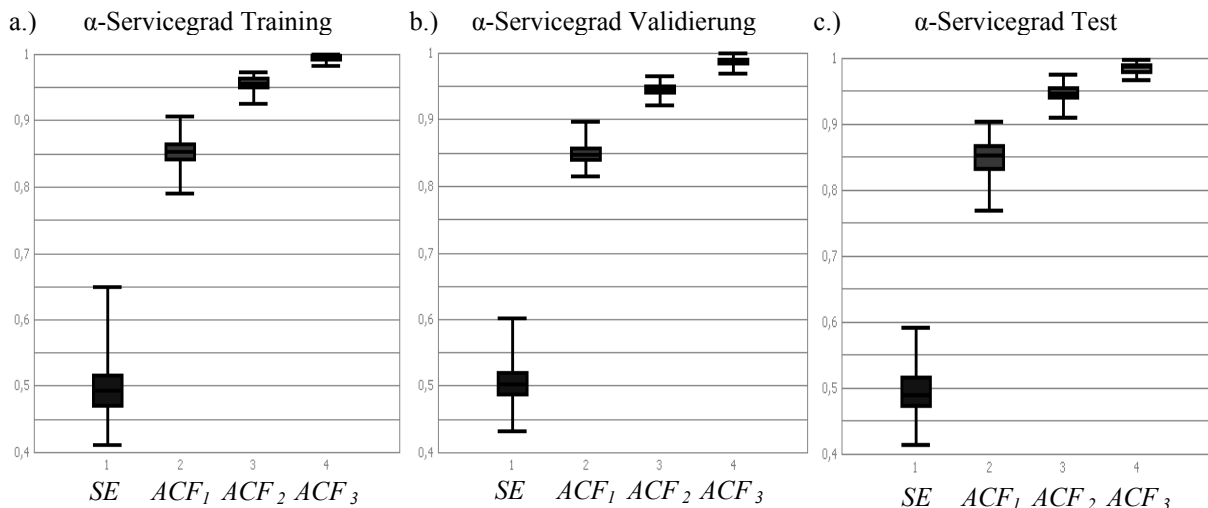


Abbildung 5.27: Verteilung der α -Servicegrade über alle MLP der Trainingsmenge (a.), der Validierungsmenge (b.) und der Testmenge (c.) für die Zielfunktionen SE , ACF_1 , ACF_2 und ACF_3

Die Abbildung bestätigt, dass sich weder die 75%-Quantile noch die Minima und Maxima der realisierten Servicegrade von SE , ACF_1 und ACF_2 auf Trainings-, Validierungs- oder Testmenge aller Zeitreihen überschneiden. Einzig die Extremwerte der Verteilung der ähnlichen Servicegrade der ACF_2 mit $\alpha_3 = 95\%$ und der ACF_3 mit $\alpha_4 = 99\%$ erlauben auf den Testdaten keine vollständige Trennung, was jedoch auf den vergleichsweise geringen Abstand der Zielsetzungen zurückzuführen ist. Zusammenfassend ist anzunehmen, dass die Unterschiede zwischen den erzielten α -Servicegraden nicht zufälliger Art sind. Unter Konstanz aller

Modellparameter sind sie vielmehr auf die Minimierung alternativer Kostenfunktionen zurückzuführen.

Aufgrund der Bedeutung dieser Erkenntnis wird nun die Signifikanz der veränderten Ausgabewerte durch asymmetrische Zielfunktionen statistisch geprüft. Eine multifaktorielle Varianzanalyse (ANOVA) prüft die Differenzen der mittleren Servicegrade auf den einzelnen Faktorstufen auf statistisch signifikante Unterschiede.¹ Die Varianzanalyse belegt, dass alle erzielten Servicegrade der $MLP.SE$, $MLP.ACF_1$, $MLP.ACF_2$ und $MLP.ACF_3$ über alle Zeitreihen signifikant, bis auf wenige Ausnahmen sogar hoch signifikant unterschiedlich sind. Daher kann anhand der erzielten Servicegrade auch statistisch bestätigt werden, dass MLPs in der Lage sind, beliebige lineare, asymmetrische Zielfunktionen mit unterschiedlichen Fehlerkosten robust zu minimieren, und dass dieses zu einer abweichenden Verhaltensweise des Verfahrens hin zu einer Warendisposition unter Erzielung signifikant unterschiedlicher Servicegrade führt.

Um diese Erkenntnis für unterschiedliche Zeitreihen zu überprüfen, werden die Mittelwerte und Standardabweichungen der α -Servicelevel für unterschiedliche Zeitreihenmuster, Schwankungsbreiten und Wahrscheinlichkeitsfunktionen in Tabelle 5.7 gegenübergestellt.

Tabelle 5.7: α - Servicegrade aller MLP je Datenmenge und Zeitreihenmuster

	$MLP.SE$ ($\alpha = 50\%$)			$MLP.ACF_1$ ($\alpha = 85\%$)			$MLP.ACF_2$ ($\alpha = 95\%$)			$MLP.ACF_3$ ($\alpha = 99\%$)		
	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
Mw stationär	49,10%	50,30%	48,90%	84,50%	84,40%	83,80%	95,60%	94,90%	94,90%	99,10%	98,90%	98,70%
Mw saisonal	49,70%	50,40%	49,30%	85,10%	84,50%	84,70%	95,40%	94,70%	94,60%	99,40%	98,90%	98,60%
StAb stationär	2,90%	2,00%	2,40%	2,50%	1,50%	3,30%	1,20%	0,60%	1,40%	0,40%	0,30%	0,60%
StAb saisonal	4,50%	1,90%	2,60%	2,40%	0,90%	2,70%	0,90%	0,60%	0,90%	0,50%	0,50%	0,60%
Mw niedrig	48,90%	50,52%	50,62%	83,85%	84,15%	85,97%	95,95%	94,73%	95,78%	99,30%	98,70%	98,57%
Mw mittel	49,18%	49,78%	47,35%	85,07%	85,23%	83,18%	95,63%	94,73%	94,12%	99,08%	98,98%	98,48%
Mw hoch	50,10%	50,95%	49,43%	85,40%	84,07%	83,60%	94,93%	95,05%	94,38%	99,32%	99,15%	98,95%
StAb niedrig	4,41%	2,43%	2,68%	2,63%	1,50%	3,46%	0,63%	0,63%	0,93%	0,57%	0,35%	0,73%
StAb mittel	4,51%	1,98%	1,39%	2,66%	0,90%	2,78%	0,95%	0,76%	0,99%	0,49%	0,31%	0,66%
StAb hoch	2,25%	1,21%	2,18%	1,99%	0,96%	2,09%	1,25%	0,39%	0,86%	0,50%	0,45%	0,37%
Mw NV.	48,17%	50,42%	49,13%	84,72%	84,10%	84,27%	95,77%	94,90%	95,23%	99,37%	99,12%	98,53%
Mw NB.	51,43%	50,85%	49,57%	82,95%	84,12%	82,10%	95,27%	94,87%	94,95%	99,05%	98,98%	98,98%
Mw PV.	48,58%	49,98%	48,70%	86,65%	85,23%	86,38%	95,48%	94,75%	94,10%	99,28%	98,73%	98,48%
StAb NV.	3,77%	1,59%	3,44%	2,20%	1,07%	2,41%	0,97%	0,64%	0,75%	0,38%	0,48%	0,71%
StAb NB.	2,52%	1,69%	2,12%	2,21%	1,31%	2,40%	1,16%	0,10%	1,23%	0,48%	0,31%	0,18%
StAb PV.	4,20%	2,51%	1,91%	1,26%	1,06%	2,63%	1,04%	0,88%	1,27%	0,63%	0,37%	0,73%
Mw gesamt	49,40%	50,40%	49,10%	84,80%	84,50%	84,20%	95,50%	94,80%	94,70%	99,20%	98,90%	98,60%
StAb gesamt	3,70%	1,90%	2,40%	2,40%	1,20%	2,90%	1,00%	0,60%	1,10%	0,50%	0,40%	0,60%

Die Servicegrade werden robust über unterschiedliche Bedarfsverteilungen erzielt, sowohl für stationäre als auch für saisonale Bedarfsmuster, über alle Intensitäten niedriger, mittlerer bis hoher Nachfrageschwankungen sowie über alle Verteilungsformen der Normalverteilung, der Negativen Binomialverteilung oder der Poissonverteilung. Das $MLP.SE$ erzielt robust einen Servicegrad um 50% sowohl über alle stationären Zeitreihen (49,1% / 50,3% / 48,9%) als auch über alle saisonalen Zeitreihen (49,7% / 50,4% / 49,3%). Ebenso wird der gewünschte Servicegrad von 85% des $MLP.ACF_1$ zuverlässig über alle stationären Zeitreihenmuster

¹ Aufgrund des Vorliegens mehrerer Gruppen von Einflussgrößen und somit vielfacher paarweiser Einzelvergleiche ist eine Beurteilung der Verfahren mittels mehrfacher t -Tests durch die mögliche Kumulierung der Alphafehler unzulässig. Daher wird nachfolgend eine multifaktorielle ANOVA in der Software SPSS 13.0 unter Berücksichtigung aller 10 Initialisierungen innerhalb einer Faktorstufe angewendet. Trotz der offensichtlich inhomogenen Varianz auf einzelnen Faktorstufen liefert die ANOVA durch die einheitliche Anzahl von Beobachtungen je Faktorstufe verlässliche Ergebnisse. Die Signifikanz aller Faktorstufen wird in Tabelle 5.6 durch die Kennzeichnung mit * (= signifikant bei 0,05 Testniveau) bzw. ** (= hoch signifikant bei 0,01 Testniveau) angezeigt.

(84,5% / 84,4% / 83,8%) sowie alle saisonalen Zeitreihenmuster (84,8% / 84,5% / 84,2%) erzielt. Die Ergebnisse des MLP.ACF₂ und des MLP.ACF₃ zeigen ebenfalls eine robuste Erreichung der Ziel-Servicelevel für saisonale sowie stationäre Zeitreihen über alle Teildatenmengen von Training, Validierung und Test.

Die konsistente Leistung einer validen und reliablen Approximation und Generalisierung zeigt sich über alle Intensitäten der Streuung, der Wahrscheinlichkeitsfunktionen und der Zielfunktionen. Insbesondere haben komplexe saisonale Zeitreihenmuster mit höherer Streuung oder auch Zeitreihen mit nicht-normalverteiltem Bedarf keinen negativen Einfluss auf die Güte der erzielten Servicegrade. Dies belegt die Fähigkeit der MLP, asymmetrische Zielfunktionen der betriebswirtschaftlichen Kosten über unterschiedliche Zeitreihenmuster, Verteilungsformen und Schwankungsbreiten robust zu minimieren und so verlässliche Bestandsentscheidungen zu treffen.¹ Da jeweils ein identisches MLP mit konstanter Architektur trainiert wurde, kann das signifikant abweichende Verhalten der NN allein auf die robuste Minimierung alternativer Zielfunktionen zurückgeführt werden.

Die Analyse je Zeitreihenmuster zeigt ferner eine kontinuierlich sinkende Standardabweichung der Mittelwerte mit Zunahme der Asymmetrie der Kosten, z. B. bei saisonalen Zeitreihen von (4,5% / 1,9% / 2,6) für MLP.SE mit bis auf (0,5% / 0,5% / 0,6%) für MLP.ACF₃. Ein ähnliches Muster zeigt sich bei stationären Zeitreihen, allen Zeitreihenmustern niedriger, mittlerer und hoher Nachfrageschwankungen und Zeitreihen mit normalverteilter, negativ-binomialverteilter und poissonverteilter Nachfrage. Diese erweiterte Eigenschaft der NN aus dem Training mit asymmetrischen Kostenfunktionen deutet auf sinkende Varianz und zunehmende Konsistenz der Dispositionsergebnisse bei zunehmender Asymmetrie der Zielfunktion hin. Dies kann unter Umständen auf die bereits dargestellte Veränderung der Fehlerfläche und die damit einhergehende Vereinfachung der Lernaufgabe für das MLP zurückgeführt werden (vgl. 5.3.2.3). Alternativ könnten die höheren Schwankungen der Servicegrade des MLP.SE auf die Verwendung einer ausreißersensitiven quadratischen Kostenfunktion im Vergleich zum linearen Verlauf der sonstigen Kostenfunktionen zurückzuführen sein. Diese Eigenschaft erscheint insofern relevant, als eine herkömmliche Vorhersage durch NN oder statistische Verfahren üblicherweise mit quadratischen Zielfunktionen erfolgt und so die Fehleranfälligkeit der Anwendung von MLPs bei der konventionellen Bedarfsprognose zur Disposition erhöhen könnte. Die Anwendung asymmetrischer Kostenfunktionen verspricht über eine gesteigerte Genauigkeit hinaus möglicherweise auch eine zusätzliche Senkung der Dispositionskosten durch eine robustere, weniger fehleranfällige Entscheidungsunterstützung versprechen.²

Die statistische Beurteilung der abweichenden Servicegrade erlaubt jedoch keine direkten Rückschlüsse auf die valide Minimierung der jeweiligen Fehlermaße bzw. Kostengrößen durch unterschiedliche Zielfunktionen. Dies erfordert eine Beurteilung der Ergebnisse anhand der erzielten Zielfunktionswerte. Jede der MLP-Varianten MLP.SE, MLP.ACF₁, MLP.ACF₂ und MLP.ACF₃ wurde mit einer anderen Zielfunktion trainiert. Im Anschluss wird die Güte der Netzwerkausgaben mit allen Fehlermaßen der mittleren SE (MSE) und den Dispositionskosten der mittleren ACF₁ (MACF₁), der mittleren ACF₂ (MACF₂) und der mittleren ACF₃

¹ Vorangegangene Experimente haben gezeigt, dass MLPs darüber hinaus auch in der Lage sind, Servicegrade unter 50,0%, mit $c_u < c_o$, zu approximieren [Crone et al. 2005]. Die Fähigkeit zur Bestimmung eines bestimmten Quantils einer beliebigen Verteilungsfunktion ist somit nicht auf eine Seite der Verteilung beschränkt. Da diese Quantile allerdings im Kontext der betriebswirtschaftlichen Warendisposition unwirtschaftlich sind, erscheinen sie als wenig relevant und werden daher hier nicht dargestellt.

² Diese Erklärung sollte lediglich als Arbeitshypothese verstanden werden, die eine hinreichende mathematische Untersuchung der Eigenschaften asymmetrischer Kostenfunktionen erfordert. Die hier gewählte Form der Überprüfung anhand mehrerer empirischer Vergleichsverfahren auf synthetischen und empirischen Zeitreihen kann und soll dies nicht erbringen.

($MACF_3$) auf der Trainings-, Validierungs- und Testmenge bewertet. Darüber hinaus wird zum Vergleich mit den quadratischen Fehlerkosten des MSE der etablierte mittlere AE (MAE) berechnet. In Tabelle 5.8 ist der Mittelwert über alle saisonalen, stationären und die Summe aller Zeitreihen je Gütekriterium dargestellt. Die jeweils niedrigsten Ausgabewerte einer Zeile über alle MLP-Varianten sind hervorgehoben und weisen auf eine Minimierung dieses Kriteriums hin.

Tabelle 5.8: Fehlermaße und Kostenmaße aller MLP je Datenmenge über alle Zeitreihen

		MLP. SE			MLP. ACF_1			MLP. ACF_2			MLP. ACF_3		
		Train	Valid	Test	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
MSE	Mw stationär	103,60	98,30	107,10	201,90	195,00	197,90	399,60	390,20	388,20	715,20	701,00	694,30
	Mw saisonal	104,20	106,40	116,80	209,30	210,20	214,20	409,60	410,50	409,80	797,40	787,20	789,80
	Mw gesamt	103,90	102,30	111,90	205,60	202,60	206,00	404,60	400,40	399,00	756,30	744,10	742,00
MAE	Mw stationär	5,47	5,46	5,52	8,02	8,11	8,04	12,10	12,22	12,04	16,92	17,01	16,84
	Mw saisonal	5,36	5,56	5,63	7,97	8,24	8,18	11,89	12,10	12,02	17,45	17,52	17,48
	Mw gesamt	5,42	5,51	5,58	7,99	8,18	8,11	11,99	12,16	12,03	17,18	17,26	17,16
$MACF_1$	Mw stationär	3,48	3,44	3,55	1,91	1,94	1,97	2,24	2,29	2,26	3,01	3,03	3,00
	Mw saisonal	3,22	3,32	3,41	1,84	1,98	2,03	2,19	2,28	2,30	3,09	3,11	3,13
	Mw gesamt	3,35	3,38	3,48	1,87	1,96	2,00	2,22	2,28	2,28	3,05	3,07	3,06
$MACF_2$	Mw stationär	3,18	3,14	3,25	0,99	1,01	1,06	0,76	0,79	0,79	0,92	0,92	0,92
	Mw saisonal	2,90	2,98	3,08	0,92	1,04	1,11	0,74	0,80	0,83	0,93	0,95	0,97
	Mw gesamt	3,04	3,06	3,17	0,95	1,03	1,08	0,75	0,80	0,81	0,93	0,94	0,94
$MACF_3$	Mw stationär	3,08	3,04	3,15	0,67	0,69	0,74	0,25	0,28	0,28	0,20	0,20	0,20
	Mw saisonal	2,79	2,86	2,97	0,60	0,72	0,79	0,23	0,30	0,33	0,19	0,20	0,23
	Mw gesamt	2,93	2,95	3,06	0,64	0,70	0,77	0,24	0,29	0,31	0,20	0,20	0,22

Je Gruppe von Zeitreihen zeigen die MLP. SE , die auf die Minimierung des quadratischen Fehlers SE trainiert wurden, einen signifikant niedrigeren MSE auf Trainings-, Validierungs- und Testmenge als alle anderen Verfahren der MLP. ACF_1 , MLP. ACF_2 und MLP. ACF_3 . Im Gegensatz dazu zeigen MLP. ACF_1 , die auf die Minimierung der ACF_1 trainiert wurden, die niedrigsten Fehlerwerte für diese Kostenfunktion im Gegensatz zu allen anderen MLPs.¹ Diese Ergebnisse sind konsistent über alle Zielfunktionen und Zeitreihenmuster. Eine Analyse der korrespondierenden Fehlermaße und Zielfunktionskriterien MSE , MAE , $MACF_1$, $MACF_2$ und $MACF_3$ je Zeitreihe bestätigt diese Erkenntnisse für jede der 18 Zeitreihen und über alle vier Zielfunktionen. Ihre Abbildung einschließlich der Mittelwerte und Varianzen je Gruppe erfolgt aus Platzgründen im Anhang. Die Verfahren der MLP. SE zeigen neben den signifikant niedrigeren Fehlerwerten der quadratischen MSE auch niedrigere Fehlerwerte der linearen MAE als alle anderen MLP-Varianten, was auf die verwandte Zielsetzung eines α -Servicegrades von 50,0% für Mittelwert und Median der Verteilung trotz unterschiedlicher Bewertung einzelner Fehlerabweichungen zurückzuführen ist.²

Zusammenfassend kann geschlossen werden, dass MLPs durch das Training mit einer spezifischen Zielfunktion diese robust auf den Trainings- und Validierungsdaten minimieren und somit signifikant niedrigere Zielfunktionswerte erzielen als andere MLPs mit abweichenden Zielfunktionen. Darüber hinaus generalisieren die so trainierten MLPs den erlernten Zusammenhang auch auf unbekanntem Daten und zeigen so auch auf der Testmenge einen signifikant niedrigeren Zielfunktionswert als MLPs mit abweichenden Zielfunktionen. Daraus folgt eine robuste Minimierung und Generalisierung alternativer Zielfunktionen über unterschiedliche Zeitreihenmuster. Da alle Parameter der MLP-Architektur außer der

¹ Die Ursache für die veränderten Zielfunktionswerte und Servicegrade wurde bereits in der grafischen Betrachtung in Abschnitt 5.4.2.1 verdeutlicht.

² Dies bestätigt die Anwendbarkeit der quadratischen Zielfunktion auch zur Minimierung symmetrischer Zielfunktionen, z. B. der linearen Kosten der absoluten Fehler. Auf eine Untersuchung der Robustheit und somit Güte beider Zielfunktionen auf Ausreißer etc. wird hier aus Platzgründen verzichtet. Nach ersten Erkenntnissen ist jedoch eine Zielfunktion des AE der SE auf empirischen Zeitreihen vorzuziehen.

Zielfunktionen konstant gehalten wurden, kann das abweichende Verhalten allein auf die Minimierung alternativer Zielfunktionen zurückgeführt werden.

5.4.2.3 Beurteilung der Entscheidungsgüte

Nachdem die Einsatzfähigkeit von MLPs zur Minimierung asymmetrischer Zielfunktionen untersucht und belegt wurde, wird ihre relative Entscheidungsgüte in der Disposition analysiert. Dies erfordert einen Vergleich der Kosten aus der Bestandsentscheidungen der MLPs mit etablierten Prognoseverfahren bei einer konventionellen Berechnung der Bestellmengen. Dazu werden für eine Kostenfunktion ACF_2 die Prognosewerte des MLP ACF_2 ermittelt, das in den vorangehenden Experimenten auf die Minimierung der ACF_2 -Kostenfunktion trainiert wurde. Diese werden über alle Prognoseursprünge ex post anhand der asymmetrischen Kosten der Über- und Unterbestände bewertet und die Kosten des Verfahrens ermittelt.

Die Dispositionsgüte des MLP wird mit den Prognoseverfahren der Naiven Prognose und dem Expertensystem ForecastPro Extended Edition 4.3 auf den 18 synthetischen Zeitreihen verglichen.¹ Die Naive Prognose, als einfaches quantitatives Prognoseverfahren, gilt als etablierter Benchmark in Verfahrensvergleichen. Obwohl sie für stationäre und saisonale Zeitreihen aufgrund ihrer fehlerhaften Modellform nur suboptimale Prognosen geben kann, erlaubt sie Rückschlüsse auf den relativen Zugewinn der Entscheidungsqualität durch Modellselektion und Parameterbestimmung für konventionelle Prognose- und Dispositionsverfahren höherer Komplexität. Um mit statistischen Prognoseverfahren auf heterogenen, stationären und saisonalen Zeitreihen zuverlässige Prognosen erstellen zu können, ist eine automatische Modellidentifikation, -selektion und -parametrisierung von Verfahren der Exponentiellen Glättung oder ARIMA erforderlich. Das kommerzielle Expertensystem ForecastPro realisiert dies für unterschiedliche Verteilungsformen des Bedarfs. Es gilt aufgrund seiner guten Leistung bei der M3-Competition als etablierter statistischer Benchmark in Verfahrensvergleichen (vgl. Abschnitt 6).² Das Expertensystem wählt ein spezifisches Verfahren basierend auf den Eigenschaften jeder Zeitreihe mittels statistischer Tests und durch Berechnung konkurrierender Prognoseverfahren auf den Trainings- und Validierungsdaten. Tabelle 5.9 gibt eine Übersicht über die von ForecastPro zur Prognose ausgewählten Verfahren für die 18 Zeitreihen.

Das Expertensystem identifiziert für die Mehrheit der Zeitreihen ein Prognosemodell der ARIMA oder Exponentiellen Glättung, die der jeweiligen Zeitreihenstruktur entsprechen. Die Expertenroutine zeigt jedoch auch Beispiele fehlerhafter Modellselektion, z. B. durch Auswahl eines saisonalen Verfahrens für die stationäre Zeitreihe der Normalverteilung mit mittlerer Streuung NV_{mittel} , der Poissonverteilung mit hoher Streuung PV_{hoch} oder der negativen Binomialverteilung mit hoher Streuung NB_{hoch} .³ Die fehlerhafte Modellselektion in 17% aller Zeitreihen ist ein Indiz für die Gefahren der Fehlspezifikation der Modellform durch automatische Expertenroutinen selbst bei einfachen Zeitreihen. Dieses Erkenntnis verdeutlicht das Potential einer integrierten, rein datengetriebenen Modellierung durch NN ohne explizite Modellannahmen. Die ausgewählten Modelle von ForecastPro werden nachfolgend berechnet und als benchmark zur Prognose und Disposition angewendet.

¹ Auf einen umfassenden Vergleich mit statistischen Prognoseverfahren wird hier verzichtet, da dieser in Kapitel 6 auf einer repräsentativen Auswahl von realen Zeitreihen erfolgt, die für die Bestimmung der empirischen Güte eines Verfahrens im Gegensatz zu seiner theoretischen Beurteilung auf synthetischen Zeitreihen eine höhere Aussagekraft besitzen. Somit erfolgt hier nur ein reduzierter Experimentaufbau, der genügt, um die positiven Erkenntnisse über die Eignung der MLP auch für die Güte ihrer Ergebnisse zu untermauern.

² Forecast Pro wird identisch mit den NN parametrisiert, indem in den Optionen der Prognosehorizont auf 400 gesetzt wird und den Verfahren die 400 letzten Beobachtungswerte durch Spezifikation als Testmenge vorenthalten werden. Die In-sample- und Out-of-sample-Prognosewerte werden anschließend in MS Excel exportiert und dort verglichen.

³ Die Identifikation einer additiven Saisonalität auf den Zeitreihen der Poissonverteilung mit niedriger, mittlerer und hoher Streuung anstelle des wahren multiplikativen Modells erscheint hingegen akzeptabel.

Tabelle 5.9: Ausgewählte Prognosemodelle je Zeitreihe durch das Expertensystem

Zeitreihen		Wahrscheinlichkeitsverteilung des Bedarfs		
Struktur	Streuung	Poissonverteilung	Negative Binomialverteilung	Normalverteilung
Stationär	Niedrig	Poisson-Modell mit einfacher Exp. Glättung	ARIMA(0,0,0)	ARIMA(0,0,0)
	Mittel	Poisson-Modell mit einfacher Exp. Glättung	ARIMA(0,0,0) mit Log-Transformation	<i>Exp. Glättung, kein Trend, multiplikative Saisonalität</i>
	Hoch	<i>Exp. Glättung, kein Trend, multiplikative Saisonalität</i>	<i>ARIMA(0,0,0)(0,1,1)₁₂ mit Log-Transformation</i>	ARIMA(0,0,0)
Saisonal	Niedrig	Exp. Glättung, kein Trend, additive Saisonalität	Exp. Glättung, kein Trend, multiplikative Saisonalität	ARIMA(0,0,0)(0,1,1) ₁₂ mit Log-Transformation
	Mittel	Exp. Glättung, kein Trend, additive Saisonalität	ARIMA(0,0,0)(1,0,1) ₁₂ mit Log-Transformation	ARIMA(0,0,0)(2,0,1) ₁₂ mit Log-Transformation
	Hoch	Exp. Glättung, kein Trend, additive Saisonalität	Exp. Glättung, kein Trend, multiplikative Saisonalität	Exp. Glättung, kein Trend, multiplikative Saisonalität

Als drittes Prognoseverfahren wird ein konventionelles MLP mit quadratischer Zielfunktion $SE(e)$ verwendet, MLP. SE . Dieses Vorgehen erlaubt eine Beantwortung der Frage, ob eventuelle Zugewinne in der Dispositionsgüte auf die grundlegende Fähigkeit der Verfahrensklasse der NN zur nichtlinearen Funktionsapproximation oder die Fähigkeit zur Verwendung neuartiger asymmetrischer Kostenfunktionen zurückzuführen ist.

Alle Prognosewerte der statistischen Verfahren sowie des MLP. ACF_2 werden berechnet und nach MS Excel exportiert. Für alle drei Prognoseverfahren Naive, ForecastPro und MLP. SE werden Bestellmengen für das Newsboy-Modell mit einem Servicegrad von $\alpha = 95\%$ bei konventioneller Berechnung der Sicherheitsbestände mit $k = 1,645$ Standardabweichungen der Prognoseresiduen ermittelt.¹ Die Dispositionsgüte dieser drei Verfahren wird anschließend den Bestellmengen der simultanen Prognose und Disposition des MLP. ACF_2 auf den 18 Zeitreihen gegenübergestellt. Die Dispositionsgüte wird je Verfahren mit Hilfe der asymmetrischen Kosten ACF_2 für jede überbevorratete oder unterbevorratete Absatzeinheit je Dispositionsperiode in Excel ermittelt.² Aufgrund der unterschiedlichen Länge der Teildatensätze werden mittlere Kosten ($MACF_2$) über die Beobachtungen eines Datensatzes und ggf. mehrerer Zeitreihen erhoben.³ Zum Vergleich mit den vorangegangenen Analysen werden zudem die α - und β -Servicegrade abgebildet. Die Darstellung der Gütekriterien erfolgt getrennt für Trainings-, Validierungs- und Testdaten, um eine Beurteilung der Lernfähigkeit sowie der Generalisierungsfähigkeit aller Verfahren zu erlauben. Die Ergebnisse werden getrennt nach stationären und saisonalen Mustern diskutiert. Die Dispositionskosten und Servicelevel sind in Tabelle 5.10 für stationäre Zeitreihen zusammengefasst.⁴

¹ Die Annahme der normalverteilten Prognoseresiduen wurde durch die Bildung nicht normalverteilter Bedarfsverteilungen bei den Zeitreihen der PV und NB jedoch explizit verletzt. Eine solche Vorgehensweise entspricht jedoch etablierter empirischer und auch theoretischer statistischer Praxis, in der für eine große Stichprobenanzahl von $n > 100$ Beobachtungswerten die Normalverteilung als Approximation der wahren Wahrscheinlichkeitsfunktion angenommen werden kann [Backhaus et al. 1994]. Es ist jedoch darauf hinzuweisen, dass diese Annahme insbesondere bei stochastischen, instationären Bedarfsdaten in den meisten Fällen nicht zutreffend erscheint. Dennoch wird eine Überprüfung dieser Hypothese in der Praxis vielfach vernachlässigt.

² Eine Beurteilung der Prognosegüte erfolgt nicht, da die Ergebnisse des Experiments Bestandsmengen und keine Bedarfsprognosen sind. Somit kann nur die Dispositionsgüte der Verfahren verglichen werden.

³ Zu Beginn des Trainingsdatensatzes fehlt eine unterschiedliche Anzahl von Prognosewerten zur Fehlerberechnung, da jede Methode eine andere Anzahl von Startwerten zur Initialisierung benötigt, das MLP beispielsweise die Beobachtungen in t_1, t_2, \dots, t_{13} , das Naive Verfahren lediglich t_1 .

⁴ Die Ergebnisse des Verfahrens mit den niedrigsten Dispositionskosten je Zeitreihe sind zur besseren Lesbarkeit hervorgehoben.

Tabelle 5.10: Servicegrade und Dispositionskosten der Verfahren auf stationären Zeitreihen

Zeitreihe	Verfahren	α -Servicelevel			β -Servicelevel			Mittlere ACF_2		
		Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
NV _{niedrig}	MLP. ACF_2	96,6%	95,3%	96,5%	99,4%	99,0%	99,1%	0,51	0,56	0,59
	MLP. SE	96,1%	95,3%	95,8%	99,1%	98,8%	98,9%	0,51	0,56	0,59
	ForecastPro	95,9%	95,3%	96,0%	99,2%	98,8%	98,9%	0,51	0,56	0,58
	Naive	96,6%	95,3%	96,5%	99,0%	98,7%	99,3%	0,75	0,83	0,75
NV _{mittel}	MLP. ACF_2	96,9%	95,5%	95,3%	99,0%	99,0%	98,6%	1,26	1,42	1,42
	MLP. SE	94,6%	93,0%	94,0%	98,3%	98,3%	97,8%	1,28	1,43	1,45
	ForecastPro	95,6%	94,8%	94,0%	98,7%	98,7%	98,2%	1,29	1,43	1,43
	Naive	95,3%	96,0%	94,8%	98,6%	98,8%	99,0%	1,93	1,95	1,89
NV _{hoch}	MLP. ACF_2	95,1%	95,5%	95,0%	98,8%	98,8%	98,7%	2,74	2,79	2,77
	MLP. SE	94,8%	94,3%	94,0%	98,5%	98,3%	98,3%	2,75	2,84	2,81
	ForecastPro	95,3%	94,5%	94,5%	98,6%	98,5%	98,5%	2,75	2,81	2,78
	Naive	94,8%	95,8%	94,8%	98,7%	98,8%	98,6%	3,71	3,76	3,92
NB _{niedrig}	MLP. ACF_2	95,9%	94,8%	97,3%	98,8%	98,7%	99,1%	0,19	0,20	0,19
	MLP. SE	87,9%	88,0%	89,5%	94,8%	94,5%	95,6%	0,22	0,23	0,20
	ForecastPro	87,9%	88,0%	89,5%	93,4%	93,3%	94,6%	0,23	0,25	0,21
	Naive	94,6%	95,0%	95,8%	98,2%	98,7%	98,9%	0,25	0,24	0,23
NB _{mittel}	MLP. ACF_2	96,1%	94,8%	94,3%	99,0%	98,8%	98,1%	0,40	0,42	0,44
	MLP. SE	90,7%	91,5%	87,8%	95,9%	96,7%	95,4%	0,46	0,43	0,47
	ForecastPro	90,7%	91,5%	87,8%	94,8%	95,7%	94,2%	0,49	0,45	0,50
	Naive	94,1%	96,3%	93,3%	98,3%	98,9%	98,6%	0,53	0,49	0,52
NB _{hoch}	MLP. ACF_2	93,5%	95,0%	94,0%	98,7%	98,0%	98,8%	0,84	0,99	0,90
	MLP. SE	88,9%	91,3%	91,5%	96,1%	96,1%	96,9%	0,99	1,05	0,98
	ForecastPro	89,7%	91,5%	91,5%	95,8%	96,0%	96,6%	1,01	1,05	0,99
	Naive	95,3%	95,3%	93,3%	98,9%	98,2%	98,6%	1,05	1,22	1,16
PV _{niedrig}	MLP. ACF_2	96,4%	95,5%	95,3%	99,1%	98,8%	98,1%	0,16	0,18	0,21
	MLP. SE	91,5%	92,0%	93,5%	97,7%	97,5%	97,0%	0,18	0,20	0,22
	ForecastPro	87,3%	87,8%	90,8%	94,2%	94,4%	94,4%	0,21	0,22	0,23
	Naive	94,3%	95,5%	94,5%	98,4%	98,4%	98,4%	0,24	0,25	0,25
PV _{mittel}	MLP. ACF_2	95,6%	93,5%	93,5%	99,1%	98,6%	98,6%	0,23	0,25	0,25
	MLP. SE	94,3%	93,3%	92,8%	98,5%	98,0%	98,1%	0,24	0,26	0,25
	ForecastPro	89,4%	91,3%	86,3%	95,9%	96,9%	94,6%	0,27	0,28	0,30
	Naive	96,6%	96,3%	96,3%	98,6%	98,6%	98,8%	0,34	0,35	0,34
PV _{hoch}	MLP. ACF_2	94,1%	94,5%	93,0%	98,7%	98,3%	98,0%	0,28	0,31	0,31
	MLP. SE	92,8%	92,8%	92,8%	97,9%	97,8%	97,6%	0,30	0,32	0,32
	ForecastPro	92,8%	92,8%	92,8%	97,5%	97,4%	97,2%	0,30	0,32	0,32
	Naive	98,2%	95,8%	94,5%	99,6%	98,6%	98,1%	0,33	0,41	0,44

Das MLP. ACF_2 erzielt die niedrigsten Dispositionskosten auf den Trainings-, Validierungs- und vor allem Testdaten über alle stationären Zeitreihen, mit Ausnahme von NV_{niedrig}. Das MLP mit asymmetrischen Kostenfunktionen zeigt somit die höchste Dispositionsgüte aller Verfahren. Ihre Anwendung auf stationären Zeitreihen verursacht geringere Kosten durch Über- und Unterbestände und somit effizientere Warenbestände gegenüber konkurrierenden Verfahren, einschließlich des etablierten Benchmarks ForecastPro und der isolierten Anwendung eines konventionellen MLP. SE als Prognoseverfahren.

Die Dispositionskosten von MLP. ACF_2 , MLP. SE und ForecastPro sind auf allen Zeitreihen signifikant niedriger als die des Naiven Verfahrens. Dies belegt, dass alle Verfahrensansätze einen Zugewinn an Prognose- und Dispositionsgenauigkeit liefern. Somit ist ihr Einsatz in diesem Verfahrenvergleich und auf diesen Zeitreihen grundsätzlich gerechtfertigt. Die α -Servicegrade der konkurrierenden Verfahren belegen die z. T. erheblichen Abweichungen

vom Soll-Servicegrad von $\alpha^* = 95\%$ und geben so Hinweise auf mögliche Ursachen der unterschiedlichen Dispositionskosten.¹

Eine Analyse der Dispositionskosten und Servicelevel aller Verfahren auf den saisonalen Zeitreihen in Tabelle 5.11 bestätigt die Erkenntnisse hinsichtlich stationärer Zeitreihen.

Tabelle 5.11: Servicegrade und Dispositionskosten der Verfahren auf saisonalen Zeitreihen

Zeitreihe	Verfahren	α -Servicelevel			β -Servicelevel			Mittlere ACF_2		
		Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
S.NV _{niedrig}	MLP. ACF_2	94,8%	94,0%	95,5%	98,9%	97,8%	98,3%	0,50	0,60	0,59
	MLP. SE	97,7%	95,8%	97,0%	99,4%	98,0%	98,6%	0,53	0,67	0,64
	ForecastPro	84,8%	84,8%	88,0%	93,9%	92,1%	94,0%	0,64	0,80	0,71
	Naive	97,2%	94,0%	96,3%	99,7%	99,0%	99,5%	3,19	3,74	3,39
S.NV _{mittel}	MLP. ACF_2	94,8%	94,3%	94,8%	98,8%	98,6%	97,8%	1,21	1,41	1,49
	MLP. SE	96,1%	94,8%	95,0%	98,5%	98,3%	98,0%	1,34	1,54	1,53
	ForecastPro	94,8%	95,3%	94,5%	97,6%	98,0%	97,8%	1,55	1,61	1,53
	Naive	96,4%	95,8%	98,3%	98,8%	98,9%	99,5%	4,32	4,41	3,96
S.NV _{hoch}	MLP. ACF_2	96,4%	94,8%	94,3%	99,3%	98,7%	98,3%	2,56	2,85	3,04
	MLP. SE	95,3%	94,8%	94,5%	98,7%	98,5%	98,2%	2,80	2,92	3,04
	ForecastPro	94,6%	94,3%	94,8%	97,9%	98,3%	98,5%	3,05	2,94	2,89
	Naive	96,6%	96,0%	96,0%	99,0%	98,7%	98,9%	5,06	5,57	5,37
S.NB _{niedrig}	MLP. ACF_2	95,9%	94,8%	95,3%	98,3%	98,8%	98,5%	0,20	0,19	0,20
	MLP. SE	91,7%	91,3%	92,5%	96,5%	96,4%	96,7%	0,20	0,21	0,20
	ForecastPro	90,4%	89,3%	93,0%	95,1%	95,3%	96,0%	0,23	0,23	0,21
	Naive	96,4%	96,3%	96,8%	98,7%	99,1%	98,6%	0,34	0,33	0,36
S.NB _{mittel}	MLP. ACF_2	96,1%	94,8%	94,3%	98,6%	98,7%	98,3%	0,42	0,43	0,44
	MLP. SE	93,0%	92,5%	90,3%	96,7%	96,5%	94,8%	0,44	0,46	0,54
	ForecastPro	91,7%	91,5%	90,3%	94,2%	95,9%	94,8%	0,55	0,47	0,51
	Naive	95,3%	97,0%	96,3%	98,1%	99,0%	98,2%	0,69	0,63	0,71
S.NB _{hoch}	MLP. ACF_2	94,1%	95,0%	94,5%	98,8%	98,1%	98,5%	0,83	0,98	0,94
	MLP. SE	93,0%	93,8%	92,0%	97,6%	96,6%	96,8%	0,92	1,11	1,10
	ForecastPro	93,3%	92,5%	93,5%	96,7%	97,0%	97,5%	1,05	1,07	1,02
	Naive	96,9%	95,8%	94,3%	98,7%	98,2%	97,6%	1,30	1,46	1,62
S.PV _{niedrig}	MLP. ACF_2	96,1%	94,0%	94,8%	98,9%	98,2%	97,6%	0,16	0,19	0,22
	MLP. SE	89,1%	91,0%	91,5%	96,1%	95,8%	95,2%	0,19	0,22	0,24
	ForecastPro	89,7%	90,0%	91,3%	96,1%	95,7%	95,3%	0,20	0,23	0,24
	Naive	94,6%	95,8%	95,0%	98,5%	98,5%	97,8%	0,26	0,26	0,29
S.PV _{mittel}	MLP. ACF_2	94,3%	95,5%	92,5%	98,4%	98,5%	98,3%	0,26	0,26	0,26
	MLP. SE	95,6%	94,0%	94,5%	98,7%	97,7%	98,7%	0,25	0,29	0,25
	ForecastPro	95,9%	92,8%	96,3%	97,1%	95,7%	98,5%	0,37	0,48	0,31
	Naive	96,9%	96,0%	97,0%	98,4%	97,8%	99,0%	0,45	0,51	0,42
S.PV _{hoch}	MLP. ACF_2	96,4%	95,5%	95,5%	99,5%	98,9%	98,7%	0,27	0,31	0,32
	MLP. SE	95,3%	96,0%	94,5%	98,9%	98,7%	98,4%	0,29	0,31	0,32
	ForecastPro	94,1%	94,0%	94,5%	97,8%	97,3%	97,8%	0,44	0,49	0,46
	Naive	96,6%	94,8%	95,3%	99,2%	98,7%	98,8%	0,46	0,52	0,50

Erneut erzielt das MLP. ACF_2 die niedrigsten Dispositionskosten auf Trainings-, Validierungs- und Testdaten bei der Mehrheit der saisonalen Zeitreihen, mit Ausnahme von S.NV_{hoch} und S.PV_{mittel}. Das MLP mit asymmetrischen Kostenfunktionen zeigt somit erneut die höchste Dispositionsgüte aller Verfahren durch die geringsten Kosten und verbesserte Warenbestände.²

¹ Eine Betrachtung der β -Servicegrade zeigt ein höheres Niveau als das der α -Servicegrade, wie aufgrund der Verteilungsformen zu erwarten ist. Im Durchschnitt wird mit der Befriedigung von $\alpha=95\%$ aller Bedarfsergebnisse ein Anteil von $\beta=98,7\%$ am Gesamtbedarf in Mengeneinheiten befriedigt, was in der korrekten Spezifikation der Servicelevel für die Zielsetzung von MLPs zu berücksichtigen ist.

² Eine multifaktorielle Varianzanalyse belegt für die Mehrzahl der stationären und saisonalen Zeitreihen die statistisch signifikanten Unterschiede der Dispositionsgüte von MLP. ACF_2 und der anderen Verfahren. Vielfach ermittelt sie durch eine Untersuchung der „marginal means“ drei homogene Untergruppen mit

Aus den dargestellten Kosten kann das Einsparungspotenzial der Verfahren zur Prognose und Disposition als Differenz der Dispositionskosten ermittelt werden. Je nach Schwankungsbreite, Funktionsform und Muster der Zeitreihe variiert die Differenz der Dispositionskosten erheblich, z. B. für $MLP.ACF_2$ und ForecastPro zwischen 0,7% für NV_{mittel} bis zu 16,7% für PV_{mittel} . In Tabelle 5.12 erfolgt eine aggregierte Betrachtung der Dispositionskosten je Verteilungsform und Zeitreihenmuster sowie über alle Zeitreihen.

Tabelle 5.12: Mittlere Dispositionskosten der Verfahren nach Verlauf und Verteilung der Zeitreihe

Zeitreihen	Verfahren	Mittlere Kosten $MACF_2$ der Verfahren								
		Stationäre Zeitreihen			Saisonale Zeitreihen			Zeitreihen Gesamt		
		Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
NV.	$MLP.ACF_2$	1,50	1,59	1,59	1,42	1,62	1,71	1,46	1,60	1,65
	$MLP.SE$	1,51	1,61	1,62	1,55	1,71	1,74	1,53	1,66	1,68
	ForecastPro	1,51	1,60	1,60	1,75	1,78	1,71	1,63	1,69	1,65
	Naive	2,13	2,18	2,19	4,19	4,57	4,24	3,16	3,38	3,21
NB.	$MLP.ACF_2$	0,48	0,54	0,51	0,49	0,54	0,53	0,48	0,54	0,52
	$MLP.SE$	0,56	0,57	0,55	0,52	0,59	0,61	0,54	0,58	0,58
	ForecastPro	0,58	0,58	0,57	0,61	0,59	0,58	0,59	0,59	0,57
	Naive	0,61	0,65	0,64	0,78	0,81	0,90	0,69	0,73	0,77
PV.	$MLP.ACF_2$	0,22	0,25	0,26	0,23	0,25	0,27	0,23	0,25	0,26
	$MLP.SE$	0,24	0,26	0,26	0,24	0,28	0,27	0,24	0,27	0,27
	ForecastPro	0,26	0,28	0,28	0,33	0,40	0,34	0,30	0,34	0,31
	Naive	0,30	0,34	0,34	0,39	0,43	0,41	0,35	0,38	0,37
Verteilungen Gesamt	$MLP.ACF_2$	0,73	0,79	0,79	0,71	0,80	0,83	0,72	0,80	0,81
	$MLP.SE$	0,77	0,81	0,81	0,77	0,86	0,87	0,77	0,84	0,84
	ForecastPro	0,78	0,82	0,82	0,90	0,92	0,88	0,84	0,87	0,85
	Naive	1,01	1,06	1,06	1,78	1,94	1,85	1,40	1,50	1,45

Die mittleren Kosten $MACF_2$ bestätigen das $MLP.ACF_2$ als bestes Verfahren mit den niedrigsten Dispositionskosten konsistent über alle Zeitreihengruppen. Dabei erzielt es sowohl die niedrigsten Kosten auf der Trainings- und Validierungsmenge als auch auf der Testmenge und beweist so die Lernfähigkeit sowie die Generalisierungsfähigkeit des Verfahrens.

Für die Anwendung zur Vorhersage von Bestellmengen in der Disposition ist besonders die Generalisierungsleistung auf der Testmenge von Bedeutung. Hier zeigt das $MLP.ACF_2$ die niedrigsten Kosten für die Zeitreihenmuster stationärer Zeitreihen von 0,79 im Vergleich zu ForecastPro mit 0,82 und den herkömmlichen $MLP.SE$ mit 0,81. Auf den saisonalen Zeitreihen zeigt das $MLP.ACF_2$ Kosten von 0,83 im Vergleich zu 0,87 beim $MLP.SE$ und 0,88 bei ForecastPro. Ein identisches Bild ergibt sich bei einer Betrachtung der Ergebnisse nach der Verteilungsfunktion. Das $MLP.ACF_2$ verursacht die geringsten Dispositionskosten von 1,65 über alle Zeitreihen des normalverteilten Bedarfs $NV.$, für Bedarf mit negativer Binomialverteilung $NB.$ von 0,52 und für poissonverteilten Bedarf $PV.$ von 0,26. Entsprechend verursacht das MLP mit asymmetrischen Kostenfunktionen auch die niedrigsten Gesamtkosten von 0,81 über alle Verteilungen und Zeitreihenmuster, gefolgt vom $MLP.SE$ mit durchschnittlichen Kosten von 0,84, ForecastPro mit Kosten von 0,85 und dem Naiven Benchmark-Verfahren mit Kosten von 1,45 auf den Testdaten.

Als zweitbestes Verfahren kann das $MLP.SE$ mit quadratischen Zielfunktionen identifiziert werden. Es bietet jedoch nur eine insignifikant bessere Dispositionsgüte als das Expertensystem ForecastPro über alle Zeitreihen und zeigt keine Dominanz bei der Betrachtung

signifikant unterschiedlichen Dispositionskosten. Diese Gruppen bestehen aus dem Verfahren des $MLP.ACF_2$ als Gruppe mit den niedrigsten mittleren Kosten, dem Naiven Verfahren als Gruppe mit den höchsten Kosten und der Gruppe der Verfahren $MLP.SE$ und ForecastPro als Gruppe mit identischen mittleren Kosten, die keine statistisch signifikanten Unterschiede in der Dispositionsgüte erkennen lassen. Die ANOVA bietet jedoch keine konsistenten Erkenntnisse über die Dominanz des $MLP.ACF_2$ hinaus, weshalb auf eine detaillierte Dokumentation hier verzichtet wird.

unterschiedlicher Gruppen von Zeitreihen. So zeigt ForecastPro auf den Zeitreihen NV. und NB. die zweitniedrigsten Dispositionskosten gefolgt von MLP.*SE* sowie vice versa für die Zeitreihen PV.

Da die Erkenntnisse aus den metrisch skalierten mittleren Kosten durch einzelne Ausreißer verzerrt werden können, wird zusätzlich der Rang der Verfahren nach den niedrigsten Kosten je Zeitreihe ermittelt und daraus der mittlere Rang über alle Zeitreihen bestimmt. Verfahren mit einem niedrigeren mittleren Rang erzielen somit eine bessere relative Dispositionsgüte als die anderen Verfahren. Tabelle 5.13 enthält die mittleren Ränge der Verfahren nach $MACF_2$.

Tabelle 5.13: Mittlere Ränge der Dispositionskosten der Verfahren nach Verlauf und Verteilung der Zeitreihe

Zeitreihen	Verfahren	Mittler Rang der Verfahren nach $MACF_2$								
		Stationäre Zeitreihen			Saisonale Zeitreihen			Zeitreihen gesamt		
		Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
NV.	MLP. ACF_2	1,3	1,7	1,7	1,0	1,0	1,3	1,2	1,3	1,5
	MLP. <i>SE</i>	2,3	2,0	2,7	2,0	2,0	2,3	2,2	2,0	2,5
	ForecastPro	2,3	2,3	1,7	3,0	3,0	2,3	2,7	2,7	2,0
	Naive	4,0	4,0	4,0	4,0	4,0	4,0	4,0	4,0	4,0
NB.	MLP. ACF_2	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0
	MLP. <i>SE</i>	2,0	2,0	2,0	2,0	2,3	2,7	2,0	2,2	2,3
	ForecastPro	3,0	3,3	3,0	3,0	2,7	2,3	3,0	3,0	2,7
	Naive	4,0	3,7	4,0	4,0	4,0	4,0	4,0	3,8	4,0
PV.	MLP. ACF_2	1,0	1,0	1,0	1,3	1,3	1,3	1,2	1,2	1,2
	MLP. <i>SE</i>	2,0	2,0	2,0	1,7	1,7	2,0	1,8	1,8	2,0
	ForecastPro	3,0	3,0	3,0	3,0	3,0	2,7	3,0	3,0	2,8
	Naive	4,0	4,0	4,0	4,0	4,0	4,0	4,0	4,0	4,0
Verteilungen Gesamt	MLP. ACF_2	1,1	1,2	1,2	1,1	1,1	1,2	1,1	1,2	1,2
	MLP. <i>SE</i>	2,1	2,0	2,2	1,9	2,0	2,3	2,0	2,0	2,3
	ForecastPro	2,8	2,9	2,6	3,0	2,9	2,4	2,9	2,9	2,5
	Naive	4,0	3,9	4,0	4,0	4,0	4,0	4,0	3,9	4,0

Die Rangreihung der Verfahren bestätigt die Überlegenheit der MLP. ACF_2 bei der Disposition, die mit einem Rang von 1,2 sowohl das etablierte Expertensystem ForecastPro mit 2,5 als auch das MLP.*SE* mit 2,3 übertreffen. Somit können MLP. ACF_2 konsistent als beste Methode zur Disposition über alle Zeitreihen bestätigt werden.

Das Verfahren der MLP.*SE* mit konventioneller Zielfunktion wird anhand der Ränge erneut als zweitbestes Verfahren identifiziert. Die Zugewinne in der Dispositionsgüte des MLP. ACF_2 sind somit nicht auf die grundsätzlichen Fähigkeiten der Verfahrensklasse der NN bei der nichtlinearen Prognose oder der datengetriebenen Parametrisierung nichtlinearer Zeitreihen zurückzuführen. Vielmehr wird die integrierte Modellierung für die simultane Prognose und Disposition mittels asymmetrischer Kostenfunktionen als Verursacher für die verringerten Bestandskosten identifiziert werden. Die Unterschiede zwischen MLP.*SE* und ForecastPro sind jedoch auch hier gering und nicht statistisch signifikant, so dass nicht auf eine grundsätzliche Überlegenheit auf den analysierten Zeitreihen geschlossen werden kann.

5.4.3 Zusammenfassung und Interpretation

In der vorangegangenen Studie wurde für synthetische Zeitreihen gezeigt, dass ein MLP mit konstanter Netzwerkarchitektur und konstanten Lernparametern kostenminimale Bestände für unterschiedliche Wahrscheinlichkeitsverteilungen und Streuungsintensitäten valide und reliabel erlernen und diese auf unbekannte Daten generalisieren kann.

Dies entspricht der Bestimmung beliebiger Quantile einer Verteilungsfunktion durch die Minimierung einer asymmetrischen Kostenfunktionen, und bestätigt die Theorie generalisier-

ter Kostenfunktionen nach GRANGER [1969]¹ und den formalen Beweis der Fähigkeit von NN zur Approximation beliebiger Quantile einer Verteilungsfunktion nach WHITE [1992a] erstmals empirisch. Die jeweiligen Quantile können unmittelbar anhand der asymmetrischen Kosten der Überbestände und Unterbestände spezifiziert und anhand der erzielten Servicegrade gemessen werden, ein Verfahren, welches konsistent mit den Erkenntnissen der Forschungsbereiche der Warendisposition und der Absatzprognose ist. Es wurde gezeigt, dass MLPs in der Lage sind, die Quantile auf unterschiedlichen Zeitreihenmustern stationärer sowie saisonaler Zeitreihen der Normalverteilung, der Negativen Binomialverteilung oder der Poissonverteilung mit unterschiedlicher Streuung und ohne A-priori-Kenntnis oder separate Bestimmung der Verteilungsfunktion zu ermitteln. Somit sind weder eine vorangehende Untersuchung der Beobachtungswerte noch explizite Verteilungsannahmen wie für statistische Verfahren erforderlich, z. B. durch statistische Tests zur Überprüfung der Prämisse normalverteilter Beobachtungswerte. Das MLP erfordert lediglich die Spezifikation der Kosten einer Über- und Unterprognose in einer asymmetrischen Kostenfunktion. Die Entscheidungskosten aus den Servicegraden sind signifikant geringer als die herkömmlicher, statistischer Prognoseverfahren sowie führender Expertensysteme zur Prognose. Die Verbesserungen in der Güte und Effizienz der Warenbestände werden konsistent erzielt, sowohl auf den historischen Daten zur Parametrisierung als auch auf unbekanntem Testdaten aller Zeitreihenmuster. Die dargestellte Funktionalität und die Güte der Ergebnisse rechtfertigen daher den Einsatz von MLPs mit asymmetrischen Kostenfunktionen zur Disposition auf Zeitreihen bei stationärem und instationärem Bedarf unterschiedlicher Streuung und beliebiger Verteilungsformen des Bedarfs. In der empirischen Anwendung kann somit nach der Erhebung repräsentativer Dispositionskosten eine einzige Konfiguration eines MLP mit asymmetrischen Kostenfunktionen produktübergreifend zur effizienten Disposition eingesetzt werden.

Die Bewertung der Prognosegenauigkeit nach Kosten auf den synthetischen Zeitreihen erlaubt eine erste Hochrechnung des Einsparungspotenzials zwischen ForecastPro und den MLP.ACF aus der Differenz der durchschnittlichen Dispositionskosten. Diese betragen 0,81 für das MLP.ACF₂ im Vergleich zu 0,85 für den Software-Benchmark ForecastPro und 1,45 bei der Naiven Prognose. Könnte eine Kosteneinsparung in Höhe von 4,7% pro Dispositionsvorgang und disponiertem Artikel im Vergleich zu ForecastPro oder 44,2% im Vergleich zur Naiven Prognose in empirischen Anwendungen bestätigt werden, entspräche dies einer erheblichen Effizienzsteigerung durch die Anwendung von MLPs mit asymmetrischen Kostenfunktionen. Bereits für eine Handelskette mittlerer Größe ergeben sich damit erhebliche Einsparpotenziale in der Warendisposition, ohne Veränderung der unternehmensinternen und -übergreifenden Logistikketten.²

Neben einer Kostensenkung bieten MLPs weitere Vorteile in der Warendisposition, indem sie die Teilprozesse zur Ermittlung kostenminimaler Warenbestände vereinfachen und in einen integrierten, simultanten und rein datengetriebenen Trainingsprozess realisieren. Insbesondere erfordern sie weder Ex-ante-Annahmen oder statistische Tests über die Verteilung der

¹ Auf eine Gültigkeit zu der generalisierten Quantil-Regression nach KOENKER [2005] kann nicht geschlossen werden, da diese auch die Veränderung der Quantile im Zeitablauf im Sinne von ARCH- oder GARCH-Modellen zulässt. Dies ist in weiterführenden Studien mit längeren Prognosehorizonten zu untersuchen.

² Je nach Betriebsform des Handels disponiert ein Discounter ca. 780 bis 1.600 Artikel, ein Supermarkt 7.000 bis 12.000 Artikel, Verbrauchermärkte 21.000 bis 40.000 und SB-Warenhäuser 33.000 bis 63.000 Artikel. Je nach Dispositionsprozess wird zudem wöchentlich oder täglich für eine Vielzahl von Absatzstellen je Unternehmensgruppe disponiert. Da viele Handelsunternehmen einfache Prognosesysteme und -verfahren einsetzen, die noch hinter der Prognosegüte des Expertensystems ForecastPro zurückbleiben (vgl. Abschnitt 1.1), könnten die Einsparpotenziale durch Einsatz von NN für den Großteil des Handels deutlich über 5% liegen. Das theoretische Einsparpotenzial pro Dispositionsvorgang durch die Anwendung eines MLP mit asymmetrischen Kostenfunktionen für ein Unternehmen im Jahr erscheint somit erheblich.

Bedarfe bzw. der Prognoseresiduen noch eine iterative Modellbildung durch Experten oder Expertensysteme. Diese Vereinfachung der Prozesse zur Bedarfsprognose und Bestellmengenrechnung in der Warendisposition wird in Abbildung 5.28 anhand der Wertschöpfungsketten-Diagramme nach Abschnitt 2.1. dargestellt. Die Prozesskette der Warendisposition in Handelsfilialen mit der zeitlich geordneten Abfolge aller Teilfunktionen bei Anwendung eines konventionellen, statistischen Prognoseverfahrens ist zum Vergleich in Abbildung 5.28.a erneut dargestellt.¹

Die Abbildung der Prozesskette unter Anwendung des MLP.*SE* als konventionelles Prognoseverfahren in Abbildung 5.28.b verdeutlicht, dass die Fähigkeit der NN zur semi-parametrischen Modellierung den Prozessschritt der Modellselektion eliminiert, da das MLP das zugrunde liegende Erklärungsmodell selbstständig aus den präsentierten Trainingsdaten lernt. Auch wenn der vorangehende Teilprozess der Datenanalyse nicht zwingend für eine standardisierte Anwendung der MLP.*SE* in operationalen Handelsprozessen erforderlich ist, kann nicht vollständig auf sie verzichtet werden. Aufbauend auf der Prognose des MLP.*SE* erfolgt nun eine konventionelle Berechnung der Sicherheitsbestände, der kostenminimalen Warenbestände und der Bestellmengen wie bei statistischen Verfahren.

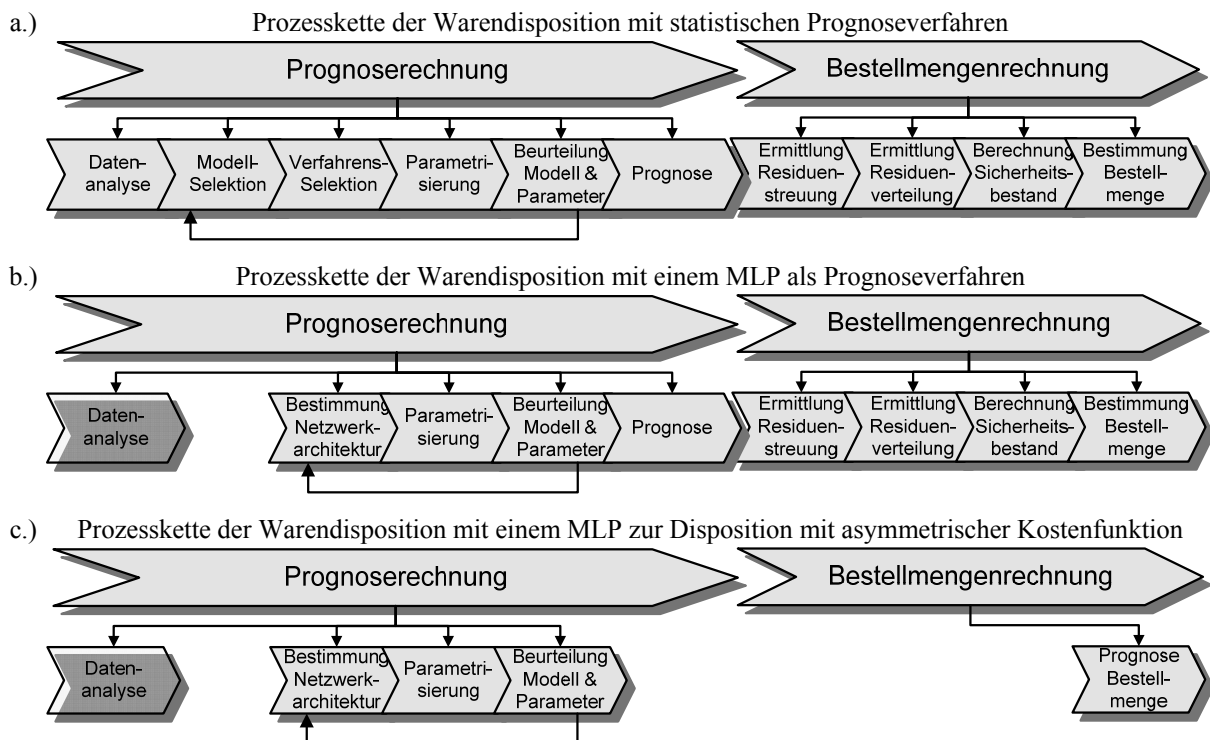


Abbildung 5.28: Prozessketten der Disposition mit statistischen Prognoseverfahren (a.), MLP.*SE* als Prognoseverfahren (b.) und MLP.*ACF* mit asymmetrischen Zielfunktionen zur Disposition (c.)

Im Vergleich zu statistischen und konventionellen MLPs erfordert die Wertschöpfungskette der MLP.*ACF* zur simultanen Prognose und Disposition in Abbildung 5.28.c deutlich weniger Prozessschritte. Insbesondere kann auf die Bestimmung der Parameter der Prognoseresiduen und ihrer Verteilungsform sowie die explizite Berechnung von Sicherheitsbeständen verzichtet werden. Somit erfolgt eine Vereinfachung des Ablaufes der Warendisposition im Vergleich zur manuellen, computergestützten sowie vollautomatisierten Disposition durch Expertensysteme. Dies ist auch unter dem Gesichtspunkt der Fehlervermeidung in einzelnen

¹ Die Betrachtung der Filialdisposition vernachlässigt somit die Teilfunktionen Liefermengenrechnung, Aufteilung usw. (vgl. Abschnitt 2.2) sowie Lieferverzögerungen, die nicht in die Struktur der synthetischen Zeitreihen eingeflossen sind. Diese könnten jedoch konsistent durch entsprechende Modellierung der Bedarfszeitreihe und der asymmetrischen Kostenfunktion erfasst werden.

Teilfunktionen der Prognose und Disposition sowie verringerter Schnittstellenprobleme zwischen einzelnen Teilfunktionen als positiv zu beurteilen. Ein MLP, das mit einer asymmetrischen Zielfunktion zur Disposition trainiert wurde, erzielt eine effiziente, sparsame Modellierung unter Vermeidung von Freiheitsgraden entlang aller Teilentscheidungen des Dispositionsprozesses. Die Modellierung erfolgt zudem näher an der betrieblichen Realität, da lediglich die betriebswirtschaftlichen Kosten der Entscheidungsfindung und die Daten des vergangenen Bedarfs verwendet werden.¹

Funktionsweise und Güte des neuartigen Verfahrens bestätigen sich auf allen Zeitreihenmustern, Verteilungsformen und Streuungen der synthetischen Zeitreihen, so dass man von einer Zustandsdominanz sprechen könnte. Diese Erkenntnisse sind jedoch nur für die vorliegenden, synthetischen Zeitreihen gültig und nur bedingt verallgemeinerbar. Insbesondere wurden lediglich Zeitreihen mit multiplikativer Saisonalität und multiplikativen Fehlertermen untersucht.² Das Problem der eingeschränkten Gültigkeit des Experimentaufbaus korrespondiert mit dem grundsätzlichen Problem empirischer Verfahrenvergleiche, in Anlehnung an das Induktionsproblem des kritischen Rationalismus nach POPPER. Demnach können formallogisch aus Einzelfällen keine allgemeinen Gesetzmäßigkeiten abgeleitet werden. Entsprechend kann selbst durch valide und reliable Verfahrenvergleiche von repräsentativen Prognoseverfahren und Zeitreihen nicht auf eine grundsätzliche Dominanz eines einzelnen Prognoseverfahrens auf allen möglichen Zeitreihen, Verteilungsformen und in beliebigen Anwendungen geschlossen werden. Entsprechend kann durch das vorangegangene Experiment auf einer Stichprobe synthetischer Zeitreihen nicht belegt werden, dass ein MLP mit asymmetrischen Zielfunktionen auf beliebigen Zeitreihen und unabhängig von der Anwendungsdomäne systematisch niedrigere Kosten erzielt als konkurrierende Vergleichsverfahren. Im Sinne des kritischen Rationalismus konnte jedoch gezeigt werden, dass im vorliegenden Experiment die Ergebnisse der konventionellen MLP.*SE* und der neuartigen MLP.*ACF* zur simultanen Disposition nicht signifikant schlechter sind als die herkömmlicher Verfahren und Expertensysteme. Dies falsifiziert die Erkenntnisse vorangegangener, einflussreicher Veröffentlichungen hinsichtlich einer geringeren Prognosegüte von NN [siehe u.a. Chatfield 2005; Chatfield 1993] oder der Unfähigkeit von NN zur Prognose saisonaler Zeitreihen ohne entsprechende Vorverarbeitung [Hill et al. 1994; Hill et al. 1996; Remus und O'Connor 2001; Zhang und Qi 2005] im Sinne willkürlicher Basissätze und trägt so zum theoretischen Erkenntnisfortschritt ungeachtet empirischer Erfolge bei.

Diese theoretischen Erkenntnisse aus der Analyse synthetischer Zeitreihen müssen fortführend im Zuge spezifischer Problemstellungen und Anwendungen empirisch durch einen wissenschaftlich fundierten Verfahrenvergleich im Einzelfall auf ihre Güte überprüft werden. Für die Warendisposition in Handelsfilialen spiegeln synthetische Zeitreihen ohne Ausreißer, Niveauverschiebungen, strukturelle Brüche und lokale Trends insbesondere nicht die Komplexität realer Bedarfszeitreihen des Handels wider. Aus diesem Grund erfolgt im nachfolgenden Abschnitt 6 eine repräsentative Beurteilung von MLPs zur Prognose und Disposition anhand realer Zeitreihen einer ausgewählten Anwendungsdomäne des stationären Einzelhandels.

¹ Dies könnte sich positiv auf die Akzeptanz des Verfahrens bei menschlichen Experten auswirken.

² Die Prognose von Zeitreihen mit multiplikativer Saisonalität stellt durch die Interaktion der Streuung mit den saisonalen Indizes eine besondere Herausforderung für einzelne statistische Verfahren dar. Die hier gefundenen Erkenntnisse sind jedoch auch auf einfachere Muster additiver Saisonalität übertragbar und wurden bereits in einer Studie, allerdings mit beschränkter Validität und Reliabilität analysiert und belegt [Crone et al. 2005].

6 Empirischer Vergleich von Neuronalen Netzen zur Prognose und Disposition

Im Anschluss an die theoretische Überprüfung der Einsatzfähigkeit und Güte von NN zur Warendisposition auf synthetischen Zeitreihen erfolgt ihre Überprüfung auf empirischen Zeitreihen des Handels. Als empirische Problemstellung wird die Disposition an Warenautomaten als vereinfachte Betriebsform des Handels gewählt.

Zunächst werden aus der Zielsetzung des Verfahrensvergleichs Arbeitshypothesen abgeleitet, die durch die Experimente mit einem validen und reliablen Versuchsaufbau überprüft werden. Anschließend wird das Logistiksystem der Warenautomaten analysiert und die Warendisposition am Automaten in die Modelle der Bedarfs- und Bestellmengenrechnung eingeordnet. Dies erlaubt die Ableitung der kontrollierbaren Entscheidungsvariablen der Warendisposition sowie die Dekomposition der Problemstellung zur Auswahl eines Modells der Bestellmengenrechnung mit minimaler Komplexität.

Im Anschluss wird die Stichprobe der Zeitreihendaten analysiert und hinsichtlich ihrer Charakteristika für den Verfahrensvergleich kategorisiert. Die zu vergleichenden Verfahren der NN und der statistischen Benchmarks zur Prognose sowie zur Bestellmengenrechnung werden separat untersucht und ihre Konfiguration im Versuchsaufbau beschrieben. Nach einer Beschreibung der Beurteilungsmaße der Prognose- und Dispositionsgenauigkeit werden die Ergebnisse der Vergleichsstudie der Prognoseverfahren sowie der Bestellmengenrechnung vorgestellt, analysiert und zusammengefasst. Dabei wird abschließend die Interaktion der Prognosefehler und der Dispositionskosten analysiert, und es werden Rückschlüsse auf die Auswahl von Prognoseverfahren in Theorie und Praxis gezogen.

6.1 Gestaltung der Vergleichsstudie

6.1.1 Zielsetzung des Verfahrensvergleichs

Primäres Ziel der Vergleichsstudie ist die Beurteilung der Genauigkeit von NN mit asymmetrischen Zielfunktionen in der Warendisposition. Der Verfahrensvergleich erfolgt auf empirischen Zeitreihen des Absatzes von Markenzigaretten an Warenautomaten, die als eine besondere Betriebsform des stationären Einzelhandels kontrollierte Experimente zulassen.

Die Studie geht aufgrund der Struktur des Versuchsaufbaus jedoch über die einfache Untersuchung von NN mit asymmetrischen Kostenfunktionen hinaus. Da die Fähigkeit von NN zur zeitreihen- und kausalanalytischen Prognose im Vergleich zu etablierten statistischen Verfahren noch immer umstritten ist und die relative Güte von Prognoseverfahren stark von dem jeweiligen Datensatz beeinflusst sein kann, werden zunächst NN als reine Prognoseverfahren beurteilt. Dies erfolgt hinsichtlich ihrer Güte zur Absatzprognose im Vergleich zu etablierten statistischen Prognoseverfahren anhand robuster Gütekriterien in Form ausgewählter statistischer Fehlermaße.

Erst im Anschluss daran werden NN hinsichtlich ihrer Güte in der Warendisposition untersucht, im Sinne des eigentlichen Entscheidungsproblems, das der Absatzprognose zugrunde liegt. Zunächst erfolgt das Training von NN als Prognoseverfahren mit einer anschließenden Berechnung von Sicherheitsbeständen und Bestellmengen. Dabei wird die konventionelle Berechnung der Bestellmengen mittels des Newsboy-Modells durchgeführt, das für die Disposition an Warenautomaten eine zulässige stochastische Modellstruktur darstellt.¹ Die

¹ Die Anwendung des Newsboy-Modells wird in Abschnitt 6.2.1.3 detailliert hergeleitet und begründet. Gemäß der Darstellung des Modells in Abschnitt 2.3.3.1 erfolgt sie unter Bestimmung des optimalen Servicegrades aus dem Kostenverhältnis von Fehlmengen und Überbeständen, der Berechnung der

Beurteilung der Dispositionsgüte der NN erfolgt anhand der entstehenden Überbestände und Unterbestände, bewertet mit den jeweiligen Entscheidungskosten. Dies entspricht einer Beurteilung der Wirtschaftlichkeit der Bestellmengen im Gegensatz zur Genauigkeit der Absatzprognose der konkurrierenden Verfahren.

Zusätzlich werden NN mit einer Zielfunktion der asymmetrischen Entscheidungskosten untersucht. Die asymmetrische Kostenfunktion spiegelt die tatsächlichen Entscheidungskosten des materialwirtschaftlichen Dispositionsproblems wider und wird abschließend auch zur Bewertung der Dispositionsgüte verwendet. Die NN berechnen somit nicht länger einen Prognosewert, sondern integrieren die einzelnen Problemstellungen der Bedarfsprognose und der Bestellmengenrechnung zu einer integrierten Disposition durch die direkte Bestimmung eines kostenminimalen Warenbestandsniveaus in Abhängigkeit von der Absatzhistorie.

Die vorangegangenen Experimente mit NN zur Prognose auf synthetischen Zeitreihen und zur Disposition begründen die Annahme, dass NN zu einer genauen und zuverlässigen Prognose und Disposition auf empirischen Zeitreihen fähig sind. Da formallogisch aus Einzelfällen oder Vergleichsstudien mit Stichproben von Zeitreihen keine allgemeinen Gesetzmäßigkeiten der grundsätzlichen Überlegenheit eines Verfahrens auf allen möglichen Zeitreihen, Verteilungsformen und in beliebigen Anwendungen abgeleitet werden können (vgl. Abschnitt 5.4), sind die Arbeitshypothesen zur Überprüfung der Prognose- und Dispositionsgüte von NN entsprechend zu formulieren.¹ Der Verfahrensvergleich untersucht somit drei wesentliche Forschungsfragen durch den Versuch einer Widerlegung der Nullhypothesen H_a , H_b und H_c :

H_a : Die Genauigkeit von NN zur Absatzprognose ist geringer als die Genauigkeit statistischer Verfahren.

H_b : Die Genauigkeit von NN zur Warendisposition bei konventioneller Berechnung von Bestellmengen ist geringer als die Genauigkeit statistischer Verfahren.

H_c : Die Genauigkeit von NN zur Warendisposition durch eine simultane Prognose und Bestellmengenrechnung mittels asymmetrischer Kostenfunktionen ist geringer als die Genauigkeit statistischer Verfahren.

Die Falsifizierung von H_c bildet den Kern dieser Arbeit.² Die Studie trägt darüber hinaus in mehrfacher Hinsicht zum Erkenntnisfortschritt bei. Es existieren nur wenige Studien, die die Güte eines Prognoseverfahrens in seinem tatsächlichen Anwendungsproblem beurteilen oder die Interaktion von Prognosefehler und Entscheidungsgüte untersuchen.³ So werden neuartige Verfahren wie die NN üblicherweise mit etablierten Verfahren hinsichtlich ihrer reinen Prognosegüte an Fehlermaßen verglichen, um losgelöst von realen Entscheidungsproblemen generalisierbare Erkenntnisse über die relative Performanz von Prognoseverfahren zu

Standardabweichung der Prognoseresiduen vergangener Absatzperioden, der Bestimmung der Sicherheitsbestände und ihrer Addition zum Prognosewert als zentralem Lageparameter.

¹ Ein statistischer Test kann anhand der Beobachtungen der Experimente aus Prognosefehlern und Dispositionskosten keine grundsätzliche Überlegenheit eines Verfahrens feststellen. Die Nullhypothesen formulieren daher die gegenwärtig vorherrschende Meinung zur Verfahrensklasse der NN in der Prognose, die durch erhebliche Skepsis hinsichtlich der Genauigkeit von NN sowie ihrer grundsätzlichen Fähigkeiten zur Absatzprognose dominiert wird. Dies wird sowohl in wissenschaftlichem Veröffentlichungen zur Prognose mit NN deutlich als auch durch ihre geringe Verbreitung und Anwendung in der Praxis. Der Untersuchungsgegenstand „Prognosegenauigkeit und Dispositionsgüte von NN“ wird als gerichtete Alternativhypothese formuliert, so dass durch eine Falsifikation der Nullhypothese diese wahrscheinlich angenommen werden kann. Insofern kann die Alternativhypothese als innovativ betrachtet werden, da eine Ablehnung der Nullhypothese die Erkenntnisse vorangegangener einflussreicher Veröffentlichungen zumindest für den hier gewählten Datensatz widerlegt.

² Der Vergleich zwischen NN und statistischen Verfahren hinsichtlich ihrer Prognosegenauigkeit (H_a) erfolgt in Abschnitt 6.3.1. Der Vergleich anhand der Kosten der Bestellmengenrechnung bei konventioneller Sicherheitsbestandsbestimmung (H_b) im Vergleich zu MLPs mit asymmetrischen Kostenfunktionen (H_c) erfolgt in Abschnitt 6.3.2.

³ Siehe hierzu Abschnitt 2.3.

erhalten. Die Untersuchung erlaubt jedoch auch die Überprüfung eines grundsätzlichen Zusammenhanges zwischen Prognose- und Dispositionsgüte:

H_d : Die Genauigkeit eines Verfahrens zur Absatzprognose dient als Indikator für die Genauigkeit eines Verfahrens in der Warendisposition

Die Hypothese H_d zielt daher auch auf eine Untersuchung der Frage ab, inwiefern die relative Prognosegüte von NN und statistischen Prognoseverfahren mit ihrer Güte in der Entscheidungsunterstützung der Warendisposition übereinstimmt und ob diese als verlässlicher Indikator verwendet werden kann. Die vorliegende Studie überwindet daher die Einschränkungen vorangegangener Analysen und erweitert diese in mehrfacher Hinsicht.

6.1.2 Aufbau der Vergleichsstudie

Ein sekundäres Ziel dieser Studie ist die Untersuchung der Bedingungen, unter welchen NN bzw. andere Verfahren der Prognose und Disposition gute Ergebnisse zeigen, um nach dem Konzept der mehrfachen Arbeitshypothesen¹ einen größtmöglichen Erkenntnisfortschritt hinsichtlich der Verfahren zu gewährleisten. Diese stammen aus vorangegangenen Ex-ante-Vergleichsstudien über statistische Prognoseverfahren und bestimmen den Versuchsaufbau dieser Studie. Alle Empfehlungen setzten eine Ex-ante-Beurteilung der Prognoseverfahren auf unbekanntem Datenmengen voraus. Die Bestätigung der Erkenntnisse aus vorangegangenen Verfahrensvergleichen erlaubt zudem einen Nachweis der externen Validität der Ergebnisse dieser Studie über den Anwendungsbereich der Disposition an Warenautomaten hinaus und dient damit als Indiz für die Gültigkeit der Experimente. Eine Falsifizierung der etablierten Hypothesen führt hingegen zu einem direkten Wissensfortschritt.

Im Versuchsaufbau sind insbesondere die Erkenntnisse der M-Competitions [Makridakis et al. 1982] für den Datensatz zu überprüfen:²

H_1 : Komplexe Prognoseverfahren sind nicht genauer als einfache Verfahren.

H_2 : Unterschiedliche Fehlermaße führen zu unterschiedlichen Rangordnungen der verschiedenen Prognoseverfahren

Diese Schlussfolgerungen wurden bereits in Abschnitt 3.3.3.1 diskutiert und vielfach durch Vergleichsstudien auf Original-Datensätzen, erweiterten Datensätzen [Fildes 1992a] und in den nachfolgenden M2- und M3-Competitions [Makridakis et al. 1993; Makridakis und Hibon 2000] hinsichtlich ihrer Validität bestätigt. Erkenntnis H_1 postuliert eine vergleichbare Prognosegenauigkeit von komplexen, nichtlinearen Verfahren der NN mit einfachen statistischen Verfahren. Da diese Erkenntnis die Erforschung komplexer Prognoseverfahren grundsätzlich in Frage stellt, erscheint sie als besonders relevant. Daher werden NN gegen mehrere Prognoseverfahren mit zunehmender Komplexität beurteilt, die sowohl empirisch in der Anwendungsdomäne der Warenautomaten als auch theoretisch durch statistische Vergleichsstudien etabliert sind. Während MAKRIDAKIS et al. die Unterscheidung in einfache und komplexe Verfahren auf zeitreihenanalytische Ansätze beschränkte, wird diese hier erweitert. Die Studie betrachtet sowohl univariate Verfahren der Zeitreihenanalyse als auch der multivariaten Interventionsanalyse. Erkenntnis H_2 erfordert die Beurteilung aller Verfahren anhand mehrerer robuster und unverzerrter Fehlermaße und sollte daher durch die vorliegende Studie nicht falsifiziert werden können. Der Versuchsaufbau berücksichtigt dementsprechend

¹ Das Konzept der *multiple working hypotheses* wurde in Abschnitt 3.3.3.1 nach den Richtlinien des International Institute of Forecasters (IIF) sowie des International Journal of Forecasting (IJF) dargestellt.

² Die darüber hinausgehenden Erkenntnisse (vgl. Abschnitt 3.3.3.1), wonach (4.) die Leistung eines Verfahrens von der Länge des Prognosehorizontes abhängt, können hier nicht untersucht werden, da die Disposition an Warenautomaten nur einen konstanten Prognosehorizont von $h=1$ erlaubt. Auf eine Untersuchung der Erkenntnis, dass (3.) kombinierte Prognosen genauer sind als individuelle Prognosen, wird aufgrund der wachsenden Komplexität der Kombination aller Prognoseverfahren hier verzichtet.

mehrere statistische Fehlermaße, die in vorherigen Vergleichsstudien etabliert wurden. Im Gegensatz zu den Vergleichsstudien der M-Competitions, die unabhängig von einer Anwendungsdomäne durchgeführt wurden, liegt hier ein konkretes betriebswirtschaftliches Entscheidungsproblem in Form der Disposition an Warenautomaten vor. Entsprechend können die Fehler- und Kostenmaße durch die Anwendungsdomäne motiviert und begründet werden, was die Validität des Experimentes für die Anwendungsdomäne verbessert, die Komplexität des Experimentdesigns jedoch erhöht. Die Beurteilung der Ergebnisse der Warendisposition erfolgt anhand der verursachten Kosten der Dispositionsentscheidungen in Form von Präsenzlücken und Überbevorratungen.¹

Durch Folgestudien [Fildes et al. 1998] wurden weitere Erkenntnisse etabliert, die ebenfalls im Versuchsaufbau zu berücksichtigen sind:

- H₃: In der Bestimmung von Prognosegüte und Dispositionskosten der einzelnen Prognoseverfahren spielen die Charakteristika der Zeitreihe eine wichtige Rolle.
- H₄: Vergleiche zwischen Verfahren sollten aus mehreren Zeitreihen-Ursprüngen heraus durchgeführt werden, um Fehler durch die Variabilität der Stichprobenziehung zu verhindern.

Nach H₃ sind die einzelnen Zeitreihen hinsichtlich ihrer Eigenschaften zu charakterisieren, und die Prognosegenauigkeit der Verfahren nach Kategorien zu untersuchen. Die Zeitreihen können für die Vergleichsstudie an Warenautomaten hinsichtlich ihrer Charakteristika nach dem Standort der Absatzstelle in Innen- oder Außenautomaten, nach der Zeitreihenstruktur in sporadische, unregelmäßige oder regelmäßige Absatzverläufe, nach der Umschlagsgeschwindigkeit in Schnelldreher oder Langsamdreher sowie nach der Umsatzhöhe gemäß ABC-Klassifikation unterschieden werden. Alle Ergebnisse des Verfahrensvergleichs werden entsprechend sowohl über alle Zeitreihen als auch für jede Kategorie der Zeitreihenstruktur überprüft.²

Nach Erkenntnis H₄ ist eine rollierende Beurteilung der Verfahren anhand der Prognosefehler und der Dispositionskosten über mehrere Zeitreihenursprünge (engl. *multiple rolling origins*) notwendig, um generalisierbare Ergebnisse zu erhalten. Der empirische Planungshorizont der Disposition an Warenautomaten im Newsboy-Modell erfordert Entscheidungen nur für die Folgeperiode, mit einem Prognosehorizont von $h = 1$. Somit werden in der Vergleichsstudie mehrfache $t+1$ Prognosen von n unterschiedlichen Prognoseursprüngen $t, t+1, t+2, \dots, t+n$ durchgeführt.³

FILDES [Fildes 1989] untersucht die Komplexität der Prognoseverfahren gemäß H₁ näher und unterscheidet Prognoseverfahren mit individueller Modellselektion je Zeitreihe von solchen mit aggregierter Modellselektion, in der ein einzelnes Prognoseverfahren auf alle Zeitreihen eines Datensatzes angewendet wird, zum Beispiel im Fall der ersten M-Competition.⁴ Er unterteilt die Verfahren mit aggregierter Modellselektion weiter in Ansätze mit festen

¹

² Die empirisch abgeleiteten Kriterien entsprechen einer statistischen Charakterisierung der Verteilungsform der Zeitreihe durch die zentralen Lageparameter (das Niveau der Zeitreihe als durchschnittlicher Absatz als 2 x als A-, B- oder C-Artikel) und den Grad der Zufallsschwankung durch die Zeitreihenstruktur anhand des Variationskoeffizienten.

³ Durch den festgelegten Prognosehorizont vereinfacht sich auch die Berechnung der Fehlermaße über unterschiedliche Prognoseursprünge.

⁴ FILDES [1989] bezeichnet sie als „aggregate modell selektion“ und unterstellt damit die bewusste Ex-ante-Wahl eines Modells zur Anwendung auf einer Vielzahl von Zeitreihen. Da somit ein ausgewähltes Prognoseverfahren zur Prognose auf allen Zeitreihen angewendet wird, kann es auch als Prognose ohne Modellselektion betrachtet werden. Um Verwechslungen einer Auswahl eines einzelnen Prognoseverfahrens durch Expertensysteme zu vermeiden, wird ein Verfahren mit aggregierter Modellselektion nachfolgend als Ansatz „ohne Modellselektion“ bezeichnet.

Parameterwerten auf allen Zeitreihen und in Ansätze mit Parameteroptimierung für die jeweilige Zeitreihenstruktur. Bei individueller Modellselektion wird hingegen durch eine konsistente Methodologie je Zeitreihe eine passende Verfahrensklasse und Modellform ausgewählt und für die jeweilige Zeitreihe parametrisiert.¹ Die individuelle Modellselektion wird durch moderne Software-Expertensysteme wie ForecastPro angewendet und bewies eine hohe Prognosegenauigkeit in der M3-Competition [Makridakis und Hibon 2000; Goodrich 2000], jedoch bei gesteigerten Anforderungen an Computer-Ressourcen. In Anlehnung an FILDES [1989] können somit folgende Hypothesen formuliert werden:

H₅: Prognoseverfahren mit individueller Modellselektion (Expertensysteme) sind nur bei längeren Prognosehorizonten genauer als Ansätze mit aggregierter Modellselektion.

H₆: Verfahren mit aggregierter Modellselektion und Parameteroptimierung sind genauer als Verfahren mit konstanten Parametern.

Die semi-parametrischen NN sind hierbei als Prognoseverfahren mit aggregierter Modellselektion und Parameteroptimierung zu betrachten, da sie im Training keinen dedizierten Prozessschritt der Modellselektion beinhalten. Liegen nur kurze Prognosehorizonte vor, wie in der vorliegenden Vergleichsstudie zur $t+1$ Prognose, so sollten Expertensysteme oder NN nach H₅ keine höhere Genauigkeit aufweisen als einfache Ansätze ohne Modellselektion. Die Nullhypothese H₆ postuliert ferner, dass NN und andere Verfahren mit Parameteroptimierung eine höhere Genauigkeit aufweisen als Verfahren wie die einfache Exponentielle Glättung mit konstantem α -Parameter ohne Parameteroptimierung, was in der Studie zu beurteilen ist. Die Auswahl der statistischen Prognoseverfahren und der NN erfolgt in Abschnitt 6.2.3. Die Verfahren zur Disposition einschließlich der MLP mit asymmetrischen Kostenfunktionen werden in Abschnitt 6.2.4 dargestellt, da dieses MLP im eigentlichen Sinne nicht länger ein Prognoseverfahren, sondern vielmehr ein Verfahren zur Disposition darstellt.

Darüber hinaus sind die Erkenntnisse zur Durchführung von Verfahrensvergleichen nach TASHMAN [2000] zu berücksichtigen. Die Bildung einer repräsentativen Stichprobe von Zeitreihen der Anwendungsdomäne wird in Abschnitt 6.2.2 dargestellt und die Charakteristika der einzelnen Zeitreihen diskutiert. Die Auswahl etablierter Benchmark-Prognoseverfahren wird in Abschnitt 6.2.3 diskutiert und die Auswahl unverzerrter, robuster Gütekriterien durch statistische Fehlermaße und monetäre Kostengrößen in Abschnitt 6.2.5 begründet. Sie werden detailliert untersucht, um die Auswahl der Verfahren und Gütekriterien hinreichend zu legitimieren und um eine Replikation der Experimente im Sinne der geforderten Reliabilität von Verfahrensvergleichen zu gewährleisten.

Während die Empfehlungen zur Gestaltung von Vergleichsstudien von Prognoseverfahren durch die intensive Diskussion im Anschluss an die M-Competitions hervorgebracht wurden, existieren nur wenige Empfehlungen zur Durchführung empirischer Vergleichsstudien zur Ex-ante-Beurteilung von NN, die bei der Berechnung valider und reliabler Experimente zu berücksichtigen sind. Wenngleich eine allgemein akzeptierte Methodologie im Sinne positiver Handlungsempfehlungen zur Modellierung von NN noch aussteht, sollten zumindest Maßnahmen in der Anwendung von NN, die zum Entstehen gravierender Fehler und invalider Experimente führen, vermieden werden. Nur wenige Studien geben derartige Hinweise. ADYA und COLLOPY [1998] geben Empfehlungen zur Berechnung valider und reliabler NN-

¹ Die feste Parameterwahl entspricht zum Beispiel der Verwendung der einfachen Exponentiellen Glättung mit einem konstanten Glättungsparameter von $\alpha=0,3$ auf allen Zeitreihen, mit optimierter Parameterwahl durch Bestimmung des optimalen α zur Anpassung an die Struktur der jeweiligen Zeitreihe. Dennoch wird nur ein Prognoseverfahren angewendet. Bei individueller Modellselektion erfolgt darüber hinaus die Identifikation eines trend-saisonalen Zeitreihenverlaufs, die Auswahl der Exponentiellen Glättung nach HOLT-WINTERS und die Bestimmung optimaler Parameter für die Zeitreihe.

Experimente mit vorbildlichem Experimentaufbauten zum Verfahrenvergleich [Liao und Fildes 2005; Hill et al. 1996]. Sie wurden durch CRONE und PREBMAR [2006] in ein einheitliches Bewertungsschema von Verfahrensvergleichen mit NN integriert, und werden im Verfahrensvergleich befolgt (siehe Abschnitt 6.2.3.3).

Die vorangehend dargelegten Zielsetzungen und die etablierten Gestaltungskriterien zur Durchführung empirischer Verfahrensvergleiche bestimmen die Struktur und den Aufbau der Experimente in der Vergleichsstudie. Die vorliegende Studie folgt einem Experimentansatz multipler Arbeitshypothesen, der alle Ergebnisse eines empirischen Verfahrensvergleichs hinsichtlich unterschiedlicher Charakteristika von Verfahren, Zeitreihen und Gütekriterien kritisch diskutiert. Die Studie berücksichtigt so die wissenschaftlichen Erkenntnisse und Gestaltungskriterien für valide, reliable und objektive empirische Vergleichsstudien von Prognoseverfahren zur Ableitung gültiger wissenschaftlicher Erkenntnisse. Sie erweitert nicht nur die Theorie von NN zur simultanen Prognose und Disposition durch asymmetrische Kostenfunktionen, sondern überprüft auch die Ergebnisse vorheriger Prognosestudien von NN und erweitert den gegenwärtigen Erkenntnisstand hinsichtlich der relativen Genauigkeit von statistischen und neuronalen Prognoseverfahren.

6.2 Durchführung der Vergleichsstudie

6.2.1 Problemstellung

6.2.1.1 Der Warenautomat als Betriebsform des Handels

Zunächst wird der Untersuchungsgegenstand der Problemstellung charakterisiert. Waren- und Verkaufsautomaten (engl. *vending machines*) sind eine Betriebsform des institutionellen, stationären Einzelhandels ohne Verkaufsraum, die sich anhand der Größe der Verkaufsfläche, des Bedienungsprinzips, der Art des Inkassos, der Sortimentspolitik und der Preispolitik von anderen Betriebsformen des Groß- und Einzelhandels unterscheidet [Müller-Hagedorn 1998].¹ Mit Ausnahme von SCHENK [1988] und TIETZ [1992; 1993a; 1993b; 1994] werden Warenautomaten in Handelsforschung, Marketing, Logistik sowie Operational Research über eine Einordnung in Betriebstypologien oder Systematiken von Distributionskanälen hinaus vernachlässigt. Entsprechend bezeichnet SCHENK den Warenautomaten als „vergessene Betriebsform des Handels“ [Schenk 1988, S. 3]. Die Untersuchung der betriebswirtschaftlichen Entscheidungsprobleme der Warendisposition erfordert daher zunächst eine kurze Darstellung des Warenautomaten, seiner Bedeutung und seiner logistischen Besonderheiten.

Ein Automat bezeichnet allgemein eine maschinelle Vorrichtung, die bestimmte Funktionen selbsttätig mechanisch, elektrisch, pneumatisch oder hydraulisch ausführt [Schenk 1988, S. 11].² Der Warenautomat repräsentiert eine besondere Form des Automaten. Er reagiert auf den Auslöseimpuls der Wahl und der Bezahlung durch den Käufer mit der selbsttätigen Herausgabe der Ware. MÜLLER-HAGEDORN bezeichnet diesen Vorgang als Automatenverkauf, „in dem neben der Warenausgabe auch das Inkasso von einer technischen Anlage übernommen wird“ [Müller-Hagedorn 1998, S. 43].³ Der Absatz erfolgt somit durch unper-

¹ Die Unterscheidung in Betriebsformen und Betriebstypen erfolgt uneinheitlich, siehe [Müller-Hagedorn 1998; Ausschuss für Begriffsdefinitionen aus der Handels- und Absatzwirtschaft 1995]. Auf eine Diskussion wird hier verzichtet.

² Der Begriff leitet sich aus der griechischen Bezeichnung *automatos* für etwas, das sich selbst bewegt, ab. Im Englischen stammt der Begriff des *vending* aus dem Lateinischen „*vendere*: to sell“ oder „*venum dare*: to give for sale“ [Trident Press International. 2003, S. 1154]. Die am Verkaufsprozess Beteiligten sind der Verkäufer (engl. *vendor*) und der Käufer oder Leistungsempfänger (engl. *vendee*).

³ Die Zuordnung des Warenautomaten zu den Betriebstypen des Handels ist nicht eindeutig geklärt. MEFFERT bezeichnet ihn als Distributionskanal [Meffert 2000, S. 826], andere Autoren als Verkaufsmethode des Einzelhandels [Arentzen 1993, S. 380], als Absatzorgan des Herstellers [Nieschlag und Kuhn 1980] oder als ladenlosen Einzelhandel mit Verkauf durch Selbstbedienung [Nieschlag et al. 2002, S. 443]. Eine eindeutige Zuordnung erscheint aufgrund der zahlreichen Facetten des Einsatzes von Warenautomaten nicht möglich.

sönlichen, zeitlich ungebundenen Verkauf an einem neutralem Ort [Ahlert 2002, S. 28; Nieschlag et al. 2002, S. 442 ff.], indem der Verkaufsvorgang alle Komponenten des Geld- und güterwirtschaftlichen Austausches durch unmittelbares Angebot, Absatz und Inkasso beinhaltet [Hadelar 2000, S. 380]. Durch den Absatz von nicht selbst be- oder verarbeiteten Gütern an Dritte liegt Handel im funktionalen Sinne vor [Ausschuss für Begriffsdefinitionen aus der Handels- und Absatzwirtschaft 1995].¹ Entsprechend wird der Absatz von Handelswaren über Automaten auch als automatischer (Einzel-)Handel bezeichnet: „Vending means automatic retailing. ... it covers the sales of products and services through a machine operated by introducing ... means of ... payment“ [European Vending Association 2004, S. 1], oder als „Retailing through an Unattended Point of Sale (UPoS)“ [NAMA 2006].

Warenautomaten sind von anderen Automatentypen abzugrenzen, die keine Betriebsformen des Handels darstellen. Grundsätzlich wird zwischen Unterhaltungsautomaten und Waren- bzw. Dienstleistungsautomaten unterschieden, durch Betriebe zur Aufstellung von Unterhaltungsautomaten trotz vergleichbarer logistischer Prozesse der Automatenbefüllung, -entleerung und -wartung nicht dem Handel, sondern der Freizeitwirtschaft zuzuordnen sind [Informationsgemeinschaft Münzspiel 2001, S. 5].² Während Warenautomaten Handelswaren absetzen, z. B. durch Zigarettenautomaten, Getränke- oder Verpflegungsautomaten, wird die Wertschöpfung bei Dienstleistungsautomaten durch den Service dominiert, z. B. bei Automaten zum Park- und Fahrscheinverkauf, öffentlichen Telefonen, Maut-Stationen oder Geldwechselgeräten. Aufgrund der ähnlichen warenlogistischen Prozesse können Geldautomaten ebenfalls als Warenautomaten betrachtet werden, auch wenn sie nicht zum Handel gezählt und in keiner Statistik erfasst werden. In Abbildung 6.1 ist eine Systematik zur Einteilung der Automatentypen abgebildet:

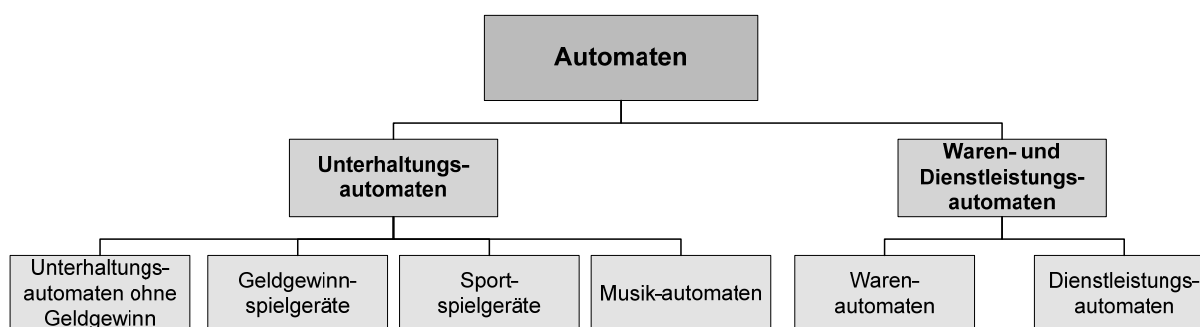


Abbildung 6.1: Systematik zur Einteilung von Automatentypen

Die grundsätzliche Idee, Waren durch Maschinen zu verkaufen, geht bis in die Antike zurück. Als erstes dokumentiertes Zeugnis beschrieb der griechische Mathematiker HERON VON ALEXANDRIA [Hero und Baldi 1589] um 150 v. Chr. einen Mechanismus zum unbeaufsichtigten, vollständig automatisierten Verkauf von heiligem Wasser in ägyptischen Tempeln nach dem Einwurf einer Münze [Schreiber 1961, S. 9; ebenso 1999]. Im 2. Jahrhundert vor Christus erwähnt PHILON VON BYZANZ ein Gerät, „das nach Einwurf einer Münze einen Bimsstein zum Händewaschen herausgibt – ein antiker Seifenautomat“ [von Kerssenbrock-

Der Warenautomat wird nachfolgend als Betriebsform des Handels zum Verkauf von Handelswaren betrachtet.

¹ Auf eine detaillierte Diskussion, ob der Absatz über Warenautomaten zum Einzelhandel im institutionalen Sinne gehört, oder ob die Tätigkeit der Automatenausstellung und -befüllung eher als Dienstleistung eines „Rack-Jobbers“ in fremden Verkaufsstellen einzuschätzen ist, wird hier nicht eingegangen. Die Definitionen in der Praxis uneinheitlich, bzw. existieren nicht aufgrund der geringen Betrachtung in der Theorie.

² Unterhaltungsautomaten ohne Geldgewinnmöglichkeit umfassen u.a. Flipper, Punktespiele und Videospiele, bei denen der Spieler über seine Geschicklichkeit den Spielverlauf entscheidend beeinflusst. Sportspielgeräte umfassen dagegen Billard, Dart, Tischfußball usw. Das Angebot von „Unterhaltung“ wird somit nicht als Dienstleistung im eigentlichen Sinne aufgefasst, obwohl es sich um einen Service handelt, der beim Konsum verzehrt wird [Vieweg 1999, S. 4].

Krosigk 1998]. Den Zeitpunkt der ersten bekannten kommerziellen Nutzung des Prinzips des unbeaufsichtigten Verkaufs datiert SCHREIBER auf das 18. Jahrhundert in England und in den amerikanischen Kolonien, und zwar in Form von münzausgelösten „honor boxes“ zum Verkauf von Pfeifen- und Schnupftabak [Schreiber, in Encyclopaedia Britannica inc. 2005].¹ Der Verkauf von Waren über Automaten spielt jedoch erst seit der zweiten Hälfte des 19. Jahrhunderts eine wirtschaftliche Rolle. Im Frühjahr 1887 nahm die Firma Stollwerck die ersten Süßwaren-Automaten in Deutschland zum Verkauf von Warenproben von Schokolade in Betrieb [Schenk 1988, S. 5]. Im selben Jahr stellte Thomas Adams in New York seinen ersten Verkaufsautomaten (engl. *vending machine*) für Kaugummis auf [Larson et al. 1976, S. 43]. Die positive Reaktion der Konsumenten führte in der Folge zunehmend dazu, Automaten nicht nur als kuriose Form der Marketingkommunikation, sondern als einen eigenständigen Vertriebsweg für Waren anzusehen. Bereits 1895 waren im Deutschen Reich über 25.000 hölzerne Warenautomaten aus 85 Typen zum Verkauf unterschiedlichster Produkte aufgestellt. Neben Schokolade und Bonbons wurden auch Stecknadeln, Zahnstocher, Briefmarken, Toilettenpapier, Parfüm, Bücher, Tabakwaren und Bleistifte über Automaten vertrieben [Koopmann 1988, S. 67]. Erst später gewannen Automaten zur Unterhaltung und zum Spiel an Verbreitung [Haberbosch und Volkmann 1994].² Im Jahr 1901 wurden Warenautomaten der Betriebsform einer offenen Verkaufsstelle zugeordnet und unterlagen damit den gesetzlichen Ladenschlusszeiten. Erst 1934 wurde diese Zuordnung durch das „Gesetz über den Verkauf aus Automaten“ aufgehoben. Gleichzeitig wurde die Residenzpflicht eingeführt, die den Betrieb von Warenautomaten nur in Verbindung mit einem Ladengeschäft im Ladeninneren zuließ, welche erst im Jahr 1956 im Rahmen des neuen Ladenschlussgesetzes wieder aufgehoben wurde [Koopmann 1988, S. 68]. Historisch betrachtet geht der Ursprung des unpersönlichen Verkaufs somit über die Anfänge der Selbstbedienung im stationären Einzelhandel bis zum Absatz durch Warenautomaten zurück.

Das empirische Erscheinungsbild des Vertriebs über Warenautomaten zeigt mit wenigen Ausnahmen im ersten und zweiten Weltkrieg sowie während der Weltwirtschaftskrise ein ungebremses Wachstum bis in die 1990er Jahre. Im Jahr 2004 waren in Deutschland mehr als 1,2 Millionen Warenautomaten aufgestellt, davon ca. 750.000 Zigarettenautomaten und 500.000 Getränke- und Verpflegungsautomaten, die insgesamt einen Umsatz von ca. 11 Mrd. Euro erzielten [EVA 2004; BDTA 2005].³ Im Jahr 2005 erwirtschafteten die 600.000 Zigarettenautomaten in Deutschland 23% des Absatzes aller Fabrikzigaretten, mit einem Umsatz von 4,5 Mrd. Euro [BDTA 2005].⁴ In Tabelle 6.1 wird die Anzahl der Verpflegungsautomaten ohne Zigarettenautomaten in Europa, die Anzahl neu aufgestellter Geräte und die Automatendichte je 1000 Einwohner abgebildet.

¹ Da diese Geräte lediglich die Warenentnahme freigaben und nicht portionierten, oblag die Verkaufsmenge sowie das Schließen des Gerätes der Ehrlichkeit des Konsumenten.

² Die Übergänge zum Warenautomaten waren fließend, da ein am Automat erzielter Gewinn vom Wirt der Gaststätte dem Gast in Form von Waren ausgezahlt werden musste [Haberbosch und Volkmann 1994, S. 5].

³ Die Validität der Zahlen ist nicht zu belegen. Sie beruhen auf Schätzungen und Umfragen des zuständigen Bundesverbandes der Deutscher Verpflegungs- und Vending-Unternehmen (BDV) und sind allgemein akzeptiert [Monßen 2001]. Zusätzlich sind im Jahr 2006 in Deutschland 132.000 Unterhaltungsautomaten ohne Geldgewinnmöglichkeit, 200.000 Geldgewinnspiele und 55.500 Sportspielgeräte aufgestellt, die bei rückläufigen Zahlen durch Verbote und Veränderungen der Steuergesetzgebung einen Umsatz von 3,8 Mrd. Euro erwirtschaften [Vieweg 2006, S. 9].

⁴ Sie zeigen einen deutlichen Rückgang von 820.000 Geräten im Jahr 1998 durch Änderungen der rechtliche Rahmenbedingungen des Jugendschutzes unter Einführung von Geldkarten mit Altersnachweis.

Tabelle 6.1: Statistik der Verpflegungsautomaten in Europa (ohne Zigarettenautomaten)¹

	Warenautomaten Bestand 2004	Warenautomaten Neuaufstellung 2004	Warenautomaten pro 1000 Einwohner
Italien	613.650	78.599	10,6
Frankreich	593.000	34.776	9,9
Großbritannien	510.911	42.718	8,6
Deutschland	502.000	35.956	6,1
Spanien	278.800	34.148	6,7
Niederlande	200.000	25.273	12,3
Schweden	88.000	11.130	9,8
Schweiz	81.500	9.176	11,1
Österreich	76.080	9.225	9,4
Portugal	50.650	8.623	4,8
Ungarn	33.000	2.211	3,3
Tschechische Rep.	12.000	2.425	1,2
Slowakei	7.070	1.273	1,3
Polen	7.000	2.328	0,2
Gesamt Europa	3.053.661	297.861	

Im Vergleich des Umsatzes von 15,5 Mrd. Euro zu den 412.351 Unternehmen im deutschen Einzelhandel, die im Jahr 2004 Lieferungen und Leistungen in Höhe von 431 Mrd. Euro erwirtschafteten [Statistisches Bundesamt Deutschland 2006a], kann somit festgestellt werden, dass der Verkauf über Warenautomaten national wie international eine nicht unbedeutende Betriebsform des Handels darstellt die in einzelnen Produktkategorien einen erheblichen Anteil des Handelsumsatzes realisiert.²

Warenautomaten zeichnen sich im Vergleich zu anderen Betriebsformen des Einzelhandels durch eine vergleichsweise geringe Sortimentsbreite und Sortimentstiefe bei geringem Umsatz aus. Sie dienen überwiegend zur ubiquitären Nahverpflegung in den Produktkategorien der „5C’s: Coffee, Cold drinks, Candy & snacks, Cold Food, Cigarettes“ [NAMA 2005] an Standorten des öffentlichen Verkaufs (engl. *public vending*) mit hoher Kundenfrequenz in Hotels, Universitäten, Schulen, Krankenhäusern, Freizeitzentren, Bahnhöfen usw. sowie in Unternehmen (engl. *private vending*). Der Umsatz von Warenautomaten an besonders hochfrequenten Standorten erreicht die der unteren Umsatzgrößenklassen der amtlichen Statistik des Statistischen Bundesamtes von €17.500. Somit erzielen Automaten durch den hohen Warenumschlag und die geringe Nettofläche von weniger als 2 Quadratmetern im Vergleich zu anderen Absatzstellen eine hohe Flächenproduktivität.³

Die Marktteilnehmer umfassen die Betreiber der Warenautomaten (engl. *Operator*), die als Automatendulder bezeichneten Eigentümer der Stellplätze der Automaten, die Hersteller der Warenautomaten und der Automatenfüllprodukte und die Konsumenten als Endkunden. Die besonderen logistischen Anforderungen durch die weite räumliche Verteilung der Absatzstätten mit geringem Umsatz führte zu einer Spezialisierung der Handelsformen als Transaktionskostenspezialisten [Müller-Hagedorn 1998, S. 124]. So entstanden spezialisierte Handelsbetriebe zur Automatenbewirtschaftung, zum Beispiel in Form der Firma Tobacco-land Automatengesellschaft „mit 2,2 Mrd. DM Umsatz und 1.700 Mitarbeitern als reine

¹ Eigene Darstellung aus Daten der EVA [2004].

² In den USA und Japan geht die Bedeutung der Warenautomaten noch darüber hinaus. In den USA wurde 2004 an Warenautomaten 30 Mrd. US-Dollar Umsatz erwirtschaftet [NAMA 2005], in Japan nahezu vergleichbar viel. Alternativ kann aufgrund der Rolle der Warenautomaten bei der Nahverpflegung auch ein Vergleich mit der Gastronomie angestellt werden, in der 198.649 Betriebe 34,4 Mrd. Euro erwirtschafteten [Statistisches Bundesamt Deutschland 2006a]. Die Rolle der Warenautomaten erscheint so noch bedeutender, dies widerspricht jedoch ihrer Betrachtung als Betriebsform des Handels.

³ Zigarettenautomaten erzielen so auf einer Stellfläche von weniger als 0,5 Quadratmetern eine höhere Flächenproduktivität als viele Facheinzelhandlungen mit geringem Umsatz [Müller-Hagedorn 1998].

Automatengesellschaft zum Vertrieb von Zigaretten über 220.000 Warenautomaten aus 14 Regionallagern“ [Eurohandelsinstitut 1999, S. 15].¹

6.2.1.2 Prognose und Disposition an Warenautomaten

Die Prognose und Disposition an Warenautomaten wird durch die Struktur determinanten ihres Logistiksystems bestimmt (vgl. Abschnitt 2.1). Warenautomaten (ATM) sind räumlich verteilte Einzelhandelsfilialen und werden im Sinne der Logistik als Lager des Warenvorrats direkt im Verkaufsraum (VR) aufgefasst. Die Warenbefüllung erfolgt aus einem Warenlager im Tourenfahrzeug (TOUR), das seinerseits periodisch in einem Regionallager (RL) aufgefüllt wird und dort das kassierte Geld einreicht. Je nach der Art und Größe des Automaten, der Art der Waren und der Zusammensetzung des Sortimentes variieren die Anfahrfrequenz, die Dispositions- und Befüllungsdauer und somit die Anzahl der Tourenziele je Tag.² In Abbildung 6.2 ist ein typisches, divergentes Lagersystem der Warenautomatenbefüllung abgebildet.

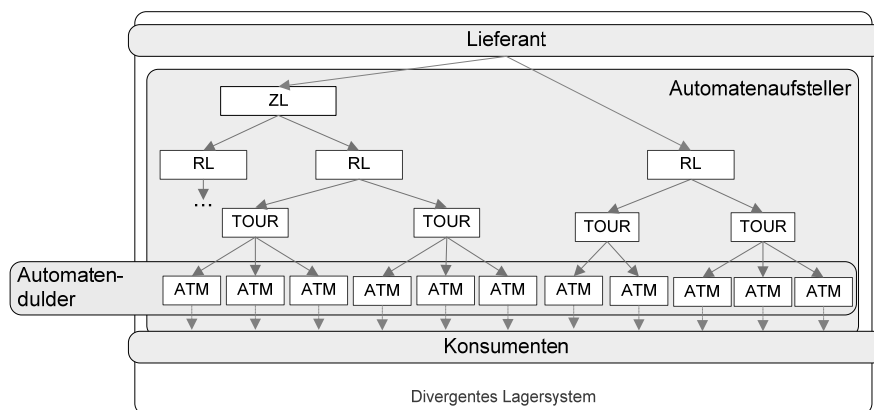


Abbildung 6.2: Lagersystem der Warenautomaten

Durch die räumliche Distanz zwischen Warenautomat und Handelszentrale ohne laufende Kommunikation zur Überwachung der Abverkäufe sind Absatz und Bestand eines Automaten bis zum physischen Aufsuchen durch den Tourenfahrer unbekannt.³ Eine Gruppe von Warenautomaten wird nacheinander auf einer Auslieferungstour angefahren, wobei jeweils der Warenabfluss ermittelt, das kassierte Geld entnommen und der Automat mit neuer Ware befüllt wird.⁴ Der Dispositionsprozess erfolgt somit je Warenautomat einschließlich Wareneinlagerung und Inkasso zeitlich und räumlich konzentriert während einer Belieferung. Die Disposition an Warenautomaten wird nach einer Unternehmensstudie von ROLAND BERGER in den Prozess der Automatenbefüllung eingeordnet, der in Abbildung 6.3, angelehnt an die Teilschritte der Disposition im Handel aus Abschnitt 2.1, dargestellt ist.

¹ Durch Initiativen des Jugendschutzes wurde die Anzahl der Automaten und Mitarbeiter in 2006 abgebaut.

² Zigarettenautomaten werden üblicherweise in wöchentlichem oder mehrwöchentlichem Lieferrhythmus und auf einer Tour von ca. 45 Automaten pro Tag angefahren, während ca. 15 Kaffeeautomaten pro Tour alle zwei Tage angefahren werden, mit dem Erfordernis zusätzlicher Dienstleistungen an der Absatzstelle wie der Reinigung des Automaten usw.

³ Moderne Automaten zeichnen sich durch mehrere Komponenten der Informationstechnologie in der Sicherheitstechnik, in bargeldlosen Zahlungssystemen bzw. einer vollelektronischen Datenerfassung bis hin zur Infrarot- oder Funkauslesung der Abverkaufdaten aus. Auch wenn eine kontinuierliche Kommunikation und somit eine Anbindung von Warenautomaten an die Handelszentrale seit einigen Jahren technisch realisierbar wäre, unterbleibt sie vielfach aus Gründen der Wirtschaftlichkeit.

⁴ Eine Absatzstelle kann dabei auch mehrere Warenautomaten umfassen.

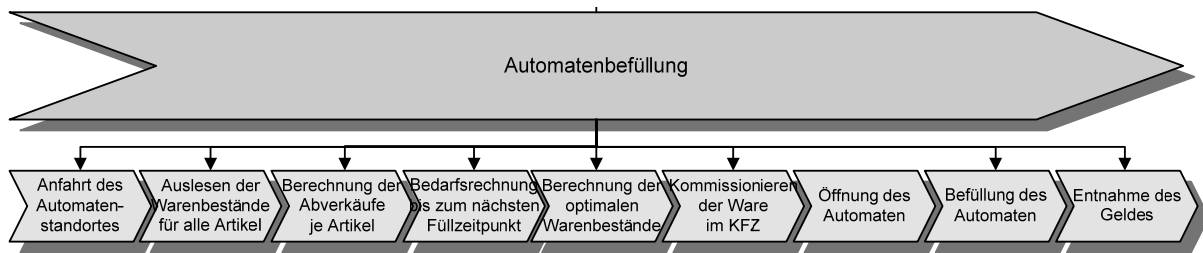


Abbildung 6.3: Prozess der WarenDisposition und Befüllung an Warenautomaten

Im Anschluss an das Anfahren eines Warenautomaten erfolgt zunächst die Auslesung der Abverkäufe seit der letzten Anfahrt im Sinne der Dispositionsperiode. Dies erfolgt üblicherweise elektronisch unter Nutzung von Geräten der mobilen Datenerfassung (MDE).¹ Anhand der erfassten Warenausgänge der vergangenen Periode wird der gegenwärtige Warenbestand ermittelt und die Zeitreihe der vergangenen Absatzhistorie fortgeschrieben, woraus anschließend die Absatzprognose für die nächste Dispositionsperiode berechnet werden kann.

Die Bedarfsprognose erfolgt üblicherweise durch das MDE-Gerät², welches durch die integrierte Berechnung eines Prognosewertes und eines Sicherheitsbestandes die Ausgabe einer empfohlenen Warenbestandshöhe für die nächste Füllperiode anzeigt. Diese wird anschließend durch den Automatenbefüller überprüft und entweder akzeptiert oder auf Basis der subjektiven Erfahrungswerte und der Kenntnis zukünftiger Ereignisse an der Absatzstelle manuell angepasst, im Sinne eines kombinierten Prognoseprozesses objektiver Berechnung und subjektiver Korrektur. Trotz der erheblichen Rechenkapazität von MDE-Geräten werden üblicherweise einfache Prognoseverfahren basierend auf gleitenden Mittelwerten oder Naiven Verfahren eingesetzt. Die Berechnung der Sicherheitsbestände erfolgt vielfach unter Vernachlässigung der Standardabweichung vergangener Prognosefehler als absolute Menge oder prozentual vom durchschnittlichen Absatzniveau, was zu gravierenden Ineffizienzen in den Warenbeständen führen kann.³

Im Anschluss wird aus dem aktuellen und dem berechneten Warenbestand die Menge der einzufüllenden Ware für die nächste Dispositionsperiode berechnet, die anschließend im Tourenfahrzeug kommissioniert und in den Automaten unter Entnahme des Geldes eingelagert wird. Im Anschluss wird der nächste Warenautomat angefahren und befüllt. Jede Automatenbefüllung umfasst also eine vollständige WarenDisposition aus Bedarfs- oder Absatzprognose und Bestellmengenrechnung und darüber hinaus die Bestellung und den gleichzeitigen Wareneingang. Die zeitliche Konzentration dieser Vorgänge erlaubt die Ableitung vereinfachter Modelle der Bestellmengenrechnung, die nachfolgend dargestellt werden.

¹ Vereinzelt erfolgt auch eine rein manuelle Sichtdisposition ohne Nutzung von Informationstechnologie, die entsprechend zur Anwendung Naiver Verfahren tendiert. Sie wird hier vernachlässigt und hat keinen Einfluss auf die nachfolgenden Erweiterungen.

² Vielfach werden zur manuellen, subjektiven Sichtdisposition die vergangenen Absatzzahlen zur Entscheidungsunterstützung hinzugezogen, nicht jedoch die vergangenen Prognosefehler zur Berechnung der Sicherheitsbestände. Empirische Erhebungen haben hinreichend bewiesen, dass dies zu ineffizienten Warenbeständen führt. Alternativ erfolgt eine rein subjektive Bestimmung der Warenbestände durch den Automatenbefüller, unter Anwendung Naiver Heuristiken und ohne explizite Trennung der Berechnung von Prognosewerten und Sicherheitsbeständen.

³ Die Verwendung eines Vielfachen des Mittelwertes ist zum Beispiel nur bei poissonverteilterm Bedarf für Waren mit langsamem Warenumschlag zulässig, nicht aber unter der Annahme normalverteilter Bedarfe. Darüber hinaus sind vielfach heuristische Regeln der Bestandsberechnung implementiert, z. B. die Berücksichtigung von Mindestbeständen, und deterministischen Sicherheitsbeständen einer zusätzlichen Anzahl von Verkaufseinheiten, die zur Prognose addiert werden. Dies führt dementsprechend zu einer Verzerrung und Unterbevorratung bei stark schwankenden Artikeln mit einem Variationskoeffizienten >1 und einer Überbevorratung bei weniger schwankenden Artikeln mit einem Variationskoeffizienten ≤ 1 .

6.2.1.3 Modelltheoretische Einordnung der Problemstellung

Warenautomaten vereinfachen die warenwirtschaftliche Analyse der Dispositionsprozesse und -modelle im Handel aufgrund der geringen Sortimentstiefe und -breite sowie aufgrund des automatisierten, unpersönlichen Verkaufs unter Ausschluss menschlicher Einflüsse auf Einstellungen und das Verhalten im Kaufprozess. Sie bieten somit im Gegensatz zu anderen Betriebsformen des Handels ein weitgehend kontrolliertes Experimentdesign.

Darüber hinaus erlauben die Struktur determinanten der Lagerhaltung eine Abbildung durch einfache Bestellmengenmodelle. Die Warendisposition für die vollständige Logistikkette der Warenautomaten entspricht einem mehrstufigen, mehrperiodigen Mehrprodukt-Modell mit stochastisch-instationärem Bedarf, Wiederbeschaffungszeiten und Liefermengen unter Ressourcenrestriktionen, wie nahezu alle realen Logistiksysteme [Arnolds et al. 2001, S. 59-60]. Derartig komplexe realitätsgetreue und divergente Lagerhaltungssysteme sind nicht analytisch lösbar. Sie können jedoch häufig in einfachere Modelle dekomponiert werden.¹

Die Auswahl zutreffender Modelle der Bedarfsrechnung und der Bestellmengenrechnung zur Disposition erfordert zunächst die Charakterisierung der Bedarfs- und Bestellstruktur. Die Konstituenten des Lagerhaltungsproblems bei der Disposition an einem einzelnen Warenautomaten mit einem Lieferanten und mehreren Kunden sind in Abbildung 6.4 nach SCHNEEWEIß [1981, S. 7] dargestellt.

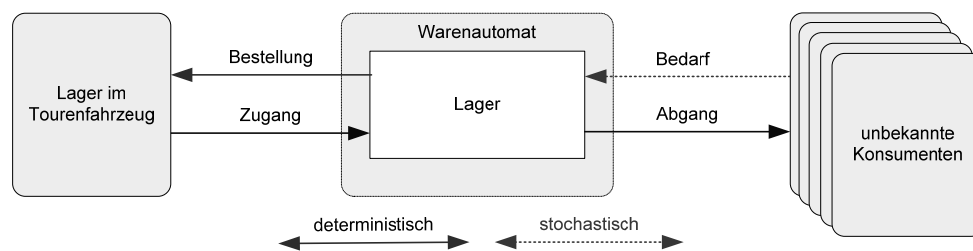


Abbildung 6.4: Konstituenten des Lagerhaltungsproblems an Warenautomaten

Als Lagerhaltungsmotiv dominiert die outputseitige, absatzbedingte Lagerung [Schneeweiß 1981, S. 3-4] zur Befriedigung einer stochastischen Nachfrage vieler unbekannter Konsumenten. Der Bedarf an einzelnen Waren je Absatzstelle ist stochastisch und schwankend, sie ist Einflüssen auf unterschiedlichen Ebenen des Marktes, sowohl der individuellen Absatzstelle als auch der Marken- und Artikelebene, ausgesetzt.² Auf der Ebene der Märkte wird der Bedarf durch grundlegende Konsummuster, Saisonalität, Wetter oder Kalendereffekte wie Feiertage und Urlaubszeiten sowie Marketingaktivitäten beeinflusst.³ Diese können lokal durch eine veränderliche Attraktivität und Frequenz der Absatzstelle sowie lokale Marketingaktivitäten verstärkt werden. Aufgrund der geringen Absatzfrequenz von Automaten im Vergleich zu Abverkäufen in Supermärkten liegen somit nur selten stationäre Bedarfsverläufe vor. Selbst etablierte Produkte zeigen vielfach je Abverkaufsstelle instationäre Bedarfsverläufe mit hohen Zufallsschwankungen, lokalen Trends, veränderbaren Saisonalitäten und Niveaueverschiebungen, was sich in einem hohen Variationskoeffizienten ausdrückt.

¹ Die wirtschaftliche Zahl der Lagerstufen, die Anzahl der Lagerstandorte, die Größe der Tourenfahrzeuge und die Zuordnung von Warenautomaten zu einer Belieferungstour wird in der taktischen Gestaltung des Logistiksystems bestimmt [Müller-Hagedorn 1998, S. 504] und gilt für die Entscheidungssituation der Warendisposition als gegeben.

² Hinsichtlich der Art der Waren kann von Markenprodukten mit geringen, homogenen Verpackungsgrößen und somit einheitlichen logischen Anforderungen ohne Rück- bzw. Abtransport ausgegangen werden.

³ So wird zum Beispiel bei wärmeren Wetter weniger Schokolade, aber mehr Zigaretten konsumiert.

Die Absatzstelle des Warenautomaten beinhaltet zudem ein Sortiment aus mehreren Produkten und Marken¹, die untereinander in komplementärer oder substitutiver Beziehung stehen können.² Die Substitutionseffekte zwischen Markenfamilien, einzelnen Marken oder Artikelgruppen des Sortimentes, z. B. durch Produktneueinführungen oder -auslistungen, beeinflussen ebenfalls den Artikelabsatz. Da Sortimente vielfach je Absatzstelle bestimmt werden, sind sie nicht über Absatzstellen hinweg vergleichbar und die lokalen Effekte nicht im Einzelnen zu bestimmen. Die unterschiedlichen markt-, absatzstellen- und artikelbezogenen Bedarfseinflüsse können je Dispositionsperiode stärkere oder schwächere Auswirkungen zeigen und machen durch ihre Interdependenz eine eindeutige Erhebung und Zuweisung je Absatzstelle unmöglich. Die Bedarfs- und Absatzverläufe an Warenautomaten spiegeln so die Charakteristika typischer Bedarfsverläufe im Handel wider.

Der Lagerabgang zur Bedarfsbefriedigung der Konsumenten ist aufgrund der Automatisierung der Warenausgabe durch den Automaten ebenfalls als deterministisch anzusehen.³ Somit sind outputseitig allein die stochastischen, instationären Bedarfsverläufe durch die zeitreihenspezifische Auswahl geeigneter Prognoseverfahren für Zeitreihen mit Trends, Saisonalität und exogenen Erklärungsvariablen zu berücksichtigen

Auf der Inputseite steuern die Entscheidungsvariablen des Zeitpunktes und der Menge einer Bestellung den Inputstrom an Handelswaren. Technische Ansätze zur laufenden Überwachung des Abverkaufs an Warenautomaten, z. B. durch Funkübertragung der Absatzdaten, sind zwar technisch realisierbar, gelten jedoch aufgrund hoher Investitionen als unwirtschaftlich. Somit scheidet Bestellpunktverfahren, die eine laufende Überwachung des Warenbestandes erfordern, für die Disposition an Warenautomaten durch die räumliche Distanz und mangelnden Informationsaustausch von Warenautomat und Handelszentrale aus. Die hohen fixen Bestellkosten der Belieferung durch einen Tourenfahrer verbieten dabei die Disposition einzelner Waren im Automaten in unterschiedlichen Bestellrhythmen.⁴ Die Verbunddisposition gleichartiger Produkte mit unterschiedlichen Bedarfsverläufen begünstigt ebenfalls die Verwendung rhythmischer Verfahren der Bestellmengenrechnung. Somit ergibt sich ein extern determinierter, konstanter Rhythmus, der die rhythmische Disposition an Warenautomaten durch eine (r, S) -Politik der Bestellrhythmusverfahren mit fixem Bestellzeitpunkt und variabler Bestellmenge zur Auffüllung auf das optimale Bestellniveau erfordert. Dieser Rhythmus wird bei Warenautomaten als Füllrhythmus bezeichnet. SILVER et al. [1998] bezeichnen den Warenautomaten entsprechend als ein typisches Beispiel für ein Bestellrhythmus-Modell ohne Kenntnis des Warenbestands innerhalb der Dispositionsperiode.

Wie auch bei der Belieferung von Einzelhandelsfilialen aus einem Regionallager üblich erfolgt somit an Warenautomaten eine rhythmische Anfahrt im Rahmen fester Touren und Routen. Diese sind vielfach durch die Automatendulder, auf deren Grund und Boden der Warenautomat betrieben wird, hinsichtlich Wochentag und Zeitfenster der Belieferung vorgegeben, zum Beispiel bei Innengeräten durch Öffnungszeiten der Stellplätze, und häufig

¹ Im Kontext der Disposition von Markenwaren an Automaten wird der handelsseitige Begriff „Ware“ mit dem logistikorientierten Begriff „Artikel“ und den marketingorientierten Begriffen „Produkt“, „Sorte“ oder „Marke“ synonym verwendet, auch wenn diese in der Marketingliteratur eine differenziertere Bedeutung besitzen. Sie bezeichnen jeweils eine Handelsware.

² So bestehen z. B. hohe komplementäre Effekte zwischen Kaffee und Schokoladenriegeln, während substitutive Beziehungen zwischen einzelnen Artikeln innerhalb der Artikelgruppe der Schokoladenartikel oder zwischen salzigen und süßen Snacks bestehen können. An Zigarettenautomaten dominieren aufgrund der hohen Markenbindung substitutive Effekte.

³ Der Lagerabgang kann durch technische Fehlfunktionen im Automaten ebenfalls stochastischer Art sein. Da in den Daten der nachfolgenden Experimente keine Störungen an den Automaten aufgetreten sind, wird hier aus Gründen der Vereinfachung ein deterministischer Warenabgang angenommen.

⁴ Es ist anzunehmen, dass auch bei laufender Überwachung der Warenbestände und Anwendung eines Bestellpunktverfahrens das mehrfache Anfahren eines Warenautomaten für jede den Bestellpunkt erreichende Handelsware innerhalb einer Planungsperiode unwirtschaftlich wäre.

vertraglich fixiert. Zudem erfordern die Risiken einer Beraubung oder von Störungen mit der daraus resultierenden Verkaufsunfähigkeit und einer verringerten Kundenzufriedenheit einen regelmäßigen Besuch jedes Automaten, der daher in der Regel nicht seltener als alle 14 Tage stattfindet. Die konstante, wirtschaftliche Lieferfrequenz einer Absatzstelle wird somit vorab bestimmt. Die vertraglich gebundenen Belieferungsfrequenzen erlauben die Bildung wirtschaftlicher Touren- und Routenpläne mit festen Belieferungstagen und -zeitpunkten, die für die Disposition an Warenautomaten als vorgegebene, exogene Größen betrachtet werden und somit keine Entscheidungsvariablen darstellen. Aufgrund der Komplexität und Instationarität der stochastischen Bedarfsverläufe und der exogen vorgegebenen Bestellrhythmen¹ erfolgt die Ermittlung der optimalen Bestellmenge immer nur für die nachfolgende Periode, mit $r = 1$, und somit im Rahmen einer $(1, S)$ - bzw. (S) -Politik. Somit gestalten sich die Dispositionsprozesse an Warenautomaten als natürliche (S) -Politik der rhythmischen Warendisposition unter nur minimalen Vereinfachungen der Konstituenten des Lagerhaltungsproblems.

Das Modell der Disposition an Automaten vereinfacht sich inputseitig weiter durch den Fortfall der Wiederbeschaffungszeit zwischen Bestellung und Wareneingang aufgrund der unmittelbaren Automatenbefüllung aus dem Tourenfahrzeug während der Warendisposition an der Absatzstelle.² Sie kann somit als deterministisch und gleich Null betrachtet werden. Die unmittelbare Belieferung je Produkt garantiert jedoch keine deterministischen Liefermengen, da durch Fehlmengen in der übergeordneten Lagerstufe des Tourenfahrzeugs auch Lieferfehlmengen möglich sind. Trotz stochastischer Absatzzahlen der Warenautomaten erzielen Tourenfahrzeuge durch hohe Warenbestände einen empirischen β -Servicelevel von 99.99%.³ Daher kann die Disposition an Warenautomaten das Auftreten stochastischer Lieferfehlmengen vernachlässigen, wodurch sich die Modellkomplexität durch die zulässige Annahme deterministischer Liefermengen auf einstufige Lagerhaltungssysteme mit ausschließlich deterministischen Inputströmen reduziert.⁴

Aufgrund des standortspezifischen Kaufverhaltens, der vergleichsweise geringen Absatzdaten und der unbekanntenen Konsumenten in Folge des unpersönlichen Verkaufs am Automaten herrscht weitgehende Unkenntnis über das Kaufverhalten und die Substitutionsbeziehungen zwischen Produkten an Warenautomaten. Die mangelnden Daten machen eine valide und reliable Bestimmung bzw. eine weiterführende Modellierung der Interaktionseffekte zwischen einzelnen Marken im Sortiment eines Warenautomaten unmöglich. Folglich vereinfacht sich die Modellauswahl auf stochastische Einprodukt-Modelle unter Vernachlässigung der Komplexität realer Entscheidungen im Sortiment aus mehreren Produkten.⁵

Zusammenfassend kann die Disposition an Warenautomaten durch einen stochastischen, instationären Bedarf und deterministische Inputströme charakterisiert werden. Die besondere Logistikstruktur erlaubt ferner eine Dekomposition mehrperiodiger Modelle in einperiodige Bestellrhythmus-Modelle ohne Wiederbeschaffungszeit mit deterministischen Liefermengen, unter Vernachlässigung mehrstufiger Dispositionssysteme und Interaktionen zwischen

¹ Hier werden Überlegungen zur Mindesthaltbarkeit aufgrund der hohen Umschlagsgeschwindigkeit bei gleichzeitiger langer Haltbarkeit der Produkte vernachlässigt, die durch die Ermittlung optimaler Tourenpläne und Belieferungsfrequenzen berücksichtigt werden.

² Die tatsächlich entstehende, als deterministisch anzunehmende Wiederbeschaffungszeit von wenigen Minuten zur Mengenermittlung, Kommissionierung der Ware im Tourenfahrzeug und Befüllung des Automaten ist geringer als der durchschnittliche Zeitraum zwischen Kaufakten und kann somit vernachlässigt werden.

³ Fehlmengen entstehen vereinzelt bei Neuprodukteinführungen sowie durch vorzeitige Listung und Auslistung einzelner Waren und sind somit zu vernachlässigen.

⁴ Die Ermittlung optimaler Warenbestände und Dispositionsstrategien der übergeordneten Lagerstufe der Tourenfahrzeuge, Regionallager und Zentrallager im Sinne einer mehrstufigen Bestandsplanung ist somit nicht erforderlich und nicht Gegenstand dieser Arbeit.

⁵ Eine Erweiterung der Einprodukt- auf Mehrprodukt-Modelle erscheint jedoch durch zusätzliche Primärerhebungen mit Konsumenten für die weiterführende Forschung interessant.

Artikeln in Mehrprodukt-Modellen. Die Bedarfsrechnung erfordert somit Verfahren der stochastischen Bedarfsprognose der nachfolgenden Periode in Form einer $t+1$ Einschrittprognose bei instationären Bedarfsverläufen unter Berücksichtigung von Trends, Saisonalitäten, Niveauverschiebungen, Strukturbrüchen und Ausreißern. Das Modell der Bestellmengenrechnung beruht daher auf einer $(1, S)$ -Bestellrhythmus-Politik eines einperiodigen Newsboy-Modells auf der Basis instationärer Bedarfsverläufe unter Analyse der Standardabweichung der Prognoseresiduen zur Ermittlung der Sicherheitsbestände. Tabelle 6.2 erlaubt eine grafische Einordnung der Modelldeterminanten durch eine Schattierung der Felder des morphologischen Kastens.

Tabelle 6.2: Determinanten der Disposition an Warenautomaten

Struktur des Bedarfs	Deterministischer Bedarf	Stochastisch-stationärer Bedarf	Stochastisch-instationärer Bedarf
Warenabgang	Deterministischer Warenabgang		Stochastischer Warenabgang
Struktur der Liefermengen	Deterministische Liefermengen	Stochastisch-stationäre Liefermengen	Stochastisch-instationäre Liefermengen
Struktur der Lieferzeiten	Deterministische Lieferzeiten	Stochastisch-stationäre Lieferzeiten	Stochastisch-instationäre Lieferzeiten
Tiefe des Lagersystems	Einstufiges System	Zweistufiges System	Mehrstufige Systeme
Anzahl der Produkte	Einprodukt-Modell		Mehrprodukt-Modelle
Anzahl der Perioden	Einperiodige Modelle		Mehrperiodige Modelle
Konstanz der Parameter	konstante Parameter		variable Parameter
Restriktionen der Ressourcen	keine Ressourcenrestriktionen		Ressourcenrestriktionen

Durch die Interdependenz von stochastischem Bedarf, stochastischer Wiederbeschaffungszeit und stochastischen Liefermengen in der Warendisposition in herkömmlichen Filialen des Einzelhandels kann den einzelnen Komponenten nur schwer eine besondere Bedeutung zugewiesen werden. Im Fall der Automatenbefüllung bestimmt hingegen die Genauigkeit der Bedarfsprognose durch den Fortfall anderer stochastischer Einflussgrößen maßgeblich die Güte der Dispositionsentscheidung, sowohl anhand der Prognose des erwarteten Bedarfsniveaus als auch über die Höhe der Sicherheitsbestände aus der Streuung vorangegangener Prognosefehler. Somit erhält die valide und reliable Berechnung der Absatzprognose an Warenautomaten einen bedeutenden Einfluss auf die gesamte Güte der Warendisposition. Dies legitimiert die nachfolgende Untersuchung einer Vielzahl konkurrierender Modelle und Verfahren zur stochastischen Bedarfsrechnung und Warendisposition.

6.2.2 Datenanalyse

6.2.2.1 Stichprobe der Zeitreihen

Alle Prognoseverfahren werden auf einer zufälligen, geschichteten Stichprobe von empirischen Absatzzeitreihen überprüft. Die Zeitreihen bezeichnen den Absatz von Verpackungseinheiten (engl. *stock keeping units*, SKUs) von Markenzigaretten in Form von Zigarettenschachteln, die eine unterschiedliche Anzahl von Zigaretten enthalten können.¹ Das Prognoseobjekt ist somit die Vorhersage von Verpackungseinheiten je Abverkaufsstelle wie

¹ Der Absatz mehrerer Wahlmöglichkeiten einer Marke an einer Absatzstelle wird in der Analyse zu einer Absatzzeitreihe aggregiert, sofern die Automaten im Verbund nebeneinander betrieben werden.

in anderen Formen des stationären Facheinzelhandels.¹ Alle Daten wurden im Rahmen eines kooperativen Forschungsprojektes der Universität Hamburg zur Beurteilung von Verfahren der Absatzprognose und Warendisposition mit einem führenden deutschen Automatenaufsteller erhoben. Die Abverkaufswerte wurden im Warenautomaten elektronisch registriert und während des Füllvorganges durch ein MDE-Gerät per Datenübertragung direkt aus dem Automaten ausgelesen.² Die Datenqualität gilt daher als valide, reliabel und robust gegen Fehler der Datenerhebung und Datenverarbeitung.

Eine repräsentative Stichprobe aller Standorttypen und Absatzverläufe erscheint nicht realisierbar, da Warenautomaten der ubiquitären Nahversorgung dienen und so nahezu überall auftreten. Die Heterogenität der Absatzstellen erfordert für ein Unternehmen daher üblicherweise eine Vollerhebung zur Sicherstellung der Repräsentativität.³ Für den empirischen Nachweis der Gültigkeit der methodischen Innovation von NN und asymmetrischen Kostenfunktionen an Warenautomaten genügt jedoch eine bedingt repräsentative Stichprobe.⁴ Das Auswahlverfahren der Stichprobenziehung erfolgte daher als bewusste Auswahl von 8 typischen Standorten mit 153 Zeitreihen durch einen internen Experten des Unternehmens. Die Schichtung der Stichprobe beschränkte sich auf die Hauptgliederungsmerkmale von Standorten.⁵ Zunächst wurden zwei geografische Nielsen-Gebiete⁶ I und VII ausgewählt, um Verzerrungen durch den Einfluss gebietsspezifischer Saisonalitäten zu vermeiden. Anschließend wurden mehrere Standorte für Aufstellplätze im Innen- oder Außenbereich ausgewählt. Stellplätze im Innenbereich, sog. Innenautomaten, finden sich überwiegend in der Gastronomie, also in Diskotheken, Restaurants, Kneipen, Trinkhallen usw. und dienen der Bedarfsbefriedigung innerhalb der Lokalität. Stellplätze im Außenbereich, sog. Außenautomaten, finden sich an öffentlich zugänglichen Aufstellplätzen in Wohngebieten oder hoch

¹ Dabei entstehen Verzerrungen, da die Simulation auf der Basis vergangener Absatzdaten anstelle von Bedarfsdaten erfolgt. Die Absatzhistorie der Vergangenheit zeigt lediglich jene Verteilung, die durch die damaligen Warenbestände auch realisiert werden konnte. Haben Absatzspitzen in der Vergangenheit zu Fehlmengen geführt, so wurde nicht das volle Absatzpotenzial in einer Periode realisiert, und die Zeitreihe weist für diese Perioden einen zu niedrigen Bedarf aus. Dies kann insbesondere in einem nachträglichen Verfahrensvergleich zu Verzerrungen der Ergebnisse führen, da ein Prognoseverfahren, das in der realen Anwendung aufgrund vorangegangener Absatzzahlen für eine bestimmte Periode eine zu niedrige Bedarfsmenge vorhergesagt und diesen Absatz durch eine Fehlmengensituation somit herbeigeführt hat, dieselben Werte in der fairen Ex-ante-Simulation vorhersagen wird, in der diese exakt richtig sind. Somit werden die Ergebnisse im Sinne einer *self-fulfilling prophecy* zugunsten des Ursprungsverfahrens verzerrt. Dieses Problem liegt jedoch nahezu allen Verfahrensvergleichen auf empirischen Daten zugrunde.

² Aufbauend auf diesen Daten wurde im MDE-Gerät je Zigarettenmarke der erwartete Absatz prognostiziert, die entsprechenden Sicherheitsbestände kalkuliert und ein Füllvorschlag berechnet. Diese Informationen stehen jedoch aus Gründen der Geheimhaltung nicht im Vergleich zu den hier berechneten Prognoseverfahren zur Verfügung, so dass keine Bewertung im Vergleich zur tatsächlichen Prognose- und Dispositionsgüte erfolgen kann. Die Erkenntnisse beruhen somit auf einem rein theoretischen Verfahrensvergleich.

³ Die Problemstellung einer repräsentativen Stichprobengröße für Verfahrensvergleiche der Zeitreihenprognose ist weitgehend ungeklärt. Die Kategorisierung von heterogenen Absatzstellen und Zeitreihen in Prognosevergleichen stellt ein eigenständiges Entscheidungsproblem dar. So wären zur Erhebung von Absatzstellen mit Saisonalität ebenso klassische Urlaubsgebiete, aber auch Fabrikanlagen mit Produktionsferien oder Universitäten mit Semesterferien zu berücksichtigen. Dagegen können Warenautomaten, die nicht in der Produktion, sondern im Bürotrakt einer Fabrik oder im Außenbereich einer Universität aufgestellt sind, ein nicht-saisonales Zeitreihenmuster aufweisen. Dazu kommen Aufstellplätze mit regionaler Saisonalität durch externe Events, z. B. das lokale Oktoberfest in München oder der regional übergreifende Karneval im Rheinland. Eine repräsentative, geschichtete Stichprobe für ein Unternehmen würde mehrere tausend Automaten und die Prognose von Hunderttausenden von Absatzzeitreihen erfordern, was weit über die Grenzen bisheriger Prognosevergleiche hinausginge.

⁴ Diese lässt explizit keine generalisierbaren Aussagen über das untersuchte Unternehmen oder den Gesamtmarkt der Warenautomaten zu. Hierzu sind weitere Untersuchungen erforderlich.

⁵ Die empirische Struktur der Absatzstellen des Unternehmens in Regionen, Niederlassungen, Touren, Aufstellplätzen sowie Informationen zur durchschnittlichen Automatendichte, zur Anzahl der Automaten je Standort, zur Tiefe und Breite des Sortiments etc. kann aus Gründen der Geheimhaltung nicht offengelegt werden.

⁶ Die Nielsen-Gebiete wurden durch die Firma ACNielsen als Hilfsmittel für die Planung von Marktforschung und Werbung verwendet, wobei das Verbraucherverhalten innerhalb eines Nielsen-Gebietes als homogen gilt.

frequenzierten öffentlichen Standorten und dienen der ubiquitären Nahversorgung rund um die Uhr. Außenautomaten zeichnen sich somit durch unterschiedliche Umsatzerwartungen und Konsummuster aus. Zur Sicherstellung einer validen und reliablen Prognose- und Dispositionsberechnung kamen nur Aufstellplätze in Betracht, die eine Absatzhistorie von mindestens zwei Jahren und einen homogenen Lieferrhythmus mit äquidistanten Dispositions- und Füllintervallen aufwiesen.¹ Es verbleiben drei Innenautomaten (IN-1, IN-2 und IN-3) und zwei Außenautomaten (AU-1, AU-2) mit insgesamt 90 Zeitreihen.

Anschließend werden alle zulässigen Zeitreihen der ausgewählten Standorte zum Verfahrensvergleich ausgewählt. Durch die Einlistung neuer Waren und die regelmäßige Sortimentsrotation im Sinne des Category Managements werden bei begrenzter Lagerkapazität und der geringen Anzahl von Wahlmöglichkeiten bei einem Automaten regelmäßig bestehende Waren ausgelistet, was zu verkürzten Zeitreihen führt. Die Prognose von Produkt-Neueinführungen und auslaufenden Artikeln ist jedoch nicht Bestandteil dieser Untersuchung, so dass alle Zeitreihen ohne Abverkaufshistorie von weniger als zwei Jahren von der Untersuchung ausgeschlossen werden.² In Tabelle 6.3 wird eine Übersicht der Standorte (POS), der zulässigen Zeitreihen je Standort und der Länge der Zeitreihen in Beobachtungswerten gegeben.

Tabelle 6.3: Stichprobe der Absatzstellen und Zeitreihen

POS	Nielsen-Gebiet	Standort	Anzahl Waren am POS	Anzahl zulässiger Waren	Länge der Zeitreihen
AU-1	1	Außen	12	12	126
AU-2	1	Außen	13	13	126
IN-1	7	Innen	25	10	126
IN-2	7	Innen	14	9	126
IN-3	7	Innen	26	7	126
Summe			90	51	
Mittelwert			18,0	10,2	

Die Absatzstellen enthalten 51 zulässige Absatzzeitreihen für den Verfahrensvergleich, jeweils mit einer unterschiedlichen Anzahl von Zeitreihen. Die Stichprobengröße gilt im Vergleich zu etablierten Publikationen von Prognosevergleichen mit NN als hinreichend [Hill et al. 1996; Liao und Fildes 2005], da in empirischen Vergleichsstudien von NN bisher nur in Einzelfällen mehr als 20 Zeitreihen untersucht wurden (vgl. Abschnitt 4.4.1.3).³ Diese werden nachfolgend einer explorativen Datenanalyse je Abverkaufsstelle und Zeitreihe unterzogen.

Durch eine derartige Ex-ante-Analyse besteht grundsätzlich die Gefahr, durch Kenntnis der Lage- und Verteilungsparameter, der Saisonalitäten und der Strukturen der analysierten Zeitreihen implizit oder explizit Entscheidungen bezüglich des Experimentdesigns, der Modellauswahl oder -parametrisierung, der Fehlermaße oder der Zielkriterien zu treffen. Dies kann einzelne Verfahren begünstigen und die Validität und Reliabilität der Ergebnisse in Frage stellen. Die Voruntersuchung der Daten entspricht jedoch einem essentiellen Schritt des Prognoseprozesses, und die erhobenen Informationen werden allen Prognoseverfahren gleichermaßen zugänglich gemacht. Zudem ist eine Analyse und Diskussion der Zeitreihen-

¹ Vereinzelt findet bei existierenden Absatzstellen eine Veränderung der Lieferrhythmen, der Liefertage oder der Lieferzeiten statt. Dies kann durch Wünsche des Automatendülders, als Reaktion auf Umsatzschwankungen, durch die Aufstellung eines Automaten mit anderer Sortimentstiefe und anderem Fassungsvermögen oder durch firmenpolitische Entscheidungen zur Lieferpolitik bedingt sein. Die entstehenden Dateninkonsistenzen könnten durch zusätzliche Entscheidungen zur Interpolation oder Summierung der Absatzzahlen Verzerrungen der Ergebnisse bewirken. Um diese für das vorliegende Experimentdesign zu vermeiden, wurden nur Absatzstellen in Betracht gezogen, die eine homogene äquidistante, wöchentliche Lieferfrequenz aufwiesen.

² Auf eine Betrachtung der Substitutionseffekte innerhalb der Markengruppen oder Konsumsegmente je POS wird in Anbetracht geringer empirischer Erkenntnisse nicht eingegangen. Das kontrollierte Experimentdesign des Absatzes an Warenautomaten bietet sich jedoch besonders für derartige Untersuchungen an.

³ Siehe hierzu die Literaturanalyse in Abschnitt 4.4.1.3.

verläufe unabdingbar, um das Experimentdesign zu dokumentieren und Entscheidungen über die Modellierung zu treffen, z. B. die Notwendigkeit hierarchischer Dekomposition aus aggregierten Zeitreihen oder aber die Berücksichtigung exogener Einflüsse durch kausale Prognosemodelle. Somit wird dem Ansatz nach HILL et al [1996] gefolgt und eine disjunkte Stichprobe einer Absatzstelle zur Vorauswahl der NN verwendet, die in Abschnitt 6.2.3.3 dargestellt wird. Zur Erhebung und Eingrenzung des Versuchsaufbaus und der Auswahl und Gestaltung der Prognoseverfahren werden jedoch die vollständigen Zeitreihen untersucht.

6.2.2.2 Explorative Datenanalyse

6.2.2.2.1 Datenanalyse je Absatzstelle

Die Struktur der Zeitreihen wird zunächst durch eine explorative Datenanalyse je Absatzstelle untersucht, um Erkenntnisse für den Untersuchungsaufbau und die nachfolgende Gestaltung der Prognoseverfahren abzuleiten. Abbildung 4.1 zeigt den aggregierten Absatzverlauf einer Absatzstelle in Verpackungseinheiten je Kalenderwoche (KW).¹

Eine visuelle Analyse der Zeitreihen zeigt neben unterschiedlichen Zeitreihenmustern zunächst ein signifikant abweichendes mittleres Absatzniveau zwischen Innen- und Außenautomaten. Außenautomaten weisen einen geringeren durchschnittlichen Absatz als Innenautomaten und eine geringere Streuung des Absatzes auf.

Die Zeitreihen können genauer durch ihre Verteilungsform anhand der Lageparameter von Mittelwert (Mw) und Median (Md), der Streuungsparameter der Standardabweichung (StdAb) und der Spannweite aus Minimum (Min) und Maximum (Max) sowie der Wölbung anhand der Kurtosis (Ku) und der Schiefe der Verteilung (Skew) charakterisiert werden. Darüber hinaus wird der Verlauf der Zeitreihen durch deskriptive Maße charakterisiert. Der Variationskoeffizient (VaKo) beschreibt das Verhältnis von Standardabweichung zu Mittelwert, welches bei stationären Zeitreihen als Indikator für das Ausmaß der Zufallsschwankungen dient. Die Anzahl von Perioden ohne Absatz, der so genannten Nullperioden (NuP), wird sowohl als Anzahl sowie in Prozent aller Beobachtungsperioden ausgedrückt. Eine Übersicht der Charakteristika der Zeitreihen ist in Tabelle 6.4 dargestellt.

Tabelle 6.4: Parameter der aggregierten Zeitreihen je Absatzstelle nach Standort

	Parameter der Zeitreihe						Verlauf der Zeitreihe			
	Min	Max	Mw	Md	StdAb	Skew	Kurt	VaKo	NuP	NuP %
AU-1 Gesamt	0,0	122,0	48,7	47,5	16,4	0,9	4,6	33,6%	2,0	1,6%
AU-2 Gesamt	0,0	100,0	57,7	58,5	17,2	-0,6	2,0	29,9%	3,0	2,4%
Außenautomaten	0,0	111,0	53,2	53,0	16,8	0,1	3,3	31,7%	2,5	2,0%
IN-1 Gesamt	0,0	501,0	171,6	169,5	96,9	0,5	1,6	56,5%	16,0	12,7%
IN-2 Gesamt	0,0	321,0	189,5	191,5	52,5	-1,1	4,2	27,7%	4,0	3,2%
IN-3 Gesamt	0,0	340,0	134,5	136,5	51,8	0,0	2,4	38,5%	6,0	4,8%
Innenautomaten	0,0	387,3	165,2	165,8	67,1	-0,2	2,7	40,9%	8,7	6,9%
Automaten Gesamt	0,0	276,8	120,4	120,7	46,9	-0,1	2,9	37,2%	6,2	4,9%

Min=Minimum; Max=Maximum; Mw=Mittelwert; Md=Median; StdAb=Standardabweichung; Skew=Schiefe; Kurt=Kurtosis; VaKo=Variationskoeffizient; NuP= Anzahl von Nullperioden; NuP %=Anzahl Nullperioden in % von allen Beobachtungswerten

Die Tabelle bestätigt die signifikanten Unterschiede im Absatzverhalten zwischen Außen- und Innenautomaten der Stichprobe. Durchschnittlich werden pro Woche an Außenautomaten 53,2 Verpackungseinheiten abgesetzt, während die Innenautomaten 165,8 Packungen absetzen. Die Streuung der Zeitreihen steigt korrespondierend mit dem Mittelwert an. Dies wird auch durch den nahezu konstanten Variationskoeffizienten bestätigt, trotz der sonst signifikanten Unterschiede der Lage- und Streuungsparameter der Zeitreihen. Der VaKo dient

¹ Der Absatz einer Absatzstelle wird durch Aggregation des Absatzes aller Zeitreihen der Waren gebildet.

nachfolgend als Indikator für die Höhe der Prognosefehler.¹ Neben geringeren Schwankungen des mittleren Absatzes weist der Absatz überwiegend rechtsschiefe, platykurtische, nicht-normalverteilte Absatzverteilungen auf, wie bei niedrigen Absatzmengen zu erwarten ist.²

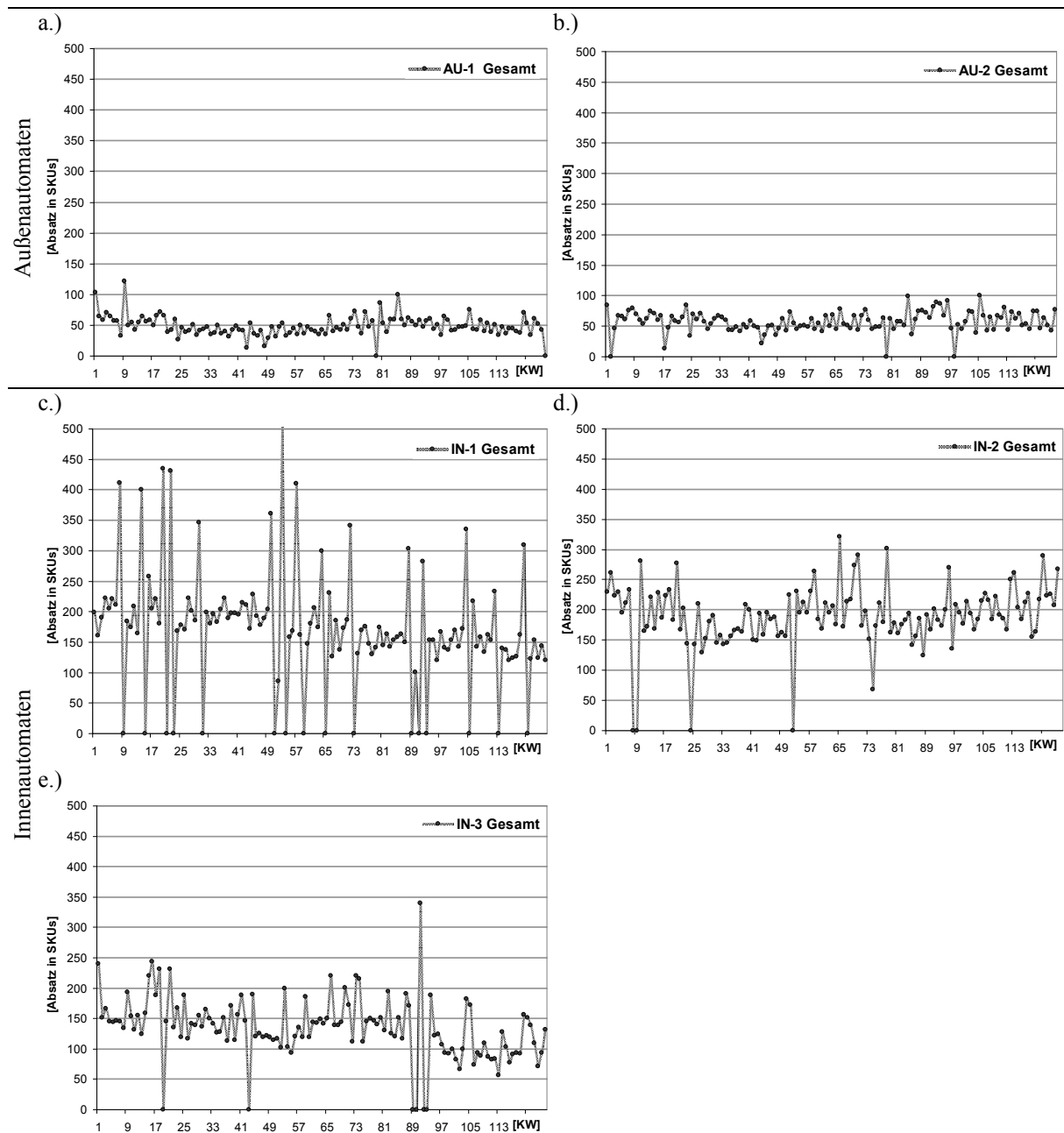


Abbildung 6.5: Zeitreihen des aggregierten Absatzes aller Zeitreihen der Absatzstellen AU-1 (a.), AU-2 (b.), IN-1 (c.), IN-2 (d.) und IN-3 (e.)

Die Absatzwerte von AU-2 und IN-2 weisen jedoch auch linksschiefe Verteilungen auf, die u. a. durch die vielen Nullperioden ohne Absatz hervorgerufen werden, was eine signifikante Verzerrung der Absatzverteilungen bewirkt. Die Zeitreihen in Abbildung 4.1 weisen insbe-

¹ Da der Variationskoeffizient für stationäre Zeitreihen das Ausmaß der Zufallsschwankungen als Verhältnis des Signals zum Rauschen (engl. *signal to noise ratio*) bezeichnet, kann er als Indikator und obere Grenze der Prognosefehler durch die Anwendung Naiver Verfahren verwendet werden. Er erlaubt einen einfachen Vergleich zwischen Absatzstellen und nachfolgend mit den Zeitreihen einzelner Waren je Absatzstelle.

² Die Schiefe der Verteilung wird durch die Skewness angegeben, die Wölbung der Verteilung durch die Kurtosis (vgl. Abschnitt 3.2.2.1.3).

sondere für Innenautomaten sichtbare Ausreißer und vielfach fehlende Werte durch einzelne Perioden ohne Abverkäufe über alle Waren auf, was auch durch den zum Teil erheblichen Anteil an Nullperioden von bis zu 12,7% in Tabelle 6.4 bestätigt wird. Diese sind durch exogene Einflüsse auf den Absatz zu erklären, z. B. durch Baustellen bei Außengeräten oder geschäftsbedingte Schließungen der Gastronomie bei Innengeräten, die einen Zugang der Konsumenten zum Automaten oder auch nur des Tourenfahrers zur Befüllung verhindern. Kann ein Warenautomat nicht im determinierten Lieferintervall angefahren werden, erlaubt die fixe Touren- und Routenplanung normalerweise keine erneute Anfahrt an einem alternativen Liefertag bis zur erneuten, planmäßigen Belieferung in der nachfolgenden Woche. Aufgrund der räumlichen Trennung und mangels informationstechnischer Anbindung der Automaten bleibt der tatsächliche Absatz der Periode unbekannt und muss der Folgeperiode zugerechnet werden. Zudem können fehlende Werte sowie negative Ausreißer durch technische Störungen am Automaten verursacht werden, die einen Abverkauf in der ganzen Dispositionsperiode oder einer Teilperiode unmöglich gemacht haben. In Abhängigkeit von der exogenen Störung können zeitlich verzögert unterschiedliche Effekte auftreten: (a.) ein erhöhter Absatz in der Folgeperiode als dynamischer Effekt der kumulierten Abverkäufe mehrerer Lieferperioden, wenn nur der Zugang zur Warendisposition in der Vorwoche versperrt war, oder (b.) ein normaler bzw. verringerter Absatz in der Folgeperiode, wenn der Abverkauf nach einer Störung der Automatentechnik für die gesamte Folgeperiode wieder möglich ist oder noch für einen Teil der Verkaufsperiode bis zur Reparatur behindert wird. Es handelt sich somit nicht ursächlich um eine Abverkaufshistorie mit sporadischem Bedarf, sondern um externe Einflüsse. Das Eintreten solcher exogener Störungen ist für die Disposition in der Folgeperiode bekannt und kann als Dummy-Variable kausalanalytisch zur Prognose in multivariaten Prognoseverfahren berücksichtigt werden.

Darüber hinaus weisen die Zeitreihen in Abbildung 4.1 leichte Instationaritäten durch globale und lokale Auf- und Abwärtstrends in unterschiedlichen Absatzstellen auf, zudem leichte zyklische Schwankungen in AU-1, AU-2 und IN-3, die auf eine Saisonalität hindeuten könnten. Um diese festzustellen, ist jedoch eine Autokorrelationsanalyse durch die Betrachtung der Autokorrelationsfunktion (AKF) und der partiellen Autokorrelationsfunktion (PAKF) nach der BOX-JENKINS-Methodologie erforderlich. Abbildung 4.1 zeigt AKF und PAKF der Absatzzeitreihe des Außenautomaten AU-1 sowie ihrer ersten Differenzen $I(1)$.¹

AKF (a.) und PAKF (b.) der Originalzeitreihe zeigen eine schwache Instationarität, welche durch einen Augmented-Dickey-Fuller-(ADF-)Test nicht bestätigt wird.² Das sinusoidale Muster der AKF belegt einen saisonalen $AR(p)$ -Prozess niedriger Ordnung, der durch die signifikanten partiellen Autokorrelationskoeffizienten in $t-4$ und $t-5$ bestätigt wird und auf einen $ARIMA(0,0,0)(1,0,0)_4$, $ARIMA(0,0,0)(2,0,0)_4$ oder $ARIMA(0,0,0)(0,0,1)_4$ -Prozess mit monatlicher Saisonalität hinweist. Die Analyse zeigt somit keine Trends oder saisonalen Prozesse höherer Ordnung durch jährliche und quartalsweise saisonale Muster³, so dass höherwertige saisonale $AR(p)$ - oder $MA(q)$ -Prozesse ausgeschlossen werden können. Der Absatz an AU-1 ist folglich schwach stationär mit monatlicher Saisonalität.

¹ ACF und PACF zeigen einen signifikanten Korrelationskoeffizienten bei einem Signifikanzniveau von 0.05 durch ein positives oder negatives Überschreiten des blau schraffierten Bereiches an, zum Beispiel für PACF $t-3$, einen hochsignifikanten Koeffizienten bei einem Signifikanzniveau von 0.01 durch ein positives oder negatives Überschreiten des grün schraffierten Bereiches, zum Beispiel für PACF $t-1$ und $t-2$.

² Somit sind ACF (c.) und PACF (d.) der ersten Differenzen hier nicht notwendig, sie bestätigen aber die ACF- und PACF-Analyse der Originalzeitreihe.

³ Diese müssten bei wöchentlichen Zeitreihen signifikante ACF- oder PACF-Koeffizienten der Lags $t-12$ für eine quartalsweise bzw. $t-52$ für eine wöchentliche, jährlich wiederkehrende Saisonalität hervorrufen.

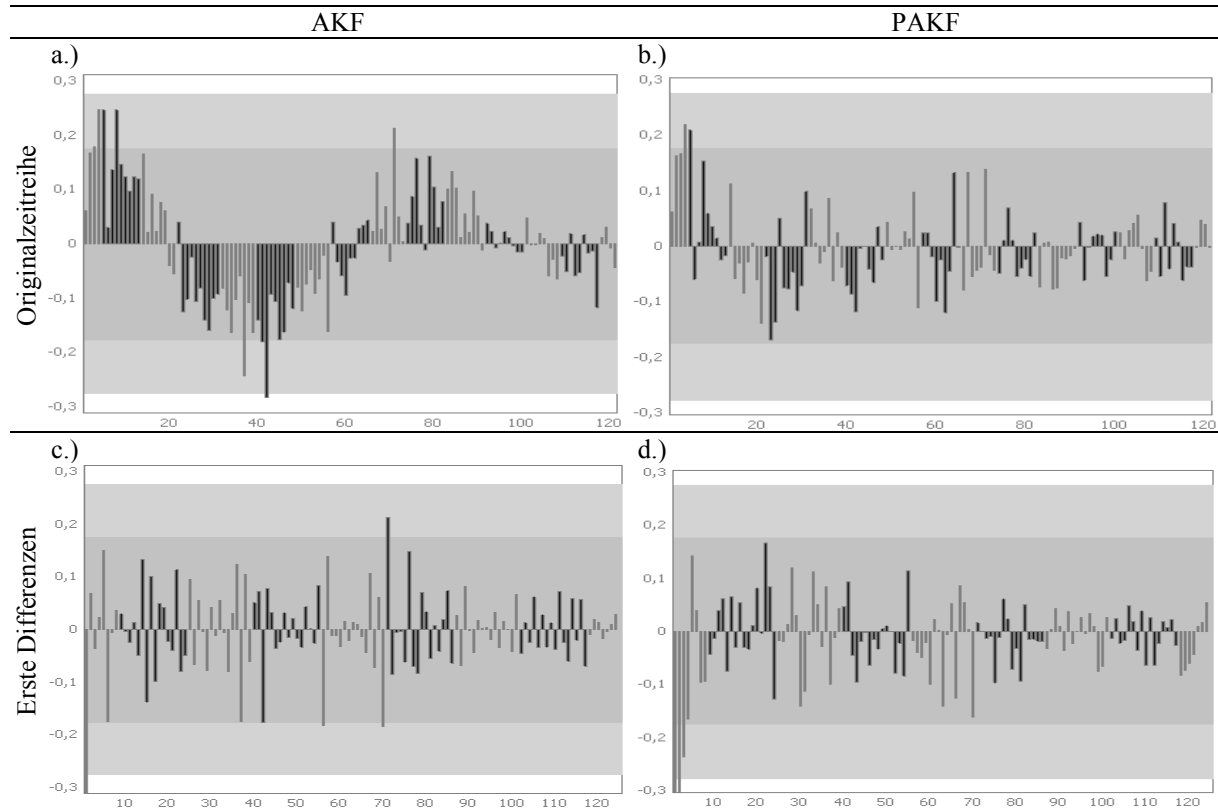


Abbildung 6.6: AKF (a.) und PAKF (b.) der Originalzeitreihe sowie AKF (c.) und PAKF (d.) der ersten Differenzen des Absatzes an Abverkaufsstelle AU-1

Die Zeitreihe des Außenautomaten AU-2 zeigt einen ähnlichen stationären Verlauf, ohne Anzeichen für Saisonalität oder Trend (siehe Abbildung 6.7).

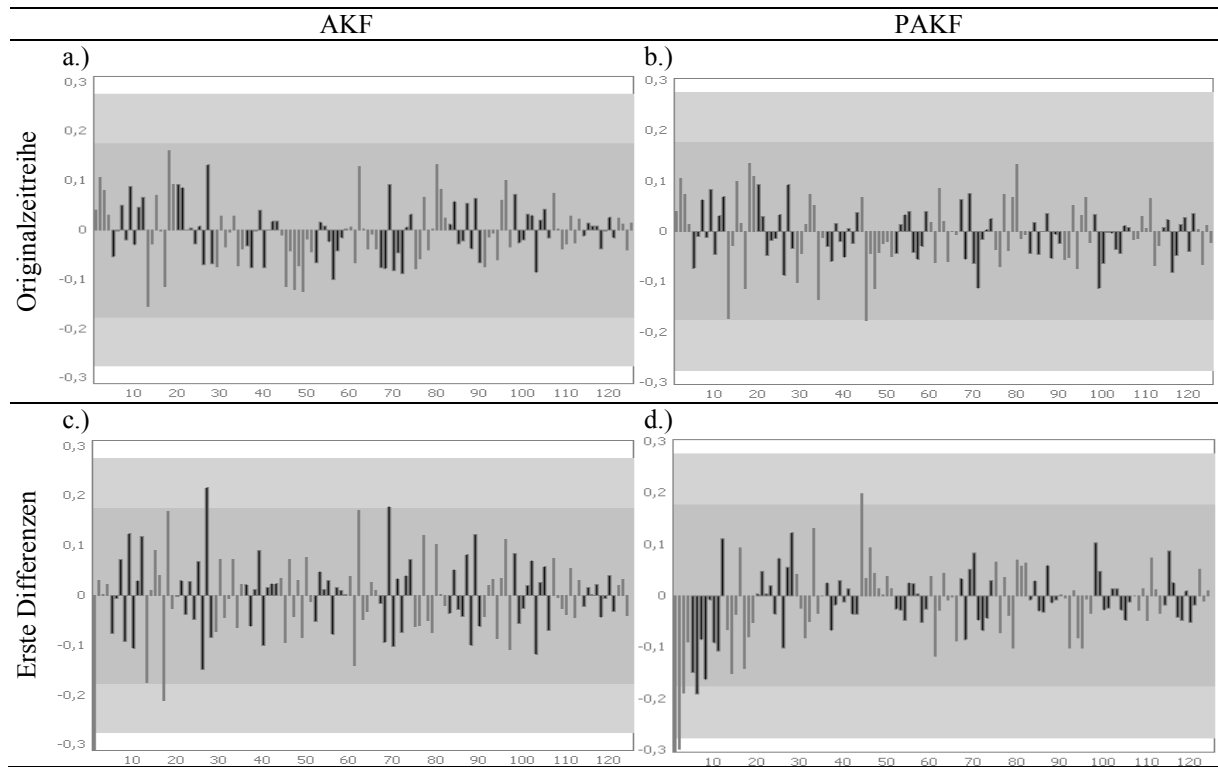


Abbildung 6.7: AKF (a.) und PAKF (b.) der Originalzeitreihe sowie AKF (c.) und PAKF (d.) der ersten Differenzen des Absatzes an Abverkaufsstelle AU-2

AKF und PAKF belegen eine stationäre Originalzeitreihe, was durch den ADF-Test bestätigt wird und weshalb die Analyse der differenzierten Zeitreihe entfällt. Der Absatz zeigt keinerlei

signifikante Autokorrelationen und weist auf einen nicht-saisonalen ARIMA (0,0,0)-Prozess ohne Trend oder höherwertige Saisonalität hin. Auch hier finden sich einzelne Ausreißer durch exogene Einflüsse auf den Absatz der Absatzstelle AU-2, die eine Analyse der Autokorrelationen verzerren könnten.

Abbildung 6.8 zeigt AKF und PAKF der Absatzzeitreihe von Innenautomaten IN-1 sowie der ersten Differenzen des Absatzes.

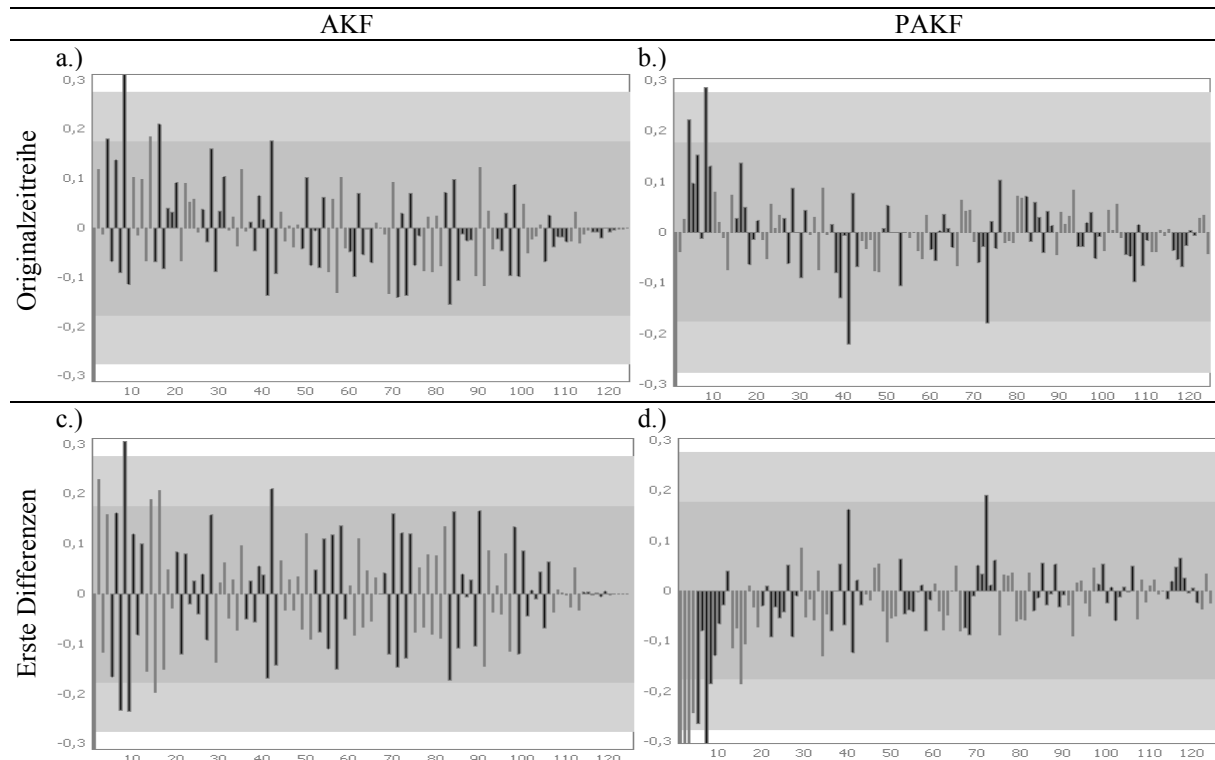


Abbildung 6.8: AKF (a.) und PAKF (b.) der Originalzeitreihe sowie AKF (c.) und PAKF (d.) der ersten Differenzen des Absatzes an Abverkaufsstelle IN-1

AKF (a.) und PAKF (b.) der Originalzeitreihe zeigen einen mittelwertstationären Verlauf, was durch einen ADF-Test bestätigt wird. Die starke Überlagerung der Zeitreihe durch Ausreißer und fehlende Werte verzerrt die Erkenntnisse der Autokorrelationsanalyse. Dennoch wird ein saisonaler $AR(p)$ -Prozess bzw. ein saisonaler $MA(q)$ -Prozess niedriger Ordnung deutlich, mit hochsignifikanten Lags der partiellen Autokorrelationen in $t-1$, $t-4$, und $t-8$. Dies weist auf einen saisonalen $ARIMA(0,0,0)(0,1,1)_4$ - oder $ARIMA(0,0,0)(2,1,0)_4$ -Prozess hin, erneut mit monatlicher und ohne jährliche Saisonalität.

AKF und PAKF in Abbildung 6.9 sowie der ADF-Test des Innenautomaten IN-2 zeigen dagegen eine stationäre Struktur ohne saisonale autoregressive Prozesse, jedoch ebenfalls überlagert von vereinzelt Ausreißern. Der Mangel an signifikanten Autokorrelationen legt einen $ARIMA(0,0,0)$ -Prozess nahe.

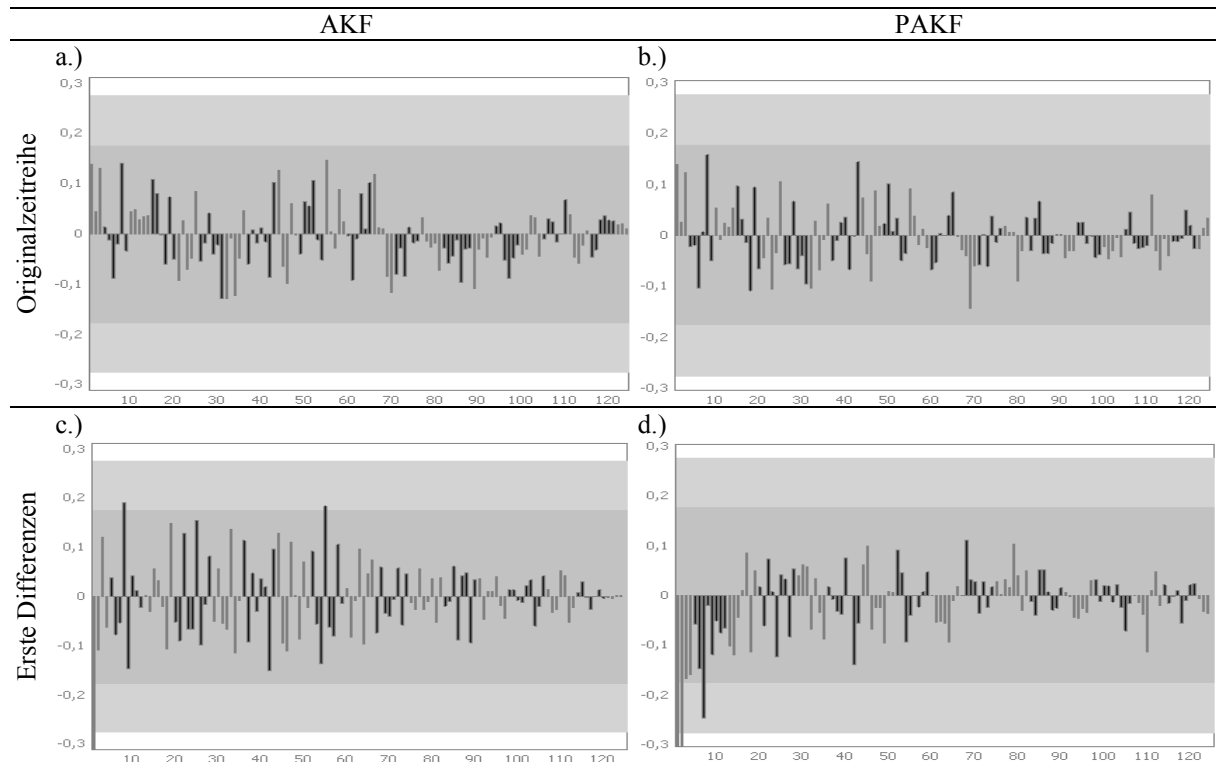


Abbildung 6.9: AKF (a.) und PAKF (b.) der Originalzeitreihe sowie AKF (c.) und PAKF (d.) der ersten Differenzen des Absatzes an Abverkaufsstelle IN-2

AKF und PAKF des Innenautomaten IN-3 in Abbildung 6.10 zeigen wiederum einen schwach stationären Verlauf mit signifikanten saisonalen Strukturen.

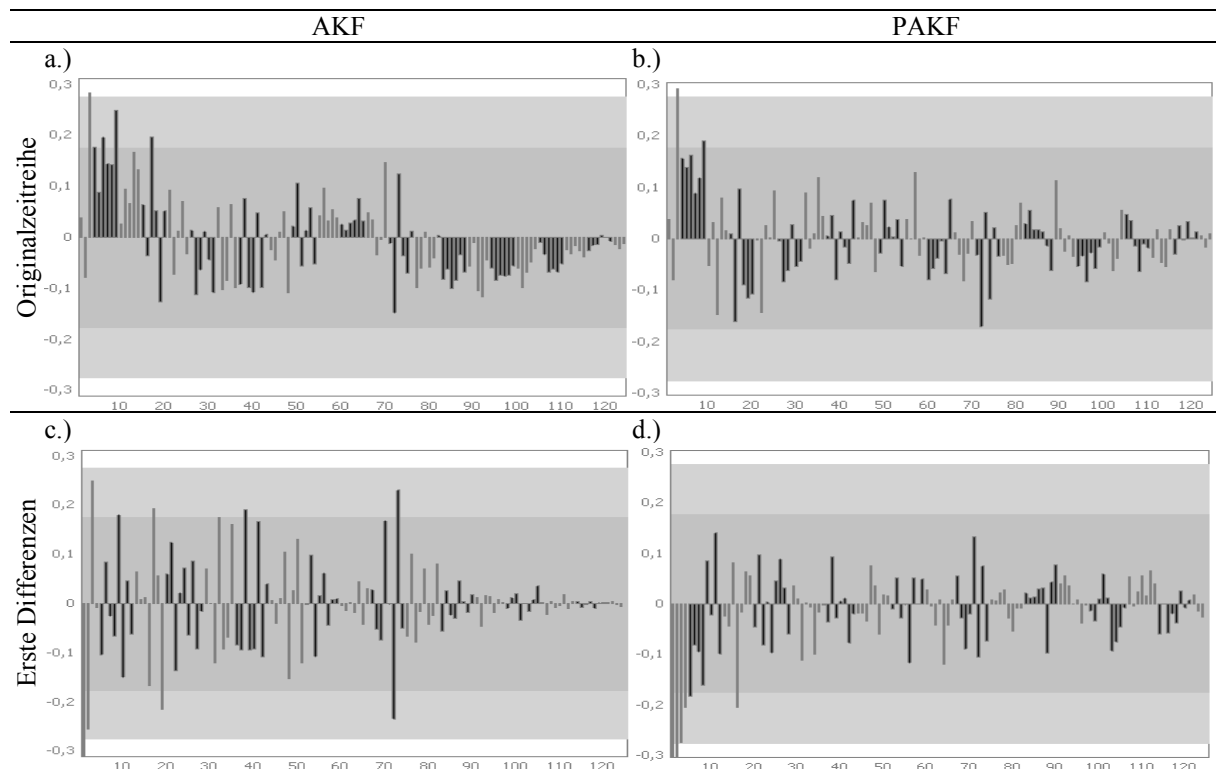


Abbildung 6.10: AKF (a.) und PAKF (b.) der Originalzeitreihe sowie AKF (c.) und PAKF (d.) der ersten Differenzen des Absatzes an Abverkaufsstelle IN-2

Während die Zeitreihe in Abbildung 4.1 einen schwach instationären Verlauf nahelegt, kann dieser nicht durch einen ADF-Test bestätigt werden. AKF (a.) und PAKF (b.) der Originalzeitreihe zeigen hochsignifikante Lags der PAKF in $t-3$ und signifikante Lags der AKF in $t-3$,

$t-6$ und $t-9$, was einen saisonalen ARIMA $(0,0,0)(1,0,0)_3$ -Prozess identifiziert, mit einer dreiwöchentlichen Saisonalität, ohne monatliche oder jährliche Saisonalität.

Zusammenfassend bestätigt sich die Heterogenität der Absatzverläufe von Warenautomaten aufgrund der individuellen Bedarfsmuster der Aufstellplätze. Sowohl die Lage- und Streuungsparameter der Absatzstellen als auch die Zeitreihenkomponenten von Trends und Saisonalität treten weder konsistent über alle Absatzstellen noch konsistent für Innen- oder Außenautomaten auf. Vielmehr zeigen sich stationäre, schwach instationäre und instationäre Zeitreihen mit oder ohne Saisonalität unterschiedlicher Länge.¹ Die geringe Homogenität der Zeitreihenstrukturen erfordert eine genauere Untersuchung der einzelnen Zeitreihen je Absatzstelle und legt eine absatzstellenspezifische Modellierung aller Prognoseverfahren nahe. Alle empirischen Zeitreihen besitzen zudem Ausreißer und Perioden ohne Absatz. Diese starken exogenen Einflüsse erfordern eine Berücksichtigung im Versuchsaufbau durch die Modellierung dynamischer Interventionsmodelle.

Die Muster der aggregierten Absatzzeitreihen erlauben jedoch noch keine Rückschlüsse auf die Struktur der einzelnen Absatzzeitreihen der zu prognostizierenden Waren und darauf, durch welche geeigneten Verfahren diese im Einzelnen zu prognostizieren sind. Die aufgezeigte Heterogenität verdeutlicht jedoch die Notwendigkeit einer detaillierten Untersuchung und Kategorisierung einzelner Zeitreihen über die Unterscheidung in Innenautomaten und Außenautomaten hinaus, um homogene Gruppen von Zeitreihen zu finden und nachfolgend reliable Erkenntnisse darüber ableiten zu können, unter welchen Bedingungen der Zeitreihenmuster ein Prognoseverfahren eine hohe Genauigkeit erzielt. Die Zeitreihen der Waren je Absatzstelle sind daher nachfolgend einzeln zu untersuchen und in homogene Kategorien zur abschließenden Bewertung der Prognose- und Dispositionsergebnisse aufzuteilen. Sie werden nachfolgend exemplarisch dargestellt, um einzelne Herausforderungen der Datenbasis und des Experimentdesigns zu verdeutlichen.

6.2.2.2.2 Datenanalyse je Zeitreihe

Anschließend werden die 51 Zeitreihen der Handelswaren der fünf Absatzstellen einer genaueren Analyse unterzogen und nach empirischen Merkmalen der Zeitreihenstruktur in homogene Gruppen kategorisiert. Einzelne Absatzzeitreihen der Absatzstelle AU-1 sind in Abbildung 6.11 stichprobenartig dargestellt.

¹ Die Erkenntnisse über die Saisonalität werden durch eine Analyse der saisonalen Diagramme für monatliche, quartalsweise und jährliche Saisonalität für jede der Zeitreihen bestätigt, die hier aus Platzgründen nicht abgebildet sind.

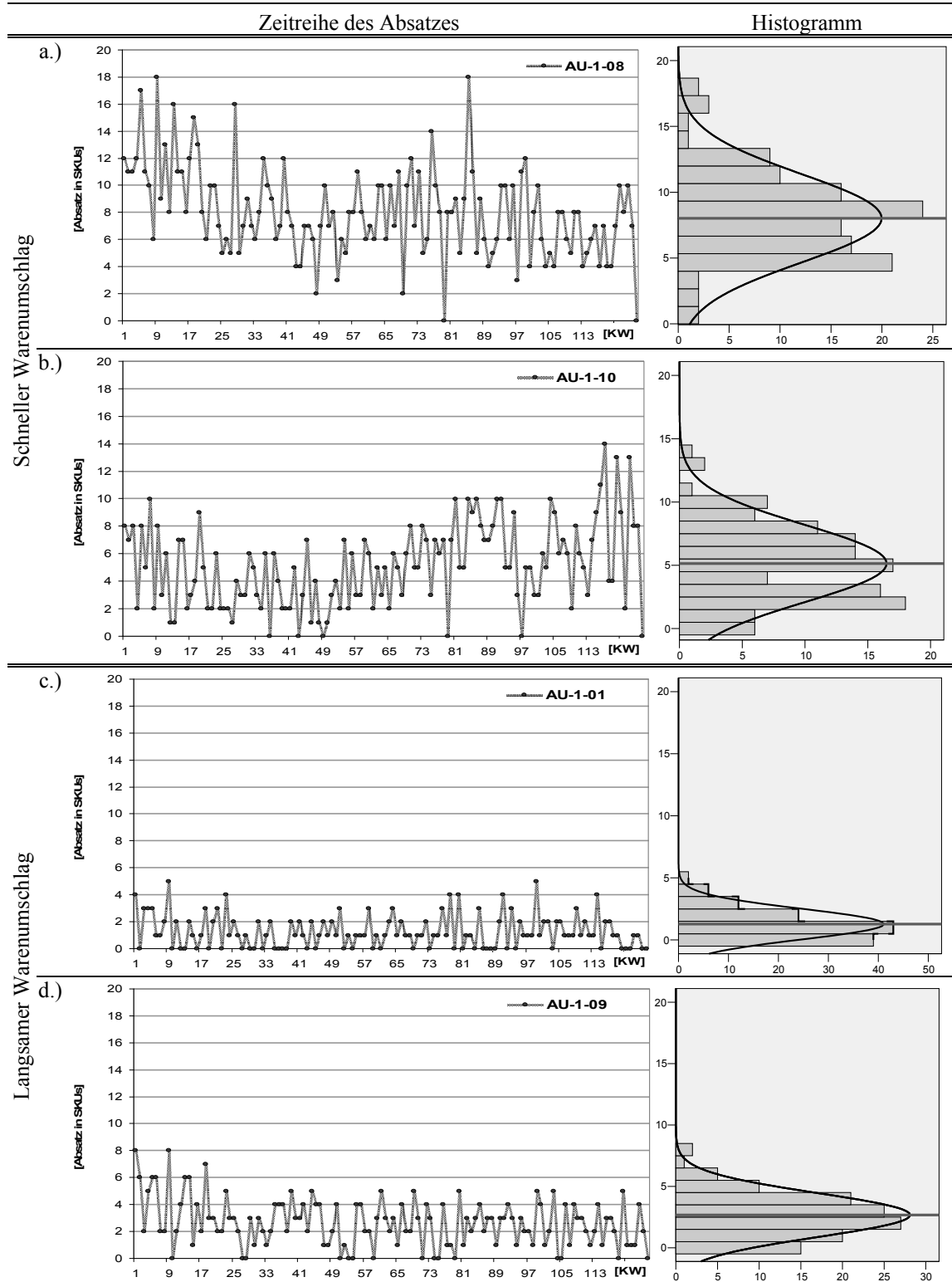


Abbildung 6.11: Zeitreihen von Artikeln mit langsamem Warenumschlag (AU-1-08 (a.) und AU-1-10 (b.)) und mit schnellem Warenumschlag (AU-1-01 (c.) und AU-1-09 (d.))

Die Abbildung zeigt jeweils zwei ausgewählte Zeitreihen mit langsamem und schnellem Warenumschlag an der Absatzstelle AU-1 sowie die Absatzverteilung als Histogramm. Die Zeitreihen weisen erneut lokale und globale Trends und eine mögliche Saisonalität auf, die durch Niveauverschiebungen, Ausreißer, Nullperioden und intensive Zufallsschwankungen überlagert werden. Die Zeitreihenverläufe verdeutlichen ferner, dass auch innerhalb einer Absatzstelle unterschiedliche Verlaufs- und Verteilungsformen von Absatzzahlen existieren. Die Häufigkeitsverteilungen des Absatzes sind überwiegend nicht oder schwach stationär und

folgen nicht einer Normalverteilung, sondern zeigen multimodale Verteilungen durch Nullperioden sowie rechtsschiefe Verteilungen durch Ausreißer und geringe Absatzmengen nahe Null. Auf eine grafische Darstellung und Diskussion aller Absatzzeitreihen wird verzichtet, da dies nur geringe zusätzliche Einsichten verspricht. Abbildungen der Zeitreihenverläufe finden sich im Anhang IX.D.¹

Da die Verteilungsform des Absatzes für die Berechnung der Sicherheitsbestände relevant ist, wird für jede Zeitreihe ein nichtparametrischer Kolmogorov-Smirnow-(KS-)Anpassungstest auf die Verteilungsformen uniforme Verteilung, Normalverteilung, Poissonverteilung und Exponentialverteilung durchgeführt. Die Ergebnisse anhand der kritischen Werte und der Signifikanz der KS-Tests sind in Tabelle 6.5 je Zeitreihe und Absatzstelle aufgeführt.

Tabelle 6.5: Kolmogorov-Smirnow-Tests der Absatzzeitreihen auf unterschiedliche Verteilungsformen

	Absatzzeitreihe												Anz.*	
AU-1	001	002	003	004	005	006	007	008	009	010	011	012		
Normal	0,000	0,051*	0,000	0,120*	0,006	0,000	0,003	0,047	0,000	0,017	0,041	0,200*	3	
Poisson	1,000*	0,334*	0,926*	0,016	0,502*	1,000*	1,000*	0,944*	0,279*	0,916*	0,039	0,468*	10	
Exponential	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0	
Uniform	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0	
AU-2	001	002	003	004	005	006	007	008	009	010	011	012	013	
Normal	0,271*	0,008	0,393*	0,024	0,000	0,001	0,022	0,001	0,115*	0,000	0,078*	0,002	0,078*	5
Poisson	0,864*	0,771*	0,836*	0,210*	0,900*	0,351*	0,757*	0,445*	0,442*	0,009	0,514*	0,387*	0,107*	12
Exponential	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0
Uniform	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000	0,000	0,000	0,012	0
IN-1	001	002	003	004	005	006	007	008	009	010				
Normal	0,000	0,054*	0,073*	0,000	0,021	0,079*	0,036	0,000	0,007	0,112*			4	
Poisson	0,061*	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,005			1	
Exponential	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000			0	
Uniform	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000			0	
IN-2	001	002	003	004	005	006	007	008	009					
Normal	0,000	0,413*	0,213*	0,383*	0,605*	0,219*	0,131*	0,034	0,000				6	
Poisson	0,300*	0,001	0,472*	0,009	0,000	0,000	0,004	0,628*	0,655*				4	
Exponential	0,001	0,000	0,000	0,000	0,000	0,000	0,001	0,001	0,000					
Uniform	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000					
IN-3	001	002	003	004	005	006	007							
Normal	0,211*	0,852*	0,808*	0,529*	0,296*	0,348*	0,014						6	
Poisson	0,000	0,001	0,000	0,006	0,087*	0,000	0,043						1	
Exponential	0,000	0,000	0,000	0,000	0,000	0,001	0,000						0	
Uniform	0,000	0,000	0,001	0,056*	0,000	0,000	0,000						1	

* KS-Test ist signifikant bei 0.05 kritischem Wert (zweiseitiger Test); Anz.* = Anzahl signifikanter KS-Tests je Verteilungsform

Die Testergebnisse identifizieren meist eindeutig eine Verteilungsform, mit vereinzelt doppelt signifikanten Tests von Normal- und Poissonverteilung, zum Beispiel von AU-1-002 oder AU-1-012. Der Bedarf an Außenautomaten folgt überwiegend einer Poissonverteilung durch ihr niedrigeres Absatzvolumen, während an Innenautomaten normalverteilte Zeitreihen überwiegen.

Ferner wird die Struktur der Stichprobe auf Interdependenzen zwischen den Zeitreihen einer Absatzstelle sowie auf hierarchische Interdependenzen mit dem aggregierten Gesamtabsatz der Absatzstelle untersucht. Dazu werden die nichtparametrischen Korrelationskoeffizienten nach SPEARMANs rho berechnet und in Tabelle 4.1 exemplarisch für AU-1 dargestellt.²

¹ Der Datensatz aller transformierten und skalierten Zeitreihen ist – zur Gewährleistung der Anonymität – auf Anfrage vom Verfasser zu beziehen, um eine Replikation der Experimente zu ermöglichen.

² Um systematische Verzerrungen zu vermeiden, entfällt die Berechnung des Korrelationskoeffizienten nach Bravais-Pearson. Korrelationsanalysen der ersten Differenzen instationärer Zeitreihen bestätigen die oben dargestellten Ergebnisse. Auf eine weiterführende Clusteranalyse der Zeitreihen wurde verzichtet, da die gefundenen Erkenntnisse durch befragte Experten mit Domänenwissen als zulässig bestätigt wurden.

Tabelle 6.6: Korrelationskoeffizienten nach SPEARMANS rho der Zeitreihen der Absatzstelle AU-1

Reihe	001	002	003	004	005	006	007	008	009	010	011	012	Gesamt
001	1,000	,053	,069	,061	,085	,180*	,151	,162	,248**	,199*	,068	,050	,237**
002	,053	1,000	,099	,294**	,316**	,350**	,278**	,284**	,234**	,110	,374**	,367**	,663**
003	,069	,099	1,000	,179*	,063	,134	,045	,176*	,122	,290**	,268**	,054	,302**
004	,061	,294**	,179*	1,000	,457**	,137	,266**	,243**	,208*	,107	,089	,128	,559**
005	,085	,316**	,063	,457**	1,000	,104	,301**	,103	,264**	,158	,216*	,255**	,553**
006	,180*	,350**	,134	,137	,104	1,000	,187*	,162	,140	,127	,231**	,250**	,384**
007	,151	,278**	,045	,266**	,301**	,187*	1,000	,362**	,338**	,217*	,029	,016	,440**
008	,162	,284**	,176*	,243**	,103	,162	,362**	1,000	,293**	,244**	,144	-,026	,500**
009	,248**	,234**	,122	,208*	,264**	,140	,338**	,293**	1,000	,211*	,169	,072	,468**
010	,199*	,110	,290**	,107	,158	,127	,217*	,244**	,211*	1,000	,103	,024	,361**
011	,068	,374**	,268**	,089	,216*	,231**	,029	,144	,169	,103	1,000	,374**	,554**
012	,050	,367**	,054	,128	,255**	,250**	,016	-,026	,072	,024	,374**	1,000	,446**
Gesamt	,237**	,663**	,302**	,559**	,553**	,384**	,440**	,500**	,468**	,361**	,554**	,446**	1,000
Anz.**	3	8	4	6	6	5	7	6	6	5	5	4	12

*Korrelation ist signifikant bei 0.05 Niveau (zweiseitiger Test); **Korrelation ist hochsignifikant bei 0.01 Niveau (zweiseitiger Test)

Anz.** = Anzahl hochsignifikanter Korrelationskoeffizienten

Wie zu erwarten sind alle Zeitreihen mit dem Gesamtabsatz des Automaten positiv korreliert. Darüber hinaus sind nur vereinzelte Gruppen von Zeitreihen positiv miteinander korreliert. Daraus sind zwei Erkenntnisse für den Verfahrensvergleich abzuleiten. Da der Absatzverlauf eines Großteils der Waren untereinander nicht korreliert ist, zeigt er eigenständige zeitliche Entwicklungen und Muster. Dies verhindert eine hierarchische Dekomposition zur Prognose (engl. *top-down modelling*), da untergeordnete Zeitreihen unterschiedliche lokale Muster aufweisen als die aggregierte Zeitreihe der Absatzstelle. Somit können hierarchische Ansätze der Absatzprognose vernachlässigt werden. Ferner sind alle signifikanten Korrelationen positiv, so dass keine Substitutionseffekte zwischen einzelnen Artikeln festzustellen sind. Daher wird nachfolgend auf die Modellierung von Interaktionen zwischen den Zeitreihen im Experimentdesign verzichtet.¹ Diese Ergebnisse werden auf allen Absatzstellen bestätigt; ihre Korrelationsanalysen finden sich aus Platzgründen im Anhang IX.E.

6.2.2.2.3 Kategorisierung der Zeitreihen

Nachfolgend werden die Zeitreihen anhand ihrer statistischen Lage- und Streuungsparameter beschrieben sowie durch die Zuordnung zu empirischen Kategorien des Absatzverlaufes charakterisiert. Die Kategorisierung erfolgt nach den warenlogistischen Kriterien der Anwendungsdomäne anhand der Struktur des Zeitreihenverlaufes, der Geschwindigkeit des Warenumschlages und der Bedeutung der Ware im Sortiment.

Der Zeitreihenverlauf gilt als sporadisch, wenn über 30% aller Absatzperioden keinen Absatz aufweisen, $y_t = 0$, was auch als Nullperiode (NuP) bezeichnet wird. Eine Zeitreihe wird als unregelmäßig kategorisiert, wenn der Variationskoeffizient (VaKo) als Verhältnis von Standardabweichung zu Mittelwert größer als 60% ist, mit $VaKo > 0,6$. Andernfalls, mit $VaKo \leq 0,6$, liegt ein regelmäßiger Absatzverlauf vor. Die Höhe des Variationskoeffizienten korreliert negativ mit dem durchschnittlichen Umsatz, so dass Waren mit vergleichsweise niedrigem Absatz einen höheren Variationskoeffizienten aufweisen und umgekehrt. Die Kategorisierung nach der Umschlaggeschwindigkeit der Waren erfolgt in so genannte Schnelldreher und Langsamdreher. Schnelldreher weisen mindestens einen monatlichen Warenumschlag und somit einen 12-fachen Warenumschlag pro Jahr auf, was in der vorliegenden Anwendung näherungsweise einem durchschnittlichen Mindestabsatz von 9,5

¹ Eine Kreuzkorrelationsanalyse der zeitlich verzögerten Korrelationen für unterschiedliche Lead- und Lag-Effekte von $t-60$ bis $t+60$ lieferte ebenfalls nur vereinzelt signifikante positive Korrelationen mit zeitgleicher Wirkung im Zeitpunkt t . Es konnten keine signifikanten dynamischen Kreuzkorrelationen identifiziert werden, die den Absatz eines einzelnen Artikels u. U. durch verzögerte Realisationen des Gesamtabsatzes der Absatzstelle erklären ließen. Dies bestätigt die oben gefundenen Erkenntnisse.

Verpackungseinheiten pro Woche entspricht. Langsamdreher haben einen geringeren Warenumschlag und einen Mindestumsatz von weniger als 9,5 Verpackungseinheiten. Die warenlogistische Bedeutung im Sortiment wird durch eine ABC-Klassifikation der Artikel kategorisiert.¹ Die umsatzstärksten A-Artikel korrespondieren mit den Schnelldreher, während die Langsamdreher weiter in B- und C-Artikel mit einem Mindestabsatz unter 5 Verpackungseinheiten pro Woche unterteilt werden können. Abbildung 6.12 zeigt die ABC-Analyse der 51 Zeitreihen.

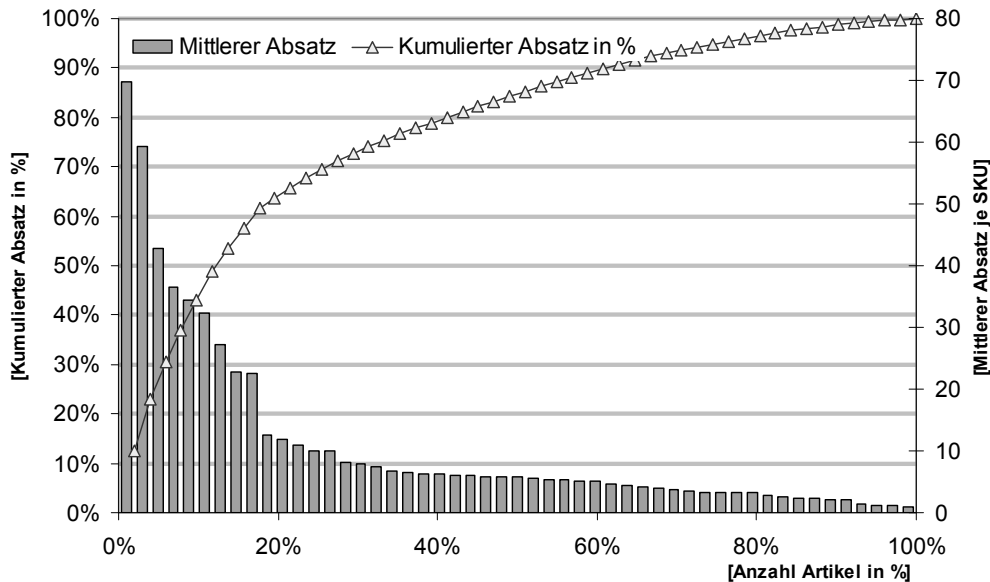


Abbildung 6.12: ABC-Analyse des mittleren Absatzes über alle Abverkaufsstellen

Die Konzentrationskurve verdeutlicht die spezifische Struktur der Zeitreihen der Stichprobe, mit einem starken Anteil von A- und C-Artikeln im Sortiment. Die oberen 20% aller Artikel erbringen 64% des Absatzes, während die unteren 60% aller Artikel lediglich 20% des Umsatzes erwirtschaften.

Die Kategorisierung der Zeitreihen nach den Lage- und Streuungsparametern sowie der Charakteristika der Zeitreihen zeigt Tabelle 6.7 absteigend sortiert nach dem mittleren Absatz.

¹ Die ABC-Analyse ist ein Standard-Instrument der Materialwirtschaft. Für eine Einführung siehe [Günther und Busch 1965; Klein 1976], im Kontext von Logistik und Handel [Tempelmeier 2003; Müller-Hagedorn 1998].

Tabelle 6.7: Lageparameter und Klassifikation der Zeitreihen nach Absatzstellen

Zeitreihe	Parameter der Zeitreihe							Verlauf der Zeitreihen					
	Min	Max	Mw	Med	StdAb	Skew	Kurt	VaKo	NuP	NuP %	Verlauf	Umschlag	ABC
AU-1-08	0	18	8,0	8,0	3,4	0,5	0,8	41,8%	2	1,6%	regelmäßig	langsam	B
AU-1-04	0	23	5,9	5,0	3,8	1,3	3,5	65,5%	7	5,6%	unregelmäßig	langsam	B
AU-1-02	0	15	5,8	5,0	3,3	0,6	0,0	56,4%	4	3,2%	regelmäßig	langsam	B
AU-1-05	0	14	5,5	5,0	3,0	0,5	-0,3	54,5%	2	1,6%	regelmäßig	langsam	B
AU-1-10	0	14	5,1	5,0	3,0	0,4	-0,2	59,3%	6	4,8%	regelmäßig	langsam	B
AU-1-11	0	10	4,4	4,0	2,5	0,3	-0,5	56,7%	5	4,0%	regelmäßig	langsam	C
AU-1-03	0	13	3,3	3,0	2,2	1,3	2,8	66,0%	7	5,6%	unregelmäßig	langsam	C
AU-1-07	0	8	2,9	3,0	1,8	0,7	0,5	64,3%	10	7,9%	unregelmäßig	langsam	C
AU-1-09	0	8	2,7	3,0	1,8	0,5	0,1	67,1%	15	11,9%	unregelmäßig	langsam	C
AU-1-08	0	9	2,3	2,0	1,8	0,9	0,9	79,5%	18	14,3%	unregelmäßig	langsam	C
AU-1-06	0	8	2,0	2,0	1,4	1,0	1,7	70,8%	14	11,1%	unregelmäßig	langsam	C
AU-1-01	0	5	1,3	1,0	1,2	0,9	0,4	95,8%	39	31,0%	sporadisch	langsam	C
AU-2-04	0	15	8,0	8,0	2,9	-0,4	0,2	36,4%	3	2,4%	regelmäßig	langsam	B
AU-2-09	0	14	6,7	6,0	3,1	0,2	-0,4	46,5%	3	2,4%	regelmäßig	langsam	B
AU-2-13	0	14	6,6	6,0	3,4	0,2	-0,7	51,7%	4	3,2%	regelmäßig	langsam	B
AU-2-03	0	15	6,3	6,0	2,8	0,0	0,2	44,8%	4	3,2%	regelmäßig	langsam	B
AU-2-01	0	15	6,1	6,0	3,0	0,4	0,6	48,6%	3	2,4%	regelmäßig	langsam	B
AU-2-07	0	12	5,2	5,0	2,8	0,4	-0,3	53,1%	4	3,2%	regelmäßig	langsam	B
AU-2-11	0	10	4,0	4,0	2,3	0,3	-0,4	57,8%	6	4,8%	regelmäßig	langsam	C
AU-2-02	0	10	3,7	3,0	2,3	0,4	-0,4	61,3%	10	7,9%	unregelmäßig	langsam	C
AU-2-08	0	12	3,4	3,0	2,3	0,9	1,0	67,2%	9	7,1%	unregelmäßig	langsam	C
AU-2-12	0	10	3,3	3,0	2,2	0,6	-0,1	66,4%	10	7,9%	unregelmäßig	langsam	C
AU-2-06	0	8	2,2	2,0	1,9	0,9	0,7	84,2%	24	19,0%	unregelmäßig	langsam	C
AU-2-05	0	6	1,5	1,0	1,3	0,8	0,3	88,5%	33	26,2%	unregelmäßig	langsam	C
AU-2-10	0	7	1,2	1,0	1,6	1,7	2,8	133,8%	57	45,2%	sporadisch	langsam	C
IN-1-02	0	235	69,9	70,0	40,6	0,6	2,1	58,1%	16	12,7%	regelmäßig	schnell	A
IN-1-05	0	115	22,8	18,5	20,1	2,0	6,0	88,0%	16	12,7%	unregelmäßig	schnell	A
IN-1-10	0	36	12,6	13,0	7,3	0,2	0,4	57,7%	16	12,7%	regelmäßig	schnell	A
IN-1-07	0	51	11,7	10,0	9,6	1,3	2,7	82,0%	17	13,5%	unregelmäßig	schnell	A
IN-1-06	0	27	7,3	7,0	5,6	1,0	1,5	76,4%	17	13,5%	unregelmäßig	langsam	B
IN-1-09	0	25	6,3	5,0	6,0	1,0	0,5	96,4%	27	21,4%	unregelmäßig	langsam	B
IN-1-08	0	34	5,5	3,0	6,8	1,6	2,6	124,1%	41	32,5%	sporadisch	langsam	B
IN-1-03	0	18	4,7	4,0	3,9	0,9	0,5	84,2%	21	16,7%	unregelmäßig	langsam	C
IN-1-04	0	28	4,3	3,0	5,1	2,0	5,2	120,0%	31	24,6%	unregelmäßig	langsam	C
IN-1-01	0	9	0,9	0,0	1,4	2,6	10,2	148,5%	54	42,9%	sporadisch	langsam	C
IN-2-02	0	106	59,3	61,0	19,2	-0,5	1,7	32,5%	4	3,2%	regelmäßig	schnell	A
IN-2-06	0	75	36,5	35,0	13,6	0,0	0,9	37,3%	4	3,2%	regelmäßig	schnell	A
IN-2-04	0	65	34,4	35,0	11,5	-0,4	1,5	33,4%	4	3,2%	regelmäßig	schnell	A
IN-2-05	0	62	32,4	32,5	12,0	-0,2	0,7	37,0%	4	3,2%	regelmäßig	schnell	A
IN-2-07	0	19	6,1	5,5	4,1	0,5	-0,3	66,7%	9	7,1%	unregelmäßig	langsam	B
IN-2-03	0	12	5,3	5,5	2,8	0,1	-0,5	52,4%	5	4,0%	regelmäßig	langsam	B
IN-2-08	0	13	3,3	3,0	2,3	1,0	1,9	70,4%	10	7,9%	unregelmäßig	langsam	C
IN-2-01	0	11	3,3	3,0	2,5	1,0	0,7	75,3%	13	10,3%	unregelmäßig	langsam	C
IN-2-09	0	8	2,2	2,0	1,9	1,2	1,3	86,1%	22	17,5%	unregelmäßig	langsam	C
IN-3-01	0	110	42,9	42,0	17,2	0,2	2,3	40,1%	6	4,8%	regelmäßig	schnell	A
IN-3-02	0	50	27,2	26,0	11,5	-0,1	0,1	42,2%	6	4,8%	regelmäßig	schnell	A
IN-3-03	0	43	22,5	22,0	10,2	-0,1	-0,1	45,3%	7	5,6%	regelmäßig	schnell	A
IN-3-06	0	40	10,9	10,0	7,0	1,0	1,9	63,8%	7	5,6%	unregelmäßig	schnell	A
IN-3-04	0	21	10,1	10,0	5,3	0,1	-0,6	52,8%	6	4,8%	regelmäßig	schnell	A
IN-3-05	0	23	9,9	10,0	5,0	0,3	-0,1	50,4%	6	4,8%	regelmäßig	schnell	A
IN-3-07	0	26	5,7	5,0	4,1	1,7	5,6	71,3%	7	5,6%	unregelmäßig	langsam	B

Min=Minimum; Max=Maximum; Mw=Mittelwert; Md=Median; StdAb=Standardabweichung; Skew=Schiefe; Kurt=Kurtosis;
VaKo=Variationskoeffizient; NuP= Anzahl von Nullperioden; NuP %=Anzahl Nullperioden in % von allen Beobachtungswerten

Die Heterogenität der Zeitreihen durch abweichende Mittelwerte, Intensitäten der Streuung, der Verteilung in Form von Schiefe und Wölbung der Waren an einem Automatenstandort bestätigt die gering Korrelation der Zeitreihen untereinander aus der Analyse in Abschnitt

6.2.2.2.2. Dies erfordert einen eigenständigen Dispositionsprozess für jede Zeitreihe, einschließlich einer artikelgenauen Absatzprognose der einzelnen Zeitreihe durch die Identifikation des Zeitreihenverlaufes, die Auswahl eines passenden Prognosemodells und seine Parametrisierung sowie die Berechnung der Residuen der Prognose und die Analyse ihrer Verteilungsform, die Bestimmung einer passenden Berechnungsvorschrift der Sicherheitsbestände und schließlich die Kalkulation der optimalen Bestandshöhe. Dies verhindert eine Vereinfachung durch die Formulierung eines einheitlichen Prognosemodells, einheitlicher Verfahrensparameter oder einer konstanten Verteilungsform für die Bestellmengenrechnung.

Die Kategorisierung der Zeitreihen in Tabelle 6.7 erlaubt eine Zusammenfassung der Charakteristika der Zeitreihenverläufe über alle Zeitreihen einer Absatzstelle sowie über die Standorte der Innenautomaten und der Außenautomaten in Tabelle 6.8.¹

Tabelle 6.8: Mittlere Lageparameter und Zeitreihenverläufe nach Absatzstelle und Standort

Standort	Parameter der Zeitreihe							Verlauf der Zeitreihen					
	Min	Max	Mw	Med	StdAb	Skew	Kurt	VaKo	NuP	NuP %	Verlauf	Umschlag	ABC
Mw AU-1	0,0	12,1	4,1	3,8	2,4	0,8	0,8	64,8%	11	8,5%	5 regelmäßig 6 unregelmäßig 1 sporadisch	12 langsam 0 schnell	0 A 5 B 7 C
Mw AU-2	0,0	11,4	4,5	4,2	2,4	0,5	0,3	64,6%	13	10,4%	7 regelmäßig 5 unregelmäßig 1 sporadisch	13 langsam 0 schnell	0 A 6 B 7 C
Mw Außengeräte	0,0	11,7	4,3	4,0	2,4	0,6	0,5	64,7%	12	9,5%	12 regelmäßig 11 unregelmäßig 2 sporadisch	25 langsam 0 schnell	0 A 11 B 14 C
Mw IN-1	0,0	57,8	14,6	13,4	10,6	1,3	3,2	93,6%	26	20,3%	2 regelmäßig 6 unregelmäßig 2 sporadisch	6 langsam 4 schnell	4 A 3 B 3 C
Mw IN-2	0,0	41,2	20,3	20,3	7,8	0,3	0,9	54,5%	8	6,6%	5 regelmäßig 4 unregelmäßig 0 sporadisch	5 langsam 4 schnell	4 A 2 B 3 C
Mw IN-3	0,0	44,7	18,4	17,9	8,6	0,4	1,3	52,3%	6	5,1%	5 regelmäßig 2 unregelmäßig 0 sporadisch	1 langsam 6 schnell	6 A 1 B 0 C
Mw Innengeräte	0,0	47,9	17,8	17,2	9,0	0,7	1,8	66,8%	13	10,7%	12 regelmäßig 12 unregelmäßig 2 sporadisch	12 langsam 14 schnell	14 A 6 B 6 C
Mw Automaten	0,0	33,4	12,4	11,9	6,4	0,7	1,3	66,0%	13	10,2%	24 regelmäßig 23 unregelmäßig 4 sporadisch	37 langsam 14 schnell	14 A 17 B 20 C

Min=Minimum; Max=Maximum; Mw=Mittelwert; Md=Median; StdAb=Standardabweichung; Skew=Schiefe; Kurt=Kurtosis;
VaKo=Variationskoeffizient; NuP= Anzahl von Nullperioden; NuP %=Anzahl Nullperioden in % von allen Beobachtungswerten

Die Stichprobe aller Zeitreihen enthält 24 regelmäßige (47%), 23 unregelmäßige (45%) und 4 sporadische Zeitreihenverläufe (9%). Alternativ können in der Stichprobe 14 Schnelldreher (27%) und 37 Langsamdreher (73%) gemäß ihrer Warenumschlagsgeschwindigkeit unterschieden werden. Die ABC-Klassifikation unterscheidet die Langsamdreher weiter in 17 B-Artikel (33%) und 20 C-Artikel (39%) im Vergleich zu den 14 A-Artikeln (27%), die mit den Schnelldrehern übereinstimmen.

Aus der Analyse und Kategorisierung der Zeitreihen bestätigt sich, dass der Absatz an Innenautomaten signifikant andere Merkmale aufweist als der an Außenautomaten.² Die

¹ Dies wird nachfolgend durch den Zusatz „mw“ bei Parametern, Fehlermaßen und Zeitreihen angezeigt.

² Unter Ausschluss der Zeitreihe IN-3, die später für die Parametrisierung von NN verwendet wird, enthält die Stichprobe 4 sporadische Zeitreihenverläufe (9%), 19 regelmäßige Verläufe (43%) und 21 unregelmäßige

Verteilung der Zeitreihenverläufe regelmäßiger, unregelmäßiger und sporadischer Verläufe ist mit (12-11-2) zu (12-12-2) nahezu identisch. Allerdings enthalten Außenautomaten ausschließlich Waren mit langsamem Artikelumschlag (25), während Innengeräte sowohl Langsamdreher (12) als auch Schnelldreher (14) aufweisen. Der signifikant höhere durchschnittliche Absatz pro Woche an Innengeräten von 17,2 Verpackungseinheiten im Vergleich zu 4,3 Verpackungseinheiten an Außengeräten ist somit auf die Schnelldreher zurückzuführen. Die Zeitreihen der Absatzstellen beider Standorte besitzen Nullperioden mit Wochen ohne Verkauf einer einzigen Verpackungseinheit. Standardabweichung und Variationskoeffizient sind bei Außenautomaten nahezu konstant, schwanken jedoch bei Innengeräten und unterscheiden sich signifikant voneinander. Die hohen Variationskoeffizienten um 66% verdeutlichen die erhöhte Volatilität der Absatzzeitreihen der einzelnen Waren im Vergleich zum Gesamtabsatz eines Automaten von 37% (siehe Tabelle 6.4) und somit die gesteigerte durch die artikelgenaue Absatzprognose auf Filialebene.

Die Zuordnung der einzelne Zeitreihen nach den Kategorisierungsmerkmalen Absatzverlauf, Warenumsschlag und ABC-Klassifikation sind in Tabelle 6.9 dargestellt.

Tabelle 6.9: Mittlere Lageparameter der Kategorien von Zeitreihen

Waren- umschlag	Mittlere Lageparameter							Klassifikation des Zeitreihenverlaufes					
	Min	Max	Mw	Med	StdAb	Kurt	VaKo	NuP	NuP %	Verlauf	Umschlag	ABC	
Schnell	0,0	235	28,8	28,2	13,6	1,4	0,5	8,5	6,7%	11 regelmäßig	14 schnell	13 A	
										3 unregelmäßig	0 langsam	0 B	
										0 sporadisch		0 C	
Langsam	0,0	34	4,4	3,9	2,9	1,1	0,7	15,0	11,9%	13 regelmäßig	0 schnell	0 A	
										20 unregelmäßig	37 langsam	17 B	
										4 sporadisch		20 C	
A-Artikel	0,0	73,7	28,8	13,6	0,5	1,4	28,2	8,5	6,7%	11 regelmäßig	14 schnell	14 A	
										3 unregelmäßig	0 langsam	0 B	
										0 sporadisch		0 C	
B-Artikel	0,0	18,4	6,2	3,8	0,6	0,7	5,6	8,7	6,9%	11 regelmäßig	0 schnell	0 A	
										5 unregelmäßig	17 langsam	17 B	
										1 sporadisch		0 C	
C-Artikel	0,0	10,6	2,8	2,2	0,8	1,5	2,5	20,4	16,2%	2 regelmäßig	0 schnell	0 A	
										15 unregelmäßig	20 langsam	20 B	
										3 sporadisch		0 C	
Sporadisch	0,0	16,7	2,5	3,2	1,4	5,2	1,3	50,7	40,2%	0 regelmäßig	0 schnell	0 A	
										0 unregelmäßig	3 langsam	1 B	
										3 sporadisch		2 C	
Unregelmäßig	0,0	20,9	5,1	4,0	0,8	1,6	4,5	16,4	13,0%	0 regelmäßig	3 schnell	3 A	
										23 unregelmäßig	21 langsam	5 B	
										0 sporadisch		16 C	
Regelmäßig	0,0	41,8	18,1	8,0	0,5	0,3	17,9	5,4	4,3%	24 regelmäßig	11 schnell	11 A	
										0 unregelmäßig	13 langsam	11 B	
										0 sporadisch		2 C	

Min=Minimum; Max=Maximum; Mw=Mittelwert; Md=Median; StdAb=Standardabweichung; Skew=Schiefe; Kurt=Kurtosis; VaKo=Variationskoeffizient; NuP= Anzahl von Nullperioden; NuP %=Anzahl Nullperioden in % von allen Beobachtungswerten

Die drei Kategorien erlauben eine mehrfache Aufteilung der Zeitreihen in der Stichprobe und somit die Beurteilung der Prognose- und Dispositionsverfahren in unterschiedlichen Gruppen von empirisch relevanten Zeitreihen. Die gebildeten Kategorien sind jedoch nicht vollständig disjunkt, wie ihre Darstellung als morphologischer Kasten in Tabelle 6.10 verdeutlicht.

Verläufe (48%). Darüber hinaus enthält sie 8 Schnelldreher (18%) und 36 Langsamdreher (82%) sowie 20 C-Artikel (45%) und 16 B-Artikel (36%) im Vergleich zu 8 A-Artikeln (18%).

Tabelle 6.10: Systematik der Zeitreihenkategorien

Automaten-Stellplatz	Außenautomaten		Innenautomaten	
Abverkaufsstelle	AU-1	AU-2	IN-1	IN-2
Absatzverlauf der Zeitreihe	Regelmäßiger Verlauf		Unregelmäßiger Verlauf	Sporadischer Verlauf
Umschlagsgeschwindigkeit	Schnelldreher	Langsamdreher		
ABC-Klassifikation	A-Artikel	B-Artikel		C-Artikel

Vielmehr repräsentieren die Klassen eine schrittweise Verfeinerung der Kategorisierung zum Auffinden homogener Gruppen von Zeitreihen, auf denen einzelne Verfahren oder Verfahrensklassen eine hohe Prognosegenauigkeit zeigen.

6.2.2.3 Aufteilung der Datenbasis

Die Datenbasis der Zeitreihen muss für einen validen Quasi-ex-ante-Verfahrensvergleich vor Beginn der Experimente in entsprechende Datenteilmengen der Kalibrierungsstichprobe und der Teststichprobe unterteilt werden [Küsters 2006].

Die Testdaten dienen der Out-of-sample-Beurteilung der Generalisierbarkeit aller Verfahren. Die Kalibrierungsstichprobe kann hingegen abweichend verwendet werden. Für NN wird jede Zeitreihe der Kalibrierungsstichprobe in zwei Datenteilmengen der Trainingsmenge und der Validierungsmenge aufgeteilt. Die Trainingsmenge wird zur Parametrisierung und die Validierungsmenge zur Kontrolle des Lernprozesses durch *early stopping* und die Auswahl der Verfahrenskandidaten aus mehrfachen Initialisierungen verwendet. Für statistische Prognoseverfahren mit aggregierter Modellselektion ist keine Validierungsmenge erforderlich, so dass die gesamte Kalibrierungsstichprobe zum Training zur Verfügung steht. Verfahren der individuellen Modellselektion je Zeitreihe können in den internen Routinen des Expertensystems zur Parameterfindung und Modellselektion ebenfalls unterschiedliche, proprietäre Aufteilungen der Kalibrierungsstichprobe verwenden.

Mit der Stichprobe liegen für 51 Zeitreihen artikelgenaue wöchentliche Absatzzahlen über einen Zeitraum von 126 Wochen vor, was einer Dauer von 31 Monaten bzw. 2½ Jahren entspricht. Die ersten 105 Beobachtungswerte aus den Jahren 2001 und 2002 werden für die Parametrisierung der Verfahren in der In-sample-Kalibrierungsstichprobe verwendet. Hiervon werden die ersten 69 Beobachtungswerte zum Training und die nachfolgenden 36 Werte zur Validierung benutzt. Die letzten 21 Beobachtungswerte, in Form der ersten 21 Wochen des Jahres 2003, werden für die Teststichprobe zur Out-of-sample-Beurteilung verwendet.

Die Beurteilung der Generalisierungsfähigkeit erfolgt gemäß der Datenaufteilung auf 21 Beobachtungswerte durch mehrfache $t+1$ Prognosen aus 21 Prognoseursprüngen t_u . Damit wird der Kritik der M3-Competition Rechnung getragen, wonach die Ergebnisse eines Verfahrensvergleichs nachhaltig durch die Auswahl des Zeitreihenursprungs beeinflusst werden. Daher sollten mittlere Gütekriterien über mehrere Zeitreihenursprünge berechnet werden, um so die Auswirkungen unterschiedlicher Datenteilmengen zu verringern.¹

¹ Es erfolgt jedoch keine erneute Parametrisierung der Verfahren an den neuen Zeitreihenursprüngen, da dies zu einem unverhältnismäßig hohen Rechenaufwand geführt hätte.

6.2.3 Verfahren zur Bedarfsprognose

6.2.3.1 Auswahl der Prognoseverfahren

Ein valider Verfahrensvergleich von NN zur Prognose nach Hypothese H_a und zur Disposition nach H_b und H_c erfordert die Beurteilung ihrer Genauigkeit in Relation zu etablierten Benchmark-Verfahren, unter Berücksichtigung der gegebenen Datenkonstellation.

Nach FILDES [1989] und den Hypothesen H_5 und H_6 können univariate Verfahren der Zeitreihenanalyse nach individueller und aggregierter Modellselektion, mit konstanten oder optimalen Parametern, unterschieden werden. Die explorative Datenanalyse hat darüber hinaus den Einfluss exogener Ereignisse auf den Absatzverlauf an allen Absatzstellen und auf alle Zeitreihen verdeutlicht. Diese Effekte können durch Verfahren der Zeitreihenanalyse vernachlässigt oder explizit durch Verfahren der Kausalanalyse als Interventionen modelliert werden. Die Studie berücksichtigt sowohl univariate Verfahren der Zeitreihenprognose als auch Verfahren der kausalanalytischen Interventionsmodellierung, die zusätzlich zur Zeitreihe der abhängigen Variablen auch exogene Einflüsse auf den Absatz durch technische Störungen des Automaten und urlaubsbedingte Schließungen der Absatzstelle durch Dummy-Variablen berücksichtigen. Multivariate Verfahren der Kausalanalyse erfordern grundsätzlich eine individuelle Modellierung. Die Experimente beurteilen die Prognose- und Dispositionsgüte von zwei NN im Vergleich zu zehn statistischen Prognoseverfahren bzw. Softwaresystemen für einen Prognosehorizont von $h = 1$ über mehrere rollierende Zeitreihenursprünge. Tabelle 6.11 zeigt eine Übersicht der Verfahrensklassen.

Tabelle 6.11: Dimensionen des Verfahrensvergleichs nach Prognosemodell und -verfahren

Modellselektion	Umfang exogener Zeitreihen	
	Univariate Verfahren	Multivariate Verfahren
Aggregierte Modellselektion	A.) Univariate statistische Verfahren	C.) Multivariate statistische Verfahren
	Univariate NN	Multivariate NN
Individuelle Modellselektion	B.) Univariate Expertensysteme	D.) Multivariate Expertensysteme

Dies erlaubt einen umfassenden Verfahrensvergleich zwischen NN und statistischen Verfahren zur Prognose sowie zur Disposition durch den paarweisen Vergleich von:¹

- univariaten NN mit univariaten Prognoseverfahren (A-A),
- univariaten NN mit univariaten Expertensystemen (A-B),
- multivariaten NN mit multivariaten Prognoseverfahren (C-C),
- multivariaten NN mit multivariaten Expertensystemen (C-D),
- multivariaten NN mit univariaten NN (C-D).

Während der Vergleich von univariaten, statistischen Verfahren mit NN auf einzelnen Zeitreihen verbreitet ist (A-A,B), steht eine Beurteilung auf einer repräsentativen Stichprobe von empirischen Zeitreihen des Handels noch aus.² Darüber hinaus fehlen jegliche Vergleiche zur Genauigkeit kausalanalytischer statistischer Verfahren mit entsprechenden Architekturen von NN sowie univariater mit multivariaten NN (C-A,B,C,D) auf einem konsistenten Datensatz einer Anwendungsdomäne. Mit dieser Studie wird erstmalig ein empirischer Verfahrensvergleich auf Absatzdaten an Warenautomaten durchgeführt, so dass bereits aus

¹ Ein Erkenntniszuwachs aus dem Vergleich univariater NN mit multivariaten statistischen Verfahren erscheint unwahrscheinlich und unterbleibt daher.

² Vgl. hierzu die Literaturrecherche zur Absatzprognose mit NN in Abschnitt 4.4.1.3.

der relativen Güte der statistischen Prognoseverfahren ein unmittelbarer Erkenntnisgewinn ausgeht.¹ Die Erweiterung bisheriger Studien um multivariate Verfahren erfolgt in der Tradition von MAKRIDAKIS et al. als Erweiterung von einfachen auf komplexe Verfahren, die entsprechend im Sinne der Hypothese H_1 diskutiert werden können.

Der gewählte Versuchsaufbau bedingt jedoch die Entstehung systematischer Prognosefehler. Alle univariaten Verfahren ohne Modellselektion werden ohne Berücksichtigung der zugrunde liegenden Struktur der Zeitreihe mit ex ante festgelegten und über alle Zeitreihen identischen Parametern berechnet. Aufgrund der fehlenden Anpassung an die Struktur der Zeitreihen entstehen systematische Fehler auf jenen Zeitreihen, die nicht mit der Struktur des Verfahrens übereinstimmen.² Bei Prognoseverfahren ohne Modellselektion mit konstanten Parametern besteht zudem das Risiko systematischer Fehler durch die Fehlspezifikation der Verfahrensparameter, selbst für Zeitreihen mit einer passenden Verlaufsform. Dies erfolgt jedoch bewusst in Anlehnung an Hypothese H_5 und entspricht der verbreiteten Anwendung einzelner Prognoseverfahren in der Praxis. Dagegen streben die Heuristiken des Expertensystems durch korrekte Modellspezifikation und Parametrisierung eine bestmögliche Anpassung an die jeweilige Zeitreihenstruktur an, um gerade Fehler durch Modellfehlspezifikation oder Fehlparametrisierung zu vermeiden. NN hingegen bilden die Schritte der Modellbildung und Parametrisierung an die Struktur der jeweiligen Zeitreihe im Trainingsprozess ab und umgehen so das Problem einer Fehlspezifikation des Modells.

Die auszuwählenden statistischen Verfahren sollten entweder in der Theorie als Benchmark oder in der Praxis durch regelmäßige Anwendung etabliert sein. Sie können anhand der Erkenntnisse aus früheren Vergleichsstudien (siehe Abschnitt 3.3.3) sowie der Marktstudie von in Industrie und Handel verwendeten Verfahren zur Bedarfsprognose (siehe Abschnitt 3.1.3) bestimmt werden. Demnach werden insbesondere Naive Verfahren, Verfahren der Exponentiellen Glättung (ES) und ARIMA-Verfahren berücksichtigt, die sowohl ohne Modellselektion und Parameteroptimierung als auch optimiert durch ein Expertensystem (EXP) modelliert werden. Als Kandidaten der NN werden aufgrund ihrer weiten Verbreitung und hohen Bekanntheit Architekturen der MLPs verwendet.³ Abbildung 6.13 ordnet die einbezogenen statistischen Verfahren und Modelle der NN in die Klassifikation der Prognoseverfahren aus Abbildung 3.12 ein.

Die eingesetzten Verfahren mit aggregierter Modellselektion werden grundsätzlich auf allen Zeitreihen angewendet; sie sind hell schattiert dargestellt. Verfahren mit dunkler Schattierung sind Kandidaten der individuellen Modellselektion, die gemäß der Struktur der jeweiligen Zeitreihe von dem Expertensystem ausgewählt und angewendet werden könnten.⁴ Abbildung 6.13 verdeutlicht, dass Benchmark-Verfahren aus allen relevanten Verfahrensklassen in die Vergleichsstudie einbezogen werden, mit einem Schwerpunkt auf univariaten Verfahren.

Insgesamt werden neun statistische Prognoseverfahren und Expertensysteme mit zwei Architekturen der NN auf 51 Zeitreihen verglichen. Die Beurteilung der Prognosegenauigkeit

¹ Darüber hinaus erweitert der vorliegende Versuchsaufbau den Erkenntnisstand konventioneller statistischer Prognoseverfahren durch einen Verfahrenvergleich zwischen zeitreihenanalytischen und multivariaten, kausalanalytischen Prognoseverfahren, insbesondere durch den Vergleich von univariaten mit multivariaten Prognoseverfahren gemäß MAKRIDAKIS [Makridakis und Hibon 2000] sowie multivariaten Prognoseverfahren mit multivariaten Expertensystemen gemäß FILDES [FILDES 1989]. Diese stehen jedoch nicht im Mittelpunkt der Untersuchung.

² Beispielsweise führt die Anwendung der S.ES auf einer saisonalen Zeitreihe zwangsläufig zu systematischen Prognosefehlern, da die Saisonalität der Zeitreihe nicht vom Verfahren berücksichtigt werden kann.

³ MLPs stellen lediglich eine Klasse vorwärtsgerichteter NN dar. Erkenntnisse über die Genauigkeit der Experimente in der Prognose und Disposition können somit nicht ohne Weiteres auf andere vorwärtsgerichtete Netzwerktypen, rekurrente Architekturen oder die Verfahrensklasse der NN als Ganzes übertragen werden.

⁴ Verfahren mit geteilter Schattierung sind somit Verfahren, die auf allen Zeitreihen eingesetzt werden und zudem durch Auswahl im Expertensystem zum Einsatz kommen können.

erfordert so den Vergleich von 11.781 Einzelprognosen auf den 21 Beobachtungswerten der Testmenge und 58.905 Einzelprognosen auf der Kalibrationsstichprobe zur Beurteilung der Prognosegüte. Darüber hinaus werden alle Prognoseverfahren zuzüglich eines NN mit asymmetrischen Kostenfunktionen zur Berechnung von Bestellmengen verwendet, so dass die Vergleichsstudie insgesamt 147.789 Einzelvorhersagen hinsichtlich mehrerer Kategorien von Zeitreihen und Verfahren über mehrere Beurteilungsmaße analysiert.

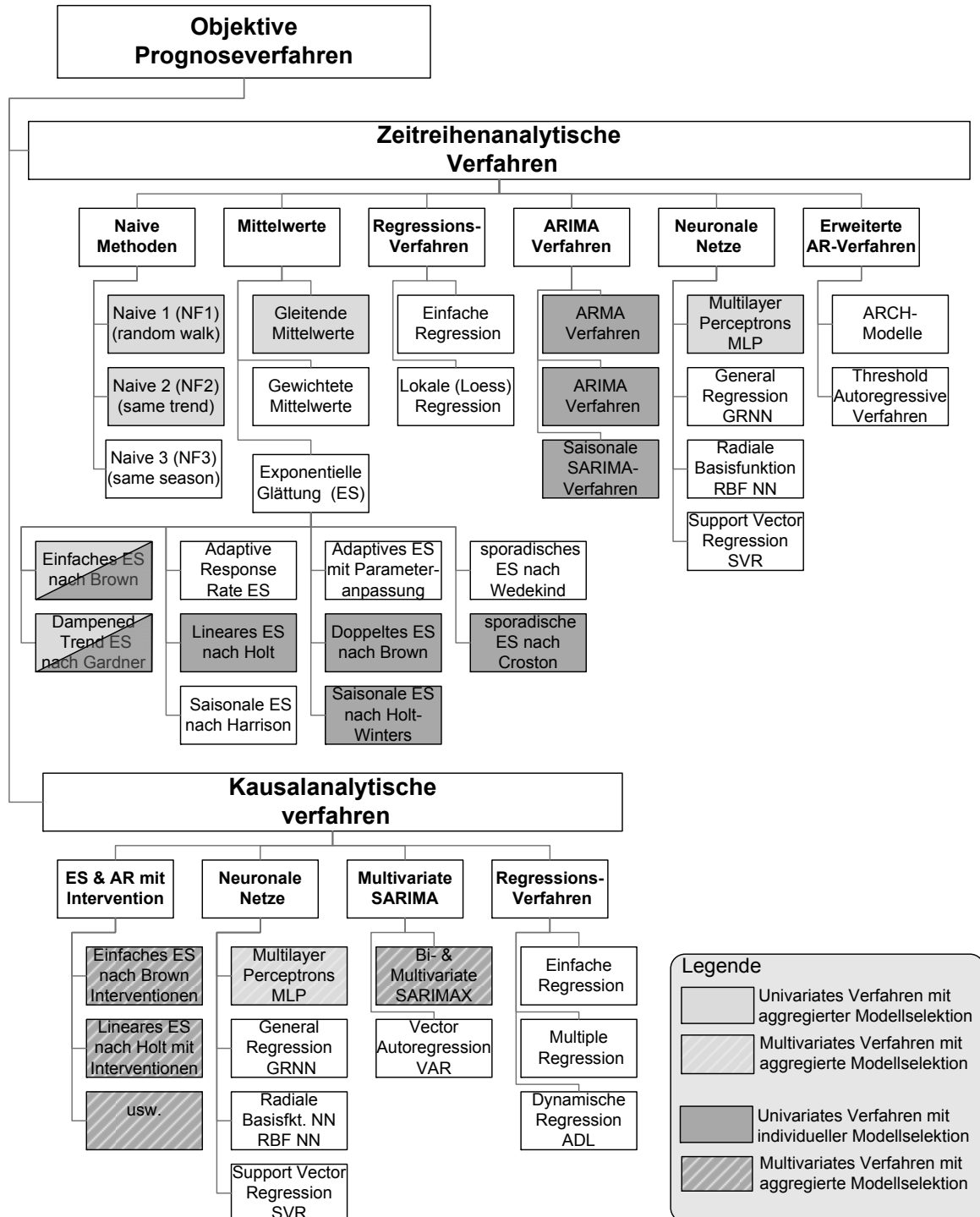


Abbildung 6.13: Übersicht der Prognoseverfahren in der Vergleichsstudie

6.2.3.2 Berechnung der statistischen Prognoseverfahren

Die statistischen Prognoseverfahren umfassen univariate und multivariate Verfahren sowie univariate und multivariate Expertensysteme. Alle Verfahren wurden bereits hinreichend in Kapitel 3 diskutiert, so dass hier lediglich die Diskussion ihrer konkreten Parametrisierung im Experiment erfolgt, um eine mögliche Replikation der Experimente sicherzustellen.

Die univariaten Verfahren mit aggregierter Modellselektion und konstanten Parametern umfassen die Verfahren Naive 1 (NF1), Naive 2 (NF2), gleitender Mittelwert (MA) und einfache Exponentielle Glättung (S.ES). Die Naiven Verfahren NF1 und NF2 dienen als Benchmark zur Bestimmung einer unteren Grenze der Prognosegenauigkeit. Darüber hinaus werden sie in der Praxis noch immer in mittelständischen Unternehmen und auch bei der Befüllung von Warenautomaten verwendet. Das Naive Verfahren (NF1) aus Formel 3.31 prognostiziert einen *random walk* unter Verwendung des letzten Beobachtungswertes. Das saisonale Naive Verfahren 2 (NF2) aus Formel 3.67 prognostiziert einen saisonalen *random walk* aus dem Beobachtungswert der vorangegangenen Saison d . NF1 benötigt keine Parameter, NF2 hingegen die Ex-ante-Bestimmung einer Saisonalität, was bei potenziell mehrfach überlagerter Saisonalität ein Auswahlproblem darstellt. NF1 ist ohne Initialisierung ab Periode t_2 anwendbar, während NF2 zunächst d Beobachtungen der vorigen Saison benötigt, um Prognosen ab dem Zeitpunkt t_{d+1} zu erstellen. Zur Initialisierung des Verfahrens NF2 werden die Werte t_2, \dots, t_d durch das Verfahren NF1 vorhergesagt.¹ Der ungewichtete gleitende Mittelwert dient als empirischer Benchmark, da er in der Bestandsplanung an Warenautomaten weit verbreitet ist und Ähnlichkeit mit dem tatsächlich eingesetzten Prognoseverfahren der Fallstudie besitzt.² Der gleitende Mittelwert MA wird nach Formel 3.33 mit der Länge $n = 3$ berechnet. Da die Berechnung eines Prognosewertes erst ab dem Beobachtungswert y_{n+1} , in diesem Fall also ab \hat{y}_4 möglich ist, kann für die ersten n Perioden keine Prognose erfolgen. Um dennoch Prognosewerte ausgeben zu können, wird das Verfahren zu Beginn der Zeitreihe durch Mittelwerte der Länge $n = \{1; 2\}$ gebildet. Die einfache Exponentielle Glättung (S.ES) aus Formel 3.34 berechnet einen exponentiell gewichteten gleitenden Mittelwert mit einem konstanten Glättungsparameter von $\alpha = 0.3$.³ Die Initialisierung des Verfahrens erfolgt durch Gleichsetzen des ersten Prognosewertes mit dem Beobachtungswert, $\hat{y}_1 = y_1$. Die Anwendung der S.ES mit konstantem Glättungsparameter über alle Zeitreihen stellt ein in der Handelspraxis weit verbreitetes Verfahren dar, das von nahezu jeder Software zur Absatzprognose unterstützt wird. Die Berechnung der Verfahren NF1, NF2, MW, S.ES erfolgt je Zeitreihe in MS Excel.⁴

Durch jüngere Vergleichsstudien univariater Prognoseverfahren [Grambsch und Stahel 1990; Liao und Fildes 2005] sind auch Verfahren der degressiven Trendextrapolation als Benchmark etabliert. Die univariate Exponentielle Glättung mit robustem oder gedämpftem Trend (DT.ES) aus Formel 3.62 wird ebenfalls ohne Modellselektion auf alle Zeitreihen angewen-

¹ Da die einzelnen Prognoseverfahren unterschiedliche Zeiträume zur Initialisierung benötigen, diese jedoch einen inhärenten Teil des Modellierungsprozesses der einzelnen Verfahren darstellen, kann es zu geringfügigen Verzerrungen der Fehlerwerte zu Beginn der Trainingsdaten kommen. Dies erscheint jedoch wenig bedeutsam, da die endgültige Verfahrensgüte auf den Testdaten erhoben wird, die nicht von dem Initialisierungszeitraum beeinflusst werden.

² Hierbei handelt es sich um ein proprietäres Prognoseverfahren aus eigener Entwicklung.

³ Der Parameterwert $\alpha=0,3$ stellt eine verbreitete heuristische Empfehlung in der Praxis und der Literatur dar und dient in diversen Softwaresystemen als Standardeinstellung, unter anderem SAP APO-DP.

⁴ Es werden 51 Excel-Dateien je Zeitreihe erstellt, in denen die univariaten Prognoseverfahren berechnet und zudem alle Prognosewerte der extern berechneten Verfahren aus ForecastPro und den NN gesammelt werden. Ferner dienen sie zur Berechnung der Standardabweichung der Prognosefehler und somit zur Berechnung der Sicherheitsbestände und zur Beurteilung der Bestellmengen. Aus Prognosewerten und Bestellmengen werden abschließend die Prognosefehler und Kostenmaße je Zeitreihe und Prognoseverfahren zur Auswertung der Experimente berechnet. Diese werden anschließend in weiteren Excel-Dateien zunächst je Absatzstelle und dann über alle Absatzstellen aggregiert und ausgewertet.

det, jedoch unter Bestimmung von optimalen Verfahrensparametern α , β und θ für jede Zeitreihe. Die Parameter der DT.ES werden durch das Software-Expertensystem ForecastPro auf den Trainings- und Validierungsdaten ermittelt, welches auch die Prognosewerte berechnet.¹

Die univariaten Verfahren mit individueller Modellselektion unterscheiden sich von den Ansätzen der aggregierten Modellselektion durch die Anwendung von heuristischen Methodologien zur Verfahrensselektion, welche softwaregestützt in einem Expertensystem abgebildet ist.² Das Expertensystem ForecastPro gilt seit seiner überragenden Leistung bei der M3-Competition [Makridakis und Hibon 2000] als Benchmark für automatisierte Softwaresysteme zur Zeitreihenprognose. Es kommt in drei Varianten zur Anwendung: zur Modellselektion aus der Verfahrensklasse der Exponentiellen Glättung (EXP.ES) und der Klasse der saisonalen ARIMA-Prozesse (EXP.ARIMA) sowie zur Modellselektion über beide Verfahrensklassen hinweg (EXP.ES&ARIMA), was die eigentliche Expertenselektion der Software ForecastPro darstellt. EXP.ES wählt gemäß den Zeitreihencharakteristika für stationäre Zeitreihen ein Verfahren der einfachen Exponentiellen Glättung nach BROWN (S.ES) oder für sporadische Zeitreihen nach CROSTON, für Zeitreihen mit Trend die Lineare Exponentielle Glättung nach HOLT oder die doppelte Glättung nach BROWN und für saisonale Muster die Exponentielle Glättung nach HOLT-WINTERS und parametrisiert das Verfahren für die Zeitreihe und den Zeitreihenursprung.³ Die Konfiguration für saisonale ARIMA-Prozesse (EXP.ARIMA) selektiert und parametrisiert passende ARIMA-Modelle für jede Zeitreihe. Die Konfiguration zur Auswahl eines Verfahrens über alle Verfahrensklassen (EXP.ES&ARIMA) selektiert und parametrisiert ein Verfahren aus den Klassen der Exponentiellen Glättung und ARIMA mit den geringsten Fehlern auf der Kalibrationsstichprobe.

Multivariate Verfahren (mv) berücksichtigen zusätzlich zu den zeitlich verzögerten Realisationen der abhängigen Variablen auch Zeitreihen von unabhängigen Variablen und erlauben so die Modellierung absatzstellenspezifischer, exogener Ereignisse (siehe Abschnitt 3.2.3). Sie werden ebenfalls durch das Expertensystem ForecastPro unter Verwendung der Verfahrensklasse der Exponentiellen Glättung (mv.EXP.ES) als Basismodell mit individueller Modellselektion berechnet.⁴ Die Verfahren berücksichtigen zwei unabhängige Variablen x_1 und x_2 als Zeitreihen von binären Dummy-Variablen, mit $x_{1,t}, x_{2,t} = \{0; 1\}$, für alle Beobachtungswerte t von y_t . Sie zeigen das Auftreten von zwei externen Effekten (engl. *events*) mit unterschiedlicher Intensität an: $x_{1,t} = 1$ identifiziert eine Periode t mit vollständiger Schließung der Absatzstelle und somit einen möglichen vollständigen Absatzverlust $y_t = 0$, zum Beispiel durch Urlaub, Umbau oder Krankheit; $x_{2,t} = 1$ identifiziert eine Periode t mit zeitweiliger Schließung der Absatzstelle und entsprechend stochastisch verringertem Absatz.

¹ Die Berechnung erfolgt in ForecastPro 4.3 Extended Edition – unter Deaktivierung des Expertenmodus – durch manuell erstellte Skriptdateien zur Steuerung der Prognosesoftware, so dass nur die optimalen Parameter des Verfahrens, nicht aber das Verfahren selbst für die Zeitreihe bestimmt wird. Die Verfahrensparameter werden für jeden Zeitreihenursprung neu bestimmt, so dass alle Prognosen unterschiedliche Verfahrensparameter aufweisen können. Die Prognosewerte auf der Kalibrationsstichprobe sowie jeweils die Prognose für einen Prognoseursprung werden aus ForecastPro in eine Excel-Datei exportiert. Somit werden 21 Dateien mit Prognosewerten je Zeitreihe exportiert, insgesamt 1071 Dateien je Prognoseverfahren.

² Das Expertensystem ForecastPro identifiziert zunächst durch statistische Tests die Struktur der Zeitreihe. Anschließend wird durch interne heuristische Expertenroutinen jeweils aus dem Vorrat möglicher Verfahrenskandidaten das am besten geeignete Verfahren für die Zeitreihe ausgewählt und anschließend durch Parametrisierung bestmöglich an die Struktur der Zeitreihe angepasst. Die grundlegende Funktionsweise der Modellidentifikation, -selektion und -parametrisierung beschreibt GOODRICH [2000].

³ Wie für die Prognose mit DT.ES werden Skriptdateien zur Steuerung der Software für die drei Verfahren erstellt, $t+1$ Prognosen für 21 Prognoseursprünge berechnet, in insgesamt 3213 Excel-Dateien exportiert und die Ergebnisse jeweils in den Excel-Dateien der Zeitreihen zum Verfahrensvergleich gesammelt.

⁴ ForecastPro erlaubt nur die Berechnung von Verfahren der ES mit Event- oder Interventions-Modellen, nicht aber von ARX- oder SARIMAX-Modellen. Somit wird auf deren Darstellung verzichtet.

Die Parameterfindung und Prognose für mv.EXP.ES erfolgt entsprechend automatisiert und analog zu DT.ES.¹

6.2.3.3 Berechnung der Neuronalen Netze zur Prognose

Korrespondierend zu den Kategorien der statistischen Verfahren werden sowohl univariate als auch multivariate Architekturen von MLPs zur Prognose konfiguriert.

Bislang existiert für MLPs keine allgemein akzeptierte Methodologie zur Bestimmung von zeitreihenspezifischen Netzwerkarchitekturen. Daher würde die Ermittlung einer geeigneten Netzwerkarchitektur je Zeitreihe unter Variation aller relevanten Parameter eines MLP $[[N_I-N_{H1}(F_I, F_A, F_O)-N_{H2}(F_I, F_A, F_O)-\dots-N_O(F_I, F_A, F_O)], W, K, [[I^N, I^P], G, P_{T,L}, B, S, O]]$ nach Kapitel 4 umfangreiche Experimente im Vorfeld der eigentlichen Untersuchung von asymmetrischen Kostenfunktionen für NN erfordern. Zudem haben die Voruntersuchungen in Abschnitt 4.4 und 5.4 hinreichend belegt, dass ein MLP mit konstanter Architektur allein durch Anpassung der Verbindungsgewichte unterschiedliche Zeitreihenmuster zu erlernen und im Sinne der Ex-ante- Prognose in die Zukunft zu generalisieren in der Lage ist. Zur Komplexitätsreduktion wird daher dem Ansatz von HILL et al. [1996] gefolgt, die eine geeignete Netzwerkarchitektur anhand einer vorab gezogenen Stichprobe repräsentativer Zeitreihen ermitteln. Die Stichprobe der Zeitreihen wird nachfolgend nicht mehr zur Evaluation im Verfahrensvergleich herangezogen, um einen Ex-ante-Verfahrensvergleich zu gewährleisten. Hierzu werden die sieben Zeitreihen der Absatzstelle IN-3 verwendet. Die Bestimmung der Netzwerkarchitektur erfordert die Berechnung unterschiedlicher Architekturvarianten von MLPs auf den Zeitreihen von IN-3 und die systematische Auswahl einer einzelnen, robusten Architektur. Aufgrund der Bedeutung und des Umfangs dieser Voruntersuchung zur Festlegung einer einheitlichen MLP-Architektur auf einer Stichprobe der Zeitreihen wird der Versuchsaufbau nachfolgend kurz skizziert.

Die Voruntersuchung ermittelt eine robuste Netzwerkarchitektur durch Experimente der Eingabeneuronen N_I , der Anzahl der versteckten Schichten H , der Anzahl ihrer Neuronen N_H und ihrer Aktivierungsfunktionen (F_I, F_A, F_O) . Der Experimentaufbau untersucht unterschiedliche Inputvektoren N_I mit 20 Lag-Strukturen zunehmender Länge $N_I = (\{y_t\}, \{y_t, y_{t-1}\}, \{y_t, y_{t-1}, y_{t-2}\}, \{y_t, y_{t-1}, \dots, y_{t-3}\}, \dots, \{y_t, y_{t-1}, \dots, y_{t-19}\})$ mit 1 bis 20 korrespondierenden Eingabeneuronen. Darüber hinaus werden unterschiedliche kausalanalytische Eingabe-Strukturen für x_1 und x_2 untersucht. Zunächst werden für x_1, x_2 20 Eingabevektoren mit identischen Lag-Strukturen zu y von $y_t, y_{t-1}, \dots, y_{t-19}$ gebildet. Darüber hinaus werden 20 Varianten der jeweiligen Lag-Struktur von y mit den Erklärungsvariablen in x_1 und x_2 nur für t und $t+1$ gebildet, mit $x_{1,t}, x_{1,t+1}, x_{2,t}, x_{2,t+1}$. Somit werden insgesamt 60 Eingabestrukturen in 3 Untergruppen untersucht. Ferner werden 20 Netzwerkarchitekturen mit bis zu 2 versteckten Schichten und einer unterschiedlichen Anzahl von versteckten Neuronen untersucht. Der Experimentaufbau berücksichtigt Architekturen mit 2 bis 20 Neuronen in einer versteckten Schicht $N_{H1} = \{2, 4, 6, \dots, 20\}$ und zusätzlich zweilagige Architekturen mit einer identischen Anzahl von Neuronen in der zweiten Schicht $N_{H2} = \{2, 4, 6, \dots, 20\}$, mit $[N_I-2-1], [N_I-2-2-1], [N_I-4-1], [N_I-4-4-1], [N_I-6-1], [N_I-6-6-1]$ usw. Alle Netzwerkarchitekturen besitzen ein einziges Ausgabeneuron zur Einschnittprognose $t+1$. Zudem wird der Einfluss unterschiedlicher Aktivierungsfunktionen des Tangens Hyperbolicus (TanH) und der Logistischen Funktion (Log) untersucht, durch Kombination der Verarbeitungsfunktionen von (Sum, TanH, Id) und (Sum, Log, Id) versteckte Neuronen und

¹ Es werden Steuerungs-Skripte erstellt, 21 $t+1$ Prognosen je Prognoseursprung und Zeitreihe erstellt, je Verfahren 1071 Excel-Dateien mit den Ergebnissen exportiert und die Ergebnisse je Zeitreihe in Excel gesammelt.

(Sum, TanH, Id), (Sum, Log, Id) oder (Sum, Id, Id) für das Ausgabeneuron. Insgesamt werden so 4800 Netzwerkarchitekturen je Initialisierung und Zeitreihe berechnet.

Die Einstellungen des Lernverfahrens wurden in anderen Experimenten eingehend untersucht, dort wurden robuste Parameterkombinationen festgestellt [siehe Crone et al. 2006a; 2006b; 2006c; 2006d]. Alle Beobachtungswerte werden linear in das Intervall $[-0,6; 0,6]$ skaliert, unter Verwendung des minimalen und des maximalen Beobachtungswertes aus der Kalibrierungsstichprobe. Das Ziehen der Datenmuster auf Trainings- und Validierungsmenge erfolgt in zufälliger Reihenfolge ohne zurücklegen, bis eine Epoche vollständig ist. Die Kalibrierungsstichprobe wird in 70% Trainingsdaten und 30% Validierungsdaten aufgeteilt. Die 21 Beobachtungswerte der Testdaten entsprechen denen der anderen Verfahren und werden nicht zur Skalierung oder Datenvorverarbeitung verwendet. Jedes MLP wird 20-fach mit zufälligen, uniform verteilten Verbindungsgewichten im Intervall $[-0,6; 0,6]$ initialisiert und für 1000 Epochen mit einer Lernrate von $\eta = 0,7$ und einem Momentum von 0,4 trainiert. Die Lernrate wird nach jeder Epoche um 1% verringert, um eine Konvergenz im Training zu erzielen. Der Trainingsprozess wird im Sinne des *early stopping* abgebrochen, wenn ein gewichteter Prognosefehler aus 20% des *MAE* auf den Trainingsdaten und 80% des *MAE* auf den Validierungsdaten in 150 Epochen nicht um 0,1% im Vergleich zum letzten gespeicherten Gewichtsvektor verbessert werden konnte. Durch die 20 Initialisierungen der 4800 Netzwerkarchitekturen für jede der 7 Zeitreihen von IN-3 werden insgesamt 672.000 MLPs berechnet und hinsichtlich ihrer Prognosegüte auf Trainings-, Validierungs- und Testmenge analysiert. Die Analyse identifiziert eine robuste Netzwerkarchitektur für die univariate sowie für die multivariate Prognose und Disposition durch die schrittweise Analyse der Verteilungsfunktionen der Fehlerwerte je Architekturvariante.¹

Die Voruntersuchung spezifiziert eine einzelne, universelle Netzwerkarchitektur eines MLP.² Diese wird für alle Vorhersagen durch NN verwendet, sowohl bei den univariaten Architekturen (MLP) und den multivariaten Architekturen (mv.MLP) zur Prognose als auch bei den MLPs mit asymmetrischen Kostenfunktionen zur Disposition (mv.MLP.ACF). Alle MLPs verwenden ein Ausgabeneuron zur Vorhersage von \hat{y}_{t+1} mit der Identitätsfunktion und zwei verdeckte Schichten mit je zwei Neuronen mit der TanH-Aktivierungsfunktion. Die univariaten Architekturen der MLPs verwenden 8 Eingabeneuronen für $y_t, y_{t-1}, \dots, y_{t-7}$, mit einer Architektur von [8-2-2-1] für univariate MLPs. Die Lernparameter bleiben unverändert. Die Architekturen des MLP und des mv.MLP zur Prognose verwenden eine Zielfunktion des quadratischen Prognosefehlers $SE(e), [x^2 | 0 | x^2]$.³ Je Zeitreihe und Architektur wurde das NN mit dem niedrigsten kombinierten Fehler auf Trainings- und Validierungsmenge gemäß der Early-stopping-Vorgabe von 80%-20% ausgewählt. Die Werte der $t+1$ Prognosen auf der Trainings-, der Validierungs- und der Testmenge werden exportiert und in Excel je Zeitreihe mit den anderen Verfahren verglichen. Bei der Prognose aus 21 Zeitreihenursprüngen der Testdaten unterbleibt eine erneute Parametrisierung für jeden der Ursprünge, da sich andernfalls der Umfang der Berechnung von MLPs erheblich erhöhen würde.⁴ Insgesamt wurden 2.040 NN zur abschließenden Prognose über alle Zeitreihen berechnet.

¹ Eine detaillierte Analyse entfällt hier aus Platzgründen. Für eine genauere Beschreibung sei auf [Crone 2005] verwiesen.

² Die Architektur lautet MLP[[8-2(Sum, TanH, Id)-2(Sum, TanH, Id)-1(Sum, Id, Id)], F, FF, [[20, [-0,6; 0,6]], Hetero, Backprop[$\eta=0,7; \theta=0,4; 1000$ Epochs;cool 99%], EarlyStop[150 Epochs;0,001]], Min Error, $SE(e)$]]

³ Die intensiven Zufallsschwankungen der Zeitreihen legen die Verwendung einer Zielfunktion mit absolutem Prognosefehler $AE(e)$ nahe, $[x|0|x]$, um auf den verrauschten Daten eine geringere Sensitivität gegenüber Ausreißern zu erreichen. Um eine Vergleichbarkeit zu vorangegangenen Experimenten mit herkömmlichen Zielfunktionen der MLPs zu erlauben, wird hier jedoch darauf verzichtet.

⁴ Zur Äquivalenz mit den Expertensystemen müsste für jeden Zeitreihenursprung eine Vielzahl von MLPs trainiert und eine geeignete Architektur und Initialisierung zur Prognose des nächsten Beobachtungswertes selektiert werden. Anstelle einer Architektur mit 20 Initialisierungen müssten k MLPs trainiert, selektiert und

Die multivariaten Topologien mv.MLP sowie mv.MLP.ACF zur kausalanalytischen Prognose verwenden zusätzlich zu der oben beschriebenen Architektur die zwei Erklärungszeitreihen exogener Dummy-Variablen x_1 und x_2 , für den Prognosezeitpunkt t und den Prognosehorizont $t+I$, mit $x_{1,t}$, $x_{1,t+1}$ sowie $x_{2,t}$ und $x_{2,t+1}$. Durch die zusätzlichen Eingabevariablen erweitert sich für multivariate MLPs die Architektur auf MLP[12-2(Sum, TanH, Id)-2(Sum, TanH, Id)-1] mit 4 zusätzlichen Eingabeneuronen und ansonsten unveränderter Architektur.

Die Experimente verwenden eine identische Architektur der MLPs zur univariaten Prognose, zur multivariaten Prognose sowie zur Disposition. Dies begünstigt den Vergleich der Ergebnisse, da alle Unterschiede der Prognose- und Dispositionsgüte allein auf Abweichungen der Eingabedaten (univariat oder multivariat) oder Abweichungen der Zielfunktion (symmetrisch oder asymmetrisch) zurückgeführt werden können.¹ Tabelle 6.12 ordnet die univariaten und multivariaten MLPs den statistischen Verfahren in dem Versuchsaufbau in Tabelle 6.11 zu.²

Tabelle 6.12: Übersicht der Verfahren zur Absatzprognose in der Vergleichsstudie

	Univariate Verfahren		Multivariate Verfahren	
	Fixe Parameter	Optimierte Parameter	Fixe Parameter	Optimierte Parameter
Aggregierte Modellselektion	- Naive1 (NF1)	- Dampened Trend Smoothing (DT.ES)		- NN zur Prognose mit Events (mv.MLP)
	- Naive2 (NF2)			
Individuelle Modellselektion	- Mittelwert (MA)	- NN zur Prognose (MLP)		- Expertensystem für Exp. Glättung mit Events (mv.EXP.ES)
	- Exp.Glättung mit fixiertem Parameter (S.ES)			
		- Expertensystem für Exp.Glättung (EXP.ES)		
		- Expertensystem für ARIMA (EXP.ARIMA)		
		- Expertensystem (EXP.ES&ARIMA)		

Da die Architektur und somit die funktionale Form der MLP über alle Zeitreihen unverändert bleibt, erfolgt keine Modellselektion im eigentlichen Sinne, sondern lediglich eine Veränderung der Parametrisierung der Verbindungsgewichte. Dies erlaubt eine Interpretation der Verfahren MLP und mv.MLP als univariate bzw. multivariate Prognoseverfahren mit aggregierter Modellselektion und optimierten Parametern.³

6.2.4 Verfahren zur Bestellmengenrechnung

6.2.4.1 Bestimmung des kostenoptimalen Servicelevels

Die Berechnung effizienter Bestellmengen zur Warendisposition erfordert zunächst eine Bestimmung kostenminimaler Servicegrade durch die Erhebung der Entscheidungskosten aus Überbevorratung c_o und Unterbevorratung c_u (siehe Abschnitt 2.3).

angewandt werden, wobei k die Anzahl möglicher MLP-Architekturen darstellt. Dies würde die Berechnung von mehreren Milliarden MLPs erfordern, was in Anbetracht des Ressourceneinsatzes und des zu erwartenden nur geringen Anstiegs der Prognosegenauigkeit unterbleibt.

¹ Die MLPs wurden unter Verwendung eines prototypischen Simulators zur Prognose mit NN berechnet, der unter der Leitung des Verfassers am Institut für Wirtschaftsinformatik der Universität Hamburg entwickelt und bereits zur Berechnung der Experimente in Abschnitt 5.4 eingesetzt wurde. Die Rechenzeit betrug 65 Minuten auf einem Intel-PC mit 3,6-GHz-Prozessor, 4 GB Hauptspeicher und 1 TB Festplattenspeicher.

² Eine Beurteilung multivariater Verfahren mit fixierten Parametern entfällt, da die Ex-ante-Bestimmung des konstanten Einflusses einer exogenen Variablen auf die abhängige Variable dem Konzept multivariater Modellierung widerspricht.

³ Würde hingegen ein MLP individuell für jede Zeitreihe trainiert und anhand statistischer oder heuristischer Regeln ausgewählt, läge wiederum ein methodologisches Vorgehen im Sinne der Expertensysteme vor, und die MLPs müssten als Verfahren mit individueller Modellselektion betrachtet werden. Die Entwicklung einer Methodologie zur Bestimmung der Architekturparameter eines MLP erfordert die Untersuchung in zukünftigen Vergleichsstudien und eine entsprechende Einordnung der Erkenntnisse.

Die Entscheidungskosten und ihre Struktur sind durch die empirische Anwendung der Disposition an Warenautomaten bestimmt. Die Kosten der Überbevorratung entstehen durch die Lagerung jeder überschüssigen Verpackungseinheit für eine weitere Planungsperiode. Sie werden unter Vernachlässigung von Problemen der Mindesthaltbarkeit unter der Annahme einer unendlichen Lagerfähigkeit von Zigaretten durch die Kosten des Kapitaldienstes quantifiziert. Dieser kann direkt aus den variablen Stückkosten pro Verkaufseinheit und dem unternehmensspezifischen, kalkulatorischen Zinssatz je SKU und Planungsperiode berechnet werden. In diesem Fall errechnen sich die Kosten der Überbevorratung $c_o = \text{€}0,0036$.

Die Kosten der Unterbevorratung c_u werden durch die Opportunitätskosten des entgangenen Deckungsbeitrags und des Goodwill-Verlustes der Kunden bestimmt. Zur Quantifizierung des Deckungsbeitrags je SKU wird die direkte Produktrentabilität (DPR) verwendet, die als lineare Approximation der artikelbezogenen Deckungsbeiträge je Verkaufseinheit auf Vollkostenbasis im Handel etabliert und verbreitet ist [Stelling und Göllnitz 2000].¹ Grundsätzlich ist an Warenautomaten von einem Verlustfall auszugehen, da unbefriedigter Bedarf nicht zu einem späteren Zeitpunkt erfüllt wird und so ein Verlust der Nachfrage auftritt. Jedoch führt ein Unterbestand in Warenautomaten mit einem Sortiment mehrerer, substituierbarer Produkte nicht zwangsläufig zu einem verlorenen Verkauf. Je nach Kaufprozess substituieren Kunden ein nicht vorrätiges durch ein konkurrierendes Produkt an derselben Absatzstelle. Nach ersten Erhebungen wird der Anteil der Substitution durch konkurrierende Produkte (engl. *switching*) an Zigarettenautomaten auf 66% geschätzt [Berger 2000]. Somit wird angenommen, dass nur jeder dritte Unterbestand tatsächlich einen Umsatzverlust in Höhe der DPR verursacht. Unter Berücksichtigung der Substitutionsrate von einer Ware auf eine andere wird der wahrscheinliche Verlust durch Unterbevorratung unter Vernachlässigung mittel- bis langfristiger Effekte des Goodwillverlustes erhoben.² Die Kosten für Unterbevorratung werden aus der DPR mit € 0,30 ermittelt.³ Aufgrund einer angenommenen Substitutionsrate bei Unterbevorratung von 66% führt nur jeder dritte Leerverkauf zu einem tatsächlichen Verlust des Umsatzes, so dass $c_u = \text{€}0,10$ als Kosten der Unterbevorratung ermittelt werden.

Die asymmetrischen, linearen Kosten der Über- bzw. Unterbevorratung bestimmen die Berechnung des Servicegrades gemäß (2.24):

¹ Während die Kosten- und Erlösstruktur von Über- und Unterbevorratungen in Produktionsbetrieben selbst durch Ansätze der Prozesskostenrechnung nur selten valide und reliabel zu ermitteln sind, gestaltet sich die Ermittlung in Handelsbetrieben unter Fortfall der Produktionsfunktionalität einfacher. Die prozessorientierte Bestimmung der Produkteinzelkosten und des Stückdeckungsbeitrages wird insbesondere durch die Verbreitung des Konzeptes der direkten Produktrentabilität in Handelsunternehmungen vereinfacht [Hallier 1995; Pichler 1989], auch wenn ihre Anwendung nicht ohne Probleme bleibt [Stelling und Göllnitz 2000]. Dennoch erfordert auch die Ermittlung der DPR einer zusätzlich verkauften Einheit entsprechende Annahmen. Auch ist anzumerken, dass die DPR trotz irreführender Bezeichnung kein Maß für die Rentabilität darstellt, sondern die Differenz aus Erlös und Kosten im Sinne eines direkten Produkt-Gewinns ermittelt. Alternativ können die Kosten auch auf Grundlage des relevanten Stückdeckungsbeitrages geschätzt werden.

² Der Goodwill-Verlust durch den Konsumenten ist aufgrund mangelnder Marktstudien über Kaufverhalten und Kundenzufriedenheit an Warenautomaten weitgehend unbekannt. Eine verlässliche Erhebung darüber erscheint insbesondere hinsichtlich der ungeklärten Auswirkungen der Substitution bei unterschiedlichen Kaufhandlungen, z. B. bei Impulskäufen, sowie hinsichtlich der Effekte eines langfristigen In- und Outswitching aus Markenfamilien problematisch, aber aufgrund des direkten Einflusses auf die Bestandshöhen dringend notwendig. Der Goodwill-Verlust kann gegenwärtig jedoch nicht quantifiziert werden und wird daher unter der Annahme kurzfristiger Entscheidungsprobleme vernachlässigt.

³ Aus Gründen der Geheimhaltung können die tatsächlichen Werte der Variablen nicht veröffentlicht werden. Andererseits erfordert ein valider Vergleich die Verwendung realistischer Entscheidungsparameter. Da zur Lösung des Entscheidungsproblems lediglich das Kostenverhältnis zwischen Unter- und Überbeständen relevant ist, werden die Parameter proportional skaliert, so dass die Ergebnisse der Studie valide sind.

$$F_D(S^*) \geq \frac{c_u}{c_o + c_u} = \frac{0,10}{0,0036 + 0,10} = 0,9662 = 96,62\% \quad , \quad (6.1)$$

Der Servicegrad bestimmt das Ziel der Disposition an Warenautomaten gemäß der empirischen Kostenstruktur durch einen Servicegrad von 96,62%.¹

6.2.4.2 Bestellmengenrechnung für statistische Prognoseverfahren

Unter Kenntnis der Entscheidungsparameter der Servicegrade können die Warenbestände der konkurrierenden Prognoseverfahren aus ihren Prognosewerten und der Verteilung ihrer Prognoseresiduen berechnet werden. Die Bestellmengenrechnung erfolgt unterschiedlich für die Prognoseverfahren mit aggregierter Modellselektion und die Expertensysteme mit individueller Modellselektion. Alle Ansätze unterstellen eine einperiodige Disposition im Verlustfall mit einer unendlichen Haltbarkeit der Güter.

Die fünf Prognoseverfahren mit aggregierter Modellselektion NF1, NF2, MA, S.ES und DT.ES berechnen einperiodige Sicherheitsbestände und Bestellmengen auf konventionelle, sequentielle Weise unter Verwendung des Newsboy-Modells aus Abschnitt 2.3.3.1. Jedes Prognoseverfahren sagt für jeden Zeitpunkt t einer Zeitreihe i den erwarteten Absatz $\hat{y}_{i,t}$ voraus, wodurch eine Zeitreihe von Einschnitt-Prognosefehlern entsteht. Auf der Basis dieser Prognosefehler $e_{i,t}$ wird rollierend die Standardabweichung $\sigma_{y,t}$ der einfachen Prognosefehler bis zum Prognoseursprung t ermittelt. Unter der Annahme normalverteilter Prognosefehler entspricht der Servicelevel von 96,62% einer Anzahl von $z = 1,82$ Standardabweichungen zur Berechnung eines kostenminimalen Sicherheitsbestandes. Die optimale Bestellmenge S_t^* der Periode t wird berechnet als

$$S_t^* = \hat{y}_t + 1,82 \cdot \sigma_{y,t} \quad . \quad (6.2)$$

Für Prognoseverfahren ohne Modellselektion erfolgt somit keine Identifikation der tatsächlichen Verteilungsform der Prognosefehler in einem zusätzlichen Prozessschritt. Dies entspricht dem Konzept einer konstanten, unveränderten Verfahrensanwendung und ist im Einsatz einfacher Prognoseverfahren und bei der Disposition an Warenautomaten üblich.²

Die Verfahren der Expertensysteme mit individueller Modellselektion werden kostenminimale Bestellmengen durch interne Routinen der Software unter Identifikation der Verteilungsform der Prognoseresiduen berechnet.³ Sie umfassen die Identifikation der Verteilungsfunktion der Prognoseresiduen des auszuwählenden Prognoseverfahrens und die Berechnung der Sicherheitsbestände für einen vorgegeben Soll-Servicegrad. Die so berechneten Bestellmengen werden exportiert und in MS Excel analysiert. Dieses Vorgehen erlaubt -

¹ Unter der Annahme abweichender Substitutionsraten würden entsprechend höhere oder niedrigere Kosten der Unterbevorratung entstehen und so abweichende Servicegrad-Ziele bestimmt werden. Wechselt beispielsweise nur jeder zweite Konsument die Marke, so würden Kosten von $c_u = \text{€ } 0,15$ einen kostenoptimalen Servicelevel von 97,66% bestimmen. Im Fall einer 100%-Markentreue würde jede Unterbevorratung zu einem Absatzverlust führen und die Opportunitätskosten von $c_u = \text{€ } 0,30$ einen Servicelevel von 98,81% erfordern. Die Sensitivität der Modellannahme zur Substitutionsrate erscheint somit verhältnismäßig gering und ihre Annahme für die nachfolgenden Experimente akzeptabel.

² Darüber hinaus ist die vereinfachte Berechnung von Sicherheitsbeständen unter der Annahme normalverteilter Prognosefehler auch in führenden Softwareprodukten weit verbreitet, trotz offensichtlicher theoretischer Mängel. Insbesondere bei Produkten mit langsamem Artikelabsatz und -umschlag müssen Poisson- und negativ Binomialverteilte Absatzverteilungen unterstellt werden. Die Residuen eines Prognoseverfahrens können in diesem Fall nicht normalverteilt sein, auch dann nicht, wenn der Stichprobenumfang hinreichend groß ist und nach einer gängigen Vereinfachung mehr als 100 Beobachtungswerte vorliegen.

³ Die internen Algorithmen und Heuristiken zur Bestellmengenrechnung von ForecastPro sind ebenfalls in [Goodrich 2000] beschrieben. Die exportierten Sicherheitsbestände der Prognoseperioden erlauben keine Rückschlüsse auf die angewandte Verteilungsform.

wie schon im Versuchsaufbau der Prognoserechnung - einen Vergleich einfacher statistischer Prognoseverfahren sowohl mit fortgeschrittenen Verfahren der Expertensysteme als auch mit der simultanen Prognose und Warendisposition durch ein MLP mit asymmetrischen Kostenfunktionen. Der Vergleich wird so auf die Beurteilung der Entscheidungskosten aus der Warendisposition erweitert.

6.2.4.3 Simultane Bestellmengenrechnung mit Neuronalen Netzen

Zum Vergleich mit den herkömmlichen Verfahren der Bestellmengenrechnung wird ein multivariates MLP mit einer asymmetrischen Kostenfunktion (mv.MLP.ACF) trainiert, um je Zeitreihe kostenminimale Bestellmengen unmittelbar vorherzusagen. Die Beurteilung seiner Prognosegüte stellt den eigentlichen Kern der Arbeit dar.

Das mv.MLP.ACF löst die Problemstellung der Disposition als simultane Absatzprognose und Bestellmengenrechnung allein unter Verwendung einer Zielfunktion der tatsächlichen Entscheidungskosten der Disposition an Warenautomaten. Aus den tatsächlichen Kosten der Überbevorratung $c_o = \text{€ } 0,0036$ und der Unterbevorratung $c_u = \text{€ } 0,10$ wird eine asymmetrische Kostenfunktion des Typs *LINLIN* (*LLC*) hergeleitet, wie in Abschnitt 5.3.1.2 beschrieben:

$$LLC_{t+h}(e) = \begin{cases} 0,10 & |y_{t+h} - \hat{y}_{t+h}| & \text{für } y_{t+h} > \hat{y}_{t+h} \\ 0,0036 & |y_{t+h} - \hat{y}_{t+h}| & \text{für } y_{t+h} \leq \hat{y}_{t+h} \end{cases} \quad (6.3)$$

Unter Verwendung von *LLC* als Zielfunktion wird ein MLP direkt auf die Minimierung der Dispositionskosten und somit auf die Bestimmung der kosteneffizienten Bestellmenge trainiert. Diese entspricht wie bei konventioneller Berechnung dem 96,62%-Quantil der jeweiligen Verteilungsfunktion der Prognoseresiduen für jede Zeitreihe.¹ Der Trainingsprozess des mv.MLP.ACF integriert alle manuellen Schritte der Prognose, der Sicherheitsbestandsberechnung und der Bestellmengenrechnung, so dass lediglich die vorhergesagten kostenminimalen Bestellmengen ausgegeben werden.²

Die Parametrisierung des mv.MLP.ACF erfolgt identisch mit der Parametrisierung des mv.MLP durch die Architektur mv.MLP.ACF[[12-2(Sum, TanH, Id)-2(Sum, TanH, Id) -1(Sum, Id, Id)], F, FF, [[20, [-0,6; 0,6]], Hetero, Backprop[$\eta=0,7$; $\theta=0,4$; 1000 Epochs; cool 9 9%], EarlyStop[150 Epochs;0,001]], Min Error, *LLC*(*e*)]. Die einzige Abweichung zu der Architektur von mv.MLP besteht in der Verwendung der asymmetrischen Kostenfunktion *LLC*(*e*) mit $[0.0036 * x | 0 | 0.1 * x]$ aus (6.3) zum Training des MLP.³ Das so konfigurierte MLP stellt nicht länger ein Prognoseverfahren dar, sondern ein Verfahren der Disposition, und wurde somit nicht in der Systematik der Prognoseverfahren in Tabelle 6.12 dargestellt. Die Prognosekomponente entspricht einer multivariaten, aggregierten Modellselektion unter Bestimmung optimaler Parameter.

¹ Die Kostenfunktion besitzt somit nur im gegebenen Anwendungskontext Gültigkeit und kann nicht grundsätzlich zur Disposition an Warenautomaten eingesetzt werden.

² Diese werden anschließend nach Excel exportiert und dort zur Ermittlung der Dispositionsgüte verglichen.

³ Zur Beschleunigung des Lernprozesses wird eine um den Faktor 10 skalierte asymmetrische Zielfunktion von $[0.036 * x | 0 | 1 * x]$ verwendet. Durch die homogene Skalierung besitzt sie einen identischen Grad der Asymmetrie c_u/c_o und bezeichnet somit einen identischen Servicegrad und ein identisches Quantil auf der Verteilungsfunktion. Die Skalierung bewirkt durch die Berechnung größerer Fehlersignale eine Beschleunigung des Lernprozesses, welche alternativ jedoch auch durch eine Anpassung der Lernrate hätte erzielt werden können. Da dies jedoch eine Veränderung der ansonsten identischen Netzwerkarchitektur bewirkt hätte, wurde auf eine Skalierung der Zielfunktion zurückgegriffen.

6.2.5 Beurteilungsmaße

6.2.5.1 Fehlermaße der Prognosegenauigkeit

Die Beurteilung von Prognoseverfahren in Vergleichsstudien erfordert die Anwendung robuster Gütekriterien zur Bestimmung der Genauigkeit von Prognose und Disposition. Die Genauigkeit der Prognose wird anhand etablierter statistischer Fehlermaße der absoluten Fehler (AE) und absoluten prozentualen Fehlern (APE) verglichen (vgl. Abschnitt 3.3.2). Die Dispositionsgüte der Prognoseverfahren wird anhand der entstehenden Kosten aus den Bestandsentscheidungen bewertet.

Der absolute Prognosefehler $AE_{i,t+h}$ einer Zeitreihe i berechnet die absolute Differenz zwischen dem Beobachtungswert y_{t+h} und dem Prognosewert $\hat{y}_{i,t+h}$ für den Prognosehorizont $h = 1$ über alle Zeitpunkte t der Trainings-, Validierungs- und Testmenge:¹

$$AE_{t+h} = |e_{t+h}| \quad . \quad (6.4)$$

Ferner werden prozentuale Fehler berechnet. Wie in der explorativen Datenanalyse in Abschnitt 6.2.2.2 deutlich wurde, existieren für alle Zeitreihen Nullperioden ohne Absatz, $y_{i,t} = 0$. Ferner können auch die Prognosewerte einer Periode korrekt einen Absatz von Null vorhersagen, $\hat{y}_{i,t} = 0$. Da der konventionelle absolute prozentuale Fehler $APE_{i,t+h}$ (Formel 3.87) für $y_{i,t} = 0$ nicht definiert ist, erfolgt eine modifizierte Fehlerberechnung des APE in Anlehnung an den symmetrischen $MAPE$ ($SMAPE$) nach ARMSTRONG [1978]:

$$APE_{i,t+h} = \begin{cases} 0 & \text{für } y_{i,t+h} = 0; \hat{y}_{i,t+h} = 0 \\ \frac{|y_{i,t+h} - \hat{y}_{i,t+h}|}{0.5(y_{i,t+h} + \hat{y}_{i,t+h})} & \text{für } y_{i,t+h} = 0; \hat{y}_{i,t+h} \neq 0 \\ \frac{|y_{i,t+h} - \hat{y}_{i,t+h}|}{y_{i,t+h}} & \text{für } y_{i,t+h} \neq 0; \hat{y}_{i,t+h} \neq 0 \end{cases} \quad . \quad (6.5)$$

Prozentuale Fehlermaße wie der $MAPE$ werden hier trotz ihrer grundlegenden Probleme berücksichtigt, da ihre Verwendung in der Anwendungsdomäne der Warenautomaten weit verbreitet ist. Die Fehlermaße der MAE entsprechen aufgrund des linear ansteigenden Fehlerverlaufs proportional zur Höhe der Prognoseabweichung am ehesten der tatsächlichen Entwicklung der Beurteilungsgröße der Kosten in der Anwendungsdomäne. Die abschließende Bewertung der Prognosegenauigkeit erfolgt somit anhand der mittleren absoluten Fehler $mwMAE$. Medianbasierte Fehlermaße $mdMAE$ und $mdMAPE$ ergänzen die Analyse der Genauigkeit als ausreißerrobuste Kriterien, die eine Beurteilung der Konsistenz der Ergebnisse zulassen. Sie bilden zwar nicht die primäre Entscheidungsgrundlage zur Bestimmung der Prognosegüte, sollten aber die Erkenntnisse der $mwMAE$ bestätigen.

Für jedes Prognoseverfahren werden anschließend die Prognosefehler über alle Zeitpunkte t der Trainings-, Validierungs- und Testdaten einer Zeitreihe aggregiert. Durch den konstanten Prognosehorizont von $h = 1$ vereinfacht sich die Aggregation der Fehlermaße zur Summierung über mehrere Zeitreihenursprünge t_u . Da summierte Fehlermaße durch unterschiedliche Verfahrensinitialisierungen² und Zeitreihen unterschiedlicher Länge zu Verzerrungen führen

¹ Die ursprünglichen Formeln in Abschnitt 3.3.2 des AE (Formel 3.83) werden lediglich um den Index der Zeitreihe i erweitert.

² Jedes Prognoseverfahren benötigt eine unterschiedliche Anzahl von Beobachtungswerten zu seiner Initialisierung und zur Berechnung von Prognosewerten. Somit könnten summierte Fehlermaße aus einer unterschiedlichen Anzahl von Trainingsfehlern entstehen. Die Bildung mittlerer Fehlerwerte ist lediglich für

können und so einen validen Vergleich verhindern, werden mittlere Fehlermaße des MAE_i und $MAPE_i$ (vgl. Formel 3.94 und 3.98) der einzelnen Datenteilmengen je Zeitreihen gebildet. Um verallgemeinerbare Aussagen über die Stichprobe der 51 Zeitreihen ableiten zu können, werden im Anschluss die mittleren Fehler eines Prognoseverfahrens je Datenmenge über alle Zeitreihen einer Absatzstelle, einer besonderen Klasse von Zeitreihen oder der gesamten Stichprobe aggregiert. Dazu werden die Summen der MAE_i ($sumMAE$)¹, die Mittelwerte der MAE_i ($mwMAE$), die Mittelwerte der $MAPE_i$ ($mwMAPE$) sowie der Median der MAE_i ($mdMAE$) und der $MAPE_i$ ($mdMAPE$), über alle Zeitreihen gebildet.

Zur besseren Lesbarkeit der relativen Fehler wird zudem der absteigende Rang der Verfahren je Datenmenge nach allen Fehlermaßen gebildet. Dabei erhält das Verfahren mit dem niedrigsten Fehlerwert den Rang 1, das mit dem zweitniedrigsten Fehler den Rang 2 und so fort. Die ordinale Rangreihung erleichtert dem Leser die visuelle Analyse und den Vergleich der Prognosegüte zwischen Verfahren für die Fehlermaße der $sumMAE$ ($Rang\ sumMAE$), der $mwMAE$ ($Rang\ mwMAE$) und der $mdMAE$ ($Rang\ mdMAE$).

Darüber hinaus kann die Genauigkeit auch direkt anhand ordinaler Gütekriterien bewertet werden, indem der absteigende Rang der Prognoseverfahren nach den jeweiligen mittleren Fehlerwerten der MAE und $MAPE$ je Datenmenge einer Zeitreihe gebildet wird und der summierte Rang ($sumRang\ MAE$ oder $sumRang\ MAPE$) oder der mittlere Rang ($mwRang\ MAE$ oder $mwRang\ MAPE$) eines Verfahrens über alle Absatzstelle, Klassen von Zeitreihen oder die gesamte Stichprobe gebildet wird. Das Prognoseverfahren mit der niedrigsten Rang-Summe repräsentiert entsprechend das beste Verfahren nach ordinalem Verfahrenvergleich. Die rangbasierten Gütemaße erlauben eine robuste Beurteilung der relativen Prognosegüte, die anders als metrische Fehlermaße nicht durch einmalige, hohe oder niedrige Ausreißer beeinflusst werden können. Sie ergänzen so herkömmliche, robuste statistische Fehlermaße² und erlauben damit wiederum die Bildung des Rangs nach der Rangsumme $Rang\ SumRang\ MAE$.³ Entsprechend können auch die rangbasierten Gütemaße durch Summierung zur Rang-Summe $sumRang$, zur Bildung von Mittelwerten zu $mwRang$ oder Medianen zu $mdRang$ für die jeweiligen Fehlermaße kumuliert werden.

6.2.5.2 Kostenmaße der Dispositionsgüte

Die Güte der Dispositionsentscheidung kann nicht anhand statistischer Fehlermaße beurteilt werden, sondern erfolgt in der Praxis üblicherweise anhand von deskriptiven Gütemaßen der Anzahl von Fehlmengen und Überbeständen (siehe Abschnitt 2.4). Es werden für die Warenbestände jedes Prognoseverfahrens die warenlogistischen Kennziffern des α - und β -Servicegrades je Datenmenge und Zeitreihe berechnet⁴, mit den entsprechenden Summen, Mittelwerten und Rangordnungen über alle Zeitreihen, berechnet als $mwALPHA$ sowie $mwBETA$.

Ferner wird die Güte der Dispositionsentscheidung anhand der resultierenden Kosten der Warenbestände bewertet (siehe Abschnitt 5.3). Aufgrund der Kenntnis der Kosten einer

die Trainingsdaten erforderlich. Sie hat keinerlei Auswirkungen auf die Testdaten mit einer konstanten Anzahl von Beobachtungswerten, erlaubt jedoch durch die Betrachtung von Fehlern pro Ware und Prognoseperiode eine einfachere Interpretation und Hochrechnung der Genauigkeit.

¹ Eine Bildung des summierten MAPE liefert keine validen Ergebnisse und unterbleibt daher.

² Während ordinale Fehlermaße durch die Rangreihung der Verfahren eine erhöhte Transparenz bieten, erlauben sie keine Rückschlüsse auf die Distanz zweier oder mehrerer Verfahren wie metrische Fehlermaße. Da in Anwendungen der Prognose vielfach die Anzahl fehlerhaft prognostizierter Absatzeinheiten maßgeblich ist, dienen ordinale Gütekriterien als robuste Zusatzinformationen.

³ Bezeichnet beispielsweise $sumRang\ MAE=188$ die Summe der Ränge eines Prognoseverfahrens und $mwRang\ MAE=4,3$ einen mittleren Rang von 4,3 über alle Zeitreihen nach dem Fehler des MAE auf dem Testdatensatz, so bezeichnet $Rang\ sumRang\ MAE=2$, dass dies das zweitbeste Prognoseverfahren nach diesem Fehlermaß ist.

⁴ Siehe Formel 2.11 und 2.12 in Abschnitt 2.3.1 bzw. Abschnitt 2.4.

Überbevorratung c_o und einer Unterbevorratung c_u aus Abschnitt 6.2.4.1 erfolgt eine exakte Bestimmung der Kostenfunktion; auftretende Überbestände und Unterbestände werden somit anhand der tatsächlich entstehenden Kosten verglichen. Die Fehlerkosten werden durch eine lineare asymmetrische Kostenfunktion LLC für den Zeitpunkt $t+h$ beschrieben:

$$LLC_{i,t+h}(e) = \begin{cases} 0,10 & |y_{i,t+h} - \hat{y}_{i,t+h}| & \text{für } y_{i,t+h} > \hat{y}_{i,t+h} \\ 0,0036 & |y_{i,t+h} - \hat{y}_{i,t+h}| & \text{für } y_{i,t+h} \leq \hat{y}_{i,t+h} \end{cases} \quad (6.6)$$

In der nachfolgenden Studie wird von einheitlichen Entscheidungskosten über alle Absatzstellen und Waren ausgegangen, so dass der Index i des Gütekriteriums entfällt. Die empirische Kostenfunktion zur Ex-post-Bewertung der Dispositionsgüte entspricht den erhobenen empirischen Kosten der Entscheidungssituation und somit auch einer asymmetrischen Kostenfunktion, die im Training von NN verwendet wird. Somit erfolgt eine einheitliche Parametrisierung und Bewertung mit einheitlichen Entscheidungskosten.

Die Aggregation der Kosten erfolgt korrespondierend zu den statistischen und rangbasierten Fehlermaßen zunächst durch Bildung der Summen der Fehlerkosten $SLLC$ über alle Zeitpunkte einer Datenmenge. Wie bei statistischen Fehlermaßen entstehen durch die Summe über eine unterschiedliche Anzahl von Beobachtungswerten mögliche Verzerrungen der Fehlerkosten durch die Berechnung der Prognosewerte und der Sicherheitsbestände.¹ Somit werden die Mittelwerte über alle Zeitpunkte einer Datenmenge berechnet ($MLLC$), einschließlich der entsprechenden Rangbildung ($Rang\ MLLC$ und $Rang\ SLLC$). Abschließend erfolgt auch hier die Aggregation über alle Zeitreihen zur Bildung der Summen ($sumSLLC$ und $sumMLLC$), der Mittelwerte ($mwSLLC$ und $mwMLLC$) und der Mediane ($mdSLLC$ und $mdMLLC$) sowie der Rangfolge der Verfahren auf Basis der Fehlerkosten.

Alle Fehlermaße und Kostengrößen werden über alle Teildatenmengen durch das Tupel ((Trainingsfehler / Validierungsfehler) Testfehler) dargestellt, um eine Überprüfung der Approximationsfähigkeit der Verfahren auf den Trainings- und Validierungsfehlern sowie der Generalisierungsleistung auf den Testfehlern zuzulassen. Die Verwendung der dargestellten Vielzahl von Gütemaßen der Prognosefehler und der Dispositionskosten gestattet vielfältige Betrachtungsweisen in der Analyse der Experimente, um die Validität und Reliabilität des Experimentdesigns und der gefundenen Ergebnisse sicherzustellen.²

6.3 Ergebnisse der Vergleichsstudie

6.3.1 Ergebnisse der Bedarfsprognose

6.3.1.1 Gesamtergebnisse

Nachfolgend werden die Ergebnisse des Verfahrensvergleichs präsentiert. Zunächst erfolgt die Analyse der Genauigkeit der Verfahren zur Bedarfsprognose. Sie wird unabhängig von ihrer Güte zur Bestandsrechnung dargestellt, um mögliche Ursachen für die unterschiedliche Performanz der einzelnen Verfahren in der Warendisposition aufzuzeigen.³

¹ Die Berechnung der Standardabweichung zur konventionellen Berechnung von Sicherheitsbeständen erfordert mindestens zwei Beobachtungswerte, während das MLP zur unmittelbaren Bestandsberechnung ab dem ersten Prognosewert eine Bestellmenge ausgeben kann. Diese Verzerrung der summierten Kostengrößen ist auf den Trainingsdaten, nicht aber den Testdaten zu berücksichtigen. Sie werden dennoch dargestellt, um einen Eindruck von der kumulierten Höhe der Entscheidungskosten auf den Testdaten zu vermitteln.

² Die Fehlerberechnung aller Experimente erfolgte einheitlich in Microsoft Excel, Version 2003 SP2 unter hierarchischer Aggregation der einzelnen Werte entsprechend der dargestellten Vorgehensweise.

³ Die dargestellte Prognosegüte ist für die Beurteilung der Eignung eines Verfahrens zur Warendisposition irrelevant; hierzu wird ausschließlich die Güte zur Bestellmengenrechnung in Abschnitt 0 herangezogen.

Die Studie zur Prognosegüte vergleicht zwei Ansätze von NN, nämlich ein univariates MLP und ein multivariates mv.MLP, mit neun statistischen Prognoseverfahren in drei Gruppen: mit univariaten Verfahren ohne Modellselektion NF1, NF2, MA, S.ES und DT.ES, mit univariaten Expertensystemen mit Modellselektion EXP.ES, EXP.ARIMA, EXP.ES&ARIMA und mit multivariaten Verfahren mit Modellselektion mv.EXP.ES.¹ Die Ergebnisse werden auf 44 Zeitreihen der vier Absatzstellen AU-1, AU-2, IN-1 und IN-2 anhand der in Abschnitt 6.2.5 dargestellten Fehlermaße des *MAE*, *MAPE* und des relativen Rangs des Verfahrens je Zeitreihe beurteilt. Abbildung 6.14 zeigt als Auszug aller Ergebnisse der Tabelle 6.14 eine Übersicht der Güte aller Prognoseverfahren auf der Testmenge über alle Zeitreihen und Zeitreihenursprünge für alle Prognoseverfahren, absteigend sortiert nach *mwMAE*.

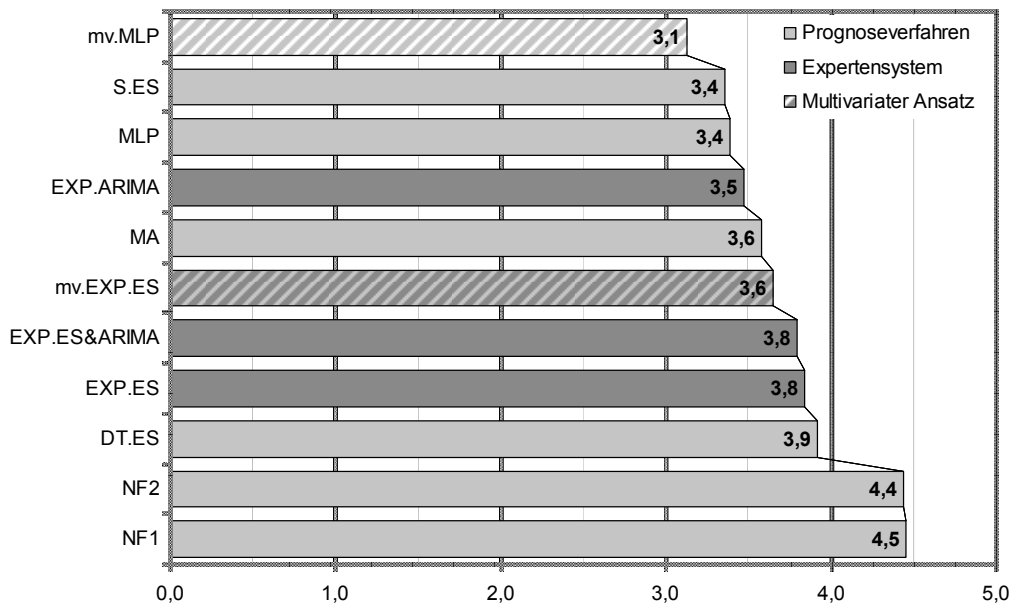


Abbildung 6.14: Prognosefehler des *mwMAE* aller Verfahren auf den Testdaten über alle Zeitreihen²

Das multivariate NN des mv.MLP ist mit einem mittleren *MAE* von 3,13 das Verfahren mit dem geringsten Prognosefehler und somit der höchsten Prognosegenauigkeit. Die einfache Exponentielle Glättung und das univariate MLP zeigen die zweit- und drittniedrigsten Fehlerwerte. Somit übertreffen die NN acht von neun konkurrierenden statistischen Verfahren auf der Stichprobe der Zeitreihen. Die Ergebnisse der Prognosegenauigkeit belegen zunächst die hohe Prognosegüte sowohl von multivariaten als auch von univariaten Ansätzen der NN im Vergleich zu den etablierten statistischen Benchmark-Verfahren.

Der Vergleich der Verfahrensgenauigkeit erfordert die Feststellung statistisch signifikanter Unterschiede der mittleren Prognosefehler der Verfahren. Aufgrund des Aufbaus der Vergleichsstudie kann hierzu nicht wie in Abschnitt 5.4 eine multifaktorielle ANOVA verwendet

¹ Die unter Abschnitt 6.2.3 dargestellten Verfahren umfassen Naive 1 (NF1), Naive 2 (NF2), gleitenden Mittelwert (MA), einfache Exponentielle Glättung (S.ES), Exponentielle Glättung des gedämpften Trends (DT.ES), ein Expertensystem für Exponentielle Glättung (EXP.ES), ein Expertensystem für saisonale Autoregressive Moving-Average-Prozesse (EXP.ARIMA) sowie ein Expertensystem für Exponentielle Glättung und SARIMA (EXP.ES&ARIMA). Die multivariaten Verfahren umfassen ein Expertensystem für Exponentielle Glättung (mv.EXP.ES). Die NN umfassen univariate Multilayer Perceptrons (MLP) und multivariate NN (mv.MLP) unter Berücksichtigung der Zeitreihe der exogenen Dummy-Variablen als kausalanalytische Prognoseverfahren, jeweils unter Verwendung einer quadratischen Zielfunktion als zeitreihenanalytisches Prognoseverfahren trainiert.

² Die Fehlerbalken folgen der farblichen Unterscheidung der Verfahren nach Abbildung 6.13. Auf allen Zeitreihen angewendete Prognoseverfahren mit aggregierter Modellselektion werden in heller Schattierung dargestellt, Verfahren der Expertensysteme mit individueller Modellselektion in dunkler Schattierung. Multivariate Ausprägungen der Ansätze mit aggregierter oder individueller Modellselektion werden schraffiert dargestellt.

werden.¹ Eine Untersuchung der Histogramme der *mwMAE* zeigt nicht-normalverteilte Fehlerwerte, was durch die Darstellung der Box-Whisker-Diagramme in Abbildung 6.15 sowie eine Reihe von Kolmogorov-Smirnov-Tests auf die Normalverteilung für alle Fehlerwerte bestätigt wird.²

Somit kann die statistische Signifikanz der mittleren Verfahrensfehler ebenfalls nicht anhand von paarweisen t-Tests ermittelt werden. Stattdessen wird eine Reihe paarweiser Wilcoxon-Vorzeichen-Rang-Tests durchgeführt, die auf Signifikanz der Fehlerwerte der Verfahren anhand ihrer ordinalen Rangordnung testen und deren Ergebnisse auf Grundlage der Prognosefehler nach dem *mwMAE* paarweise für alle Prognoseverfahren in Tabelle 6.13 dargestellt sind.³

Tabelle 6.13: Wilcoxon-Vorzeichen-Rang-Tests auf Unterschiede in den Prognosefehlern des *mwMAE*⁴

Verfahren	NF1	NF2	MA	S.ES	DT.ES	EXP.ES	EXP.ES ARIMA	EXP.ES &AR	mv. EXP.ES	MLP	mv.MLP
NF1		0,913	0,000**	0,000**	0,003**	0,003**	0,002**	0,002**	0,000**	0,000**	0,000**
NF2	0,913		0,000**	0,000**	0,011*	0,011*	0,002**	0,003**	0,001**	0,000**	0,000**
MA	0,000**	0,000**		0,000**	0,880	0,783	0,812	0,846	0,429	0,056	0,001**
S.ES	0,000**	0,000**	0,000**		0,282	0,200	0,047*	0,246	0,717	0,578	0,040*
DT.ES	0,003**	0,011*	0,880	0,282		0,715	1,000	0,559	0,042*	0,050*	0,001**
EXP.ES	0,003**	0,011*	0,783	0,200	0,715		0,689	0,727	0,065	0,036*	0,003**
EXP.ARIMA	0,002**	0,002**	0,812	0,047	1,000	0,689		0,721	0,220	0,044*	0,001**
EXP.ES&AR	0,002**	0,003**	0,846	0,246	0,559	0,727	0,721		0,033*	0,079	0,002**
mv.EXP.ES	0,000**	0,001**	0,429	0,717	0,042*	0,065	0,220	0,033*		0,270	0,039*
NN	0,000**	0,000**	0,056	0,578	0,050*	0,036*	0,044*	0,079	0,270		0,022*
mv.MLP	0,000**	0,000**	0,001**	0,040*	0,001**	0,003**	0,001**	0,002**	0,039**	0,022**	
Test mind. *	9	9	4	4	5	4	5	4	5	6	10
Test mind. **	9	7	4	3	2	2	3	3	3	3	7

* Wilcoxon-Test ist signifikant bei einem $p \leq 0.05$ Konfidenzniveau (2-seitiger Test, Berechnung mit SPSS 13.0).

** Wilcoxon-Test ist hochsignifikant bei einem $p \leq 0.01$ Konfidenzniveau (2-seitiger Test, Berechnung mit SPSS 13.0).

Eine detaillierte Aufstellung aller Fehlermaße der summierten, mittleren oder medianbasierten *MAE*⁵ und ihre korrespondierende Rangordnung⁶ liefert Tabelle 6.14. Die Tabelle erlaubt den

¹ Der Aufbau der Studie bildet Faktorstufen mit einer inhomogenen Anzahl von Beobachtungswerten und inhomogenen Varianzen durch die unterschiedlichen Gruppen von Verfahren und Zeitreihen. Somit kann durch ANOVA kein valider Test auf Signifikanz berechnet werden, und es muss auf paarweise Vergleiche der Verfahren durch Student-t-Tests oder Wilcoxon-Vorzeichen-Rang-Tests mit inhomogenen Varianzen zurückgegriffen werden.

² Einzelne Verteilungen der *mwMAE* folgen einer Exponentialverteilung mit nur schwach insignifikanten KS-Tests für NF1 ($p=0,075$) und NF2 ($p=0,070$). Die Fehler anderer Verfahren sind ebenfalls signifikant nicht-normalverteilt und folgen am nächsten einer Exponentialverteilung mit schwach signifikanten Tests für die Verfahren MA ($p=0,027$), S.ES ($p=0,024$), DT.ES ($p=0,023$), EXP.ES ($p=0,038$), EXP.ARIMA ($p=0,004$), EXP.ES&ARIMA ($p=0,024$), mv.EXP.ES ($p=0,017$), MLP ($p=0,006$) und mv.MLP ($p=0,003$). Somit kann die Prämisse der Normalverteilung für kein Verfahren bestätigt werden.

³ Alle Ergebnisse wurden ebenfalls durch den Vorzeichen- bzw. Sign-Test bestätigt. Darüber hinaus wurden die Ergebnisse, die Signifikanz der Unterschiede sowie die Rangordnung der Verfahren durch einen nichtparametrischen Friedman-Test bestätigt ($p=0,000$), der auf signifikante Unterschiede in wiederholten ordinalen Messungen im Sinne einer ANOVA der Ränge testet.

⁴ Für signifikante Werte * unter einer Irrtumswahrscheinlichkeit von $p \leq 0,05$ und hoch signifikante Werte ** unter $p \leq 0,01$ ist anzunehmen, dass die verbesserte Prognosegenauigkeit nicht auf zufällige Schwankungen zurückzuführen ist, sondern auf das Prognoseverfahren. Zusätzlich wird die Anzahl mindestens signifikanter und mindestens hoch signifikanter Tests je Prognoseverfahren dargestellt.

⁵ Die Darstellung folgt der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“. Der *mwMAE* ergibt sich beispielsweise als Mittelwert über alle *AE* aller Zeitpunkte, je Datensatz der Trainings-, Validierungs- und Testdaten berechnet und als Tupel ausgegeben. Für den dargestellten Fehler auf dem Testdatensatz entspricht dies z. B. dem Mittelwert aus 21 $t+1$ Prognosen zu unterschiedlichen Zeitreihenursprüngen, worauf nachfolgend der Mittelwert über alle 44 Zeitreihen gebildet wird.

⁶ Zur Verdeutlichung der Ergebnisse wird zu jedem Fehlermaß der relative Rang eines jeden Verfahrens berechnet. Dabei erhält das Verfahren mit dem niedrigsten Fehlerwert den Rang 1, das Verfahren mit dem nächsthöheren Fehlerwert den Rang 2 usw. Bei einer höheren Transparenz der ordinalen Rangreihung erlaubt

Vergleich der relativen Prognosegüte nach unterschiedlichen Fehlermaßen, wobei die summierten *sumMAE* im Gegensatz zum Mittelwert *mwMAE* eine unmittelbare Einschätzung der Höhe der verringerten Fehler – summiert über alle Zeitreihen des Experimentes bei identischer Rangordnung – erlauben.¹

Die statistischen Tests des *mwMAE* auf den Testdaten in Tabelle 6.13 belegen, dass die Prognosefehler des mv.MLP im Vergleich zu allen Prognoseverfahren signifikant geringer sind (mit $p \leq 0,05$), in sieben der zehn Verfahren sogar hochsignifikant (mit $p \leq 0,01$). Die Prognosefehler sind konsistent niedriger für *sumMAE* und *mwMAE* auf den Testdaten sowie auf den Trainings- und Validierungsdaten in Tabelle 6.14. Somit ist das mvMLP nicht nur in der Lage, den datengenerierenden Prozess auf der Kalibrationsstichprobe im Vergleich zu allen zehn Prognoseverfahren am besten zu erlernen, sondern auch auf den ungesesehenen Testdaten zu generalisieren. Es belegt konsistent den höchsten Rang auf allen Datenmengen (1 / 1) 1. Die überlegene Prognosegüte wird durch die ordinalen Fehlermaße der summierten Rangfolge *sumRang MAE*, *mwRang MAE* und *mdRang MAE* bestätigt, wo es mit (1 / 1) 1 konsistent als bestes Verfahren identifiziert wird. Somit kann die Nullhypothese H_a aus Abschnitt 6.1.1 für mv.MLP im Vergleich zu allen anderen Prognoseverfahren verworfen werden, da das Verfahren der NN keine schlechtere Prognosegenauigkeit erzielt als etablierte statistische Benchmark-Verfahren. Darüber hinaus kann für den gegebenen Datensatz sogar überlegene Prognosegenauigkeit festgestellt werden.

Auch die Prognosefehler des univariaten MLP sind signifikant niedriger als die von sechs anderen Verfahren, jedoch nicht statistisch signifikant unterschiedlich von jenen der S.ES ($p = 0,578$), der mv.EXP.ES ($p = 0,270$), des EXP.ES&ARIMA ($p = 0,079$) sowie des MA ($p = 0,056$). Während das MLP mit einem *mwMAE* von 3,39 im Vergleich zum S.ES mit 3,36 nur insignifikant schlechter abschneidet, zeigt es eine konstante Prognoseleistung auf allen Datenteilmengen (3 / 2) 3 als dritt- bzw. zweitbestes Verfahren. Das Verfahren der S.ES hingegen zeigt bei zweitbesten Leistung auf der Testmenge nur eine geringe Leistung auf Trainings- und Validierungsmenge als achtbestes Verfahren, bzw. (8 / 8) 2. Das S.ES ist somit nicht so gut wie das MLP in der Lage, die Zeitreihenstruktur der Trainingsdatenmenge zu erlernen und auf der Validierungsmenge wiederzugeben. Dies stellt die Konstanz seiner Prognoseleistung in Frage und deutet auf mögliche Zufallseinflüsse beim positiven Abschneiden des S.ES auf den verrauschten Daten hin. Um die Robustheit dieser Ergebnisse zu analysieren, werden die Fehlermaße des Medians und die ordinalen Gütemaße des summierten Rangs verwendet.² Während das MLP nach dem Rang des *sumMAE*, des *mwMAE* und des *mdMAE* (3 / 2) 3 kontinuierlich als zweit- bzw. drittbestes Verfahren identifiziert wird, fällt die S.ES vom zweiten auf den vierten Rang zurück (8 / 4) 4. Auch in einer rangbasierten Analyse der ordinalen Fehlermaße zeigt das MLP mit (3 / 2) 2 nach *sumRang*, (3 / 2) 2 nach *mwRang* und (2 / 2) 2 nach *mdRang* der *MAE* konsistent eine bessere Prognosegüte als das S.ES mit (8 / 8) 5, (8 / 8) 5 und (8 / 8) 2. Somit kann das neuronale Netz MLP aufgrund einer robusteren Leistung ebenfalls als valides univariates Prognoseverfahren identifiziert werden.

diese Form der Betrachtung keine Rückschlüsse mehr auf die absolute Distanz der Prognosegüte zweier oder mehrerer Verfahren, z. B. in Verpackungseinheiten oder Prozent. Somit stellen die Fehlermaße der *mwMAE* die Güte- und Entscheidungskriterien zur Beurteilung der Prognosegüte dar.

¹ Die Analyse der summierten MAE zeigt beispielsweise eine Verbesserung der Prognosegenauigkeit von 20,02 Packungseinheiten bzw. 12,7% durch Einsatz des mv.MLP gegenüber dem Branchen-Benchmark des MA. Die verbesserte Prognosegenauigkeit in Höhe von 20,02 über alle 41 Zeitreihen entspricht mit einer durchschnittlichen Verbesserung des *sumMAE* von 0,48 der des *mwMAE* mit 0,5 Packungseinheiten. Die Betrachtung summierter Fehler *sumMAE* zeigt identische Ränge wie *mwMAE*, da die *MAE* je Verfahren über eine identische Anzahl von Zeitreihen und Prognoseursprüngen aggregiert werden.

² Der Median des *MAE* zeigt als robustes Fehlermaß die Variabilität der relativen Prognosegüte der Verfahren an, während die ordinalen Gütemaße der Rangreihung durch Identifikation des besten, zweitbesten etc. Verfahrens je Zeitreihe nicht so stark durch eine besonders gute oder schlechte Leistung auf einer einzelnen Zeitreihe verzerrt werden können.

Tabelle 6.14: Fehlermaße und Rang der Fehlermaße des MAE über alle Absatzstellen

Verfahren	MAE		
	<i>sumMAE</i>	<i>mwMAE</i>	<i>mdMAE</i>
NF1	(249,72 / 214,94) 195,44	(5,68 / 4,89) 4,44	(2,85 / 3,15) 2,62
NF2	(242,62 / 206,17) 195,91	(5,51 / 4,69) 4,45	(2,88 / 3,08) 2,74
MA	(197,71 / 172,92) 157,60	(4,49 / 3,93) 3,58	(2,44 / 2,54) 2,26
S.ES	(187,59 / 165,67) 147,81	(4,26 / 3,77) 3,36	(2,35 / 2,36) 2,17
DT.ES	(175,06 / 157,19) 172,33	(3,98 / 3,57) 3,92	(2,20 / 2,35) 2,17
EXP.ES	(176,38 / 165,44) 168,89	(4,01 / 3,76) 3,84	(2,21 / 2,44) 2,29
EXP.ARIMA	(175,21 / 163,06) 153,02	(3,98 / 3,71) 3,48	(2,23 / 2,40) 2,24
EXP.ES&ARIMA	(179,16 / 164,92) 166,94	(4,07 / 3,75) 3,79	(2,20 / 2,46) 2,24
mv.EXP.ES	(165,07 / 159,86) 160,51	(3,75 / 3,63) 3,65	(2,07 / 2,40) 2,00
MLP	(169,68 / 146,36) 149,31	(3,86 / 3,33) 3,39	(2,09 / 2,22) 2,14
mv.MLP	(139,70 / 137,28) 137,62	(3,18 / 3,12) 3,13	(2,07 / 2,07) 2,07

Verfahren	Rang MAE		
	<i>Rang sumMAE</i>	<i>Rang mwMAE</i>	<i>Rang mdMAE</i>
NF1	(11 / 11) 10	(11 / 11) 10	(10 / 11) 10
NF2	(10 / 10) 11	(10 / 10) 11	(11 / 10) 11
MA	(9 / 9) 5	(9 / 9) 5	(9 / 9) 8
S.ES	(8 / 8) 2	(8 / 8) 2	(8 / 4) 4
DT.ES	(4 / 3) 9	(4 / 3) 9	(4 / 3) 4
EXP.ES	(6 / 7) 8	(6 / 7) 8	(6 / 7) 9
EXP.ARIMA	(5 / 5) 4	(5 / 5) 4	(7 / 5) 6
EXP.ES&ARIMA	(7 / 6) 7	(7 / 6) 7	(4 / 8) 6
mv.EXP.ES	(2 / 4) 6	(2 / 4) 6	(1 / 5) 1
MLP	(3 / 2) 3	(3 / 2) 3	(3 / 2) 3
mv.MLP	(1 / 1) 1	(1 / 1) 1	(2 / 1) 2

Verfahren	sumRang MAE		
	<i>sumRang MAE</i>	<i>mwRang MAE</i>	<i>mdRang MAE</i>
NF1	(453 / 436) 368	(10,3 / 9,9) 8,4	(10,0 / 10,0) 10,0
NF2	(449 / 434) 355	(10,2 / 9,9) 8,1	(10,5 / 10,0) 10,0
MA	(343 / 329) 255	(7,8 / 7,5) 5,8	(9,0 / 9,0) 5,5
S.ES	(299 / 280) 201	(6,8 / 6,4) 4,6	(7,0 / 7,0) 4,0
DT.ES	(182 / 198) 220	(4,1 / 4,5) 5,0	(4,0 / 4,0) 5,0
EXP.ES	(194 / 212) 209	(4,4 / 4,8) 4,8	(4,0 / 4,0) 4,5
EXP.ARIMA	(194 / 225) 241	(4,4 / 5,1) 5,5	(4,0 / 4,5) 5,5
EXP.ES&ARIMA	(209 / 216) 199	(4,8 / 4,9) 4,5	(5,0 / 4,0) 4,5
mv.EXP.ES	(167 / 178) 188	(3,8 / 4,0) 4,3	(3,5 / 4,0) 4,0
MLP	(182 / 107) 188	(4,1 / 2,4) 4,3	(2,0 / 2,0) 4,0
mv.MLP	(100 / 87) 157	(2,3 / 2,0) 3,6	(1,0 / 1,0) 3,0

Verfahren	Rang sumRang MAE		
	<i>Rang sumRang MAE</i>	<i>Rang mwRang MAE</i>	<i>Rang mdRang MAE</i>
NF1	(11 / 11) 11	(11 / 11) 11	(10 / 10) 10
NF2	(10 / 10) 10	(10 / 10) 10	(11 / 10) 10
MA	(9 / 9) 9	(9 / 9) 9	(9 / 9) 8
S.ES	(8 / 8) 5	(8 / 8) 5	(8 / 8) 2
DT.ES	(3 / 4) 7	(3 / 4) 7	(4 / 3) 7
EXP.ES	(5 / 5) 6	(5 / 5) 6	(4 / 3) 5
EXP.ARIMA	(5 / 7) 8	(5 / 7) 8	(4 / 7) 8
EXP.ES&ARIMA	(7 / 6) 4	(7 / 6) 4	(7 / 3) 5
mv.EXP.ES	(2 / 3) 2	(2 / 3) 2	(3 / 3) 2
MLP	(3 / 2) 2	(3 / 2) 2	(2 / 2) 2
mv.MLP	(1 / 1) 1	(1 / 1) 1	(1 / 1) 1

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

Unter Berücksichtigung der statistischen Signifikanz kann so festgestellt werden, dass die Prognosegenauigkeit des MLP nicht signifikant schlechter ist als die aller getesteten Verfahren und statistischen Benchmarks. Somit kann erneut die Nullhypothese H_a verworfen werden. Darüber hinaus zeigt das univariate MLP vielfach eine signifikant höhere Prognosegenauigkeit als die einfachen Verfahren und eine den führenden Benchmark-Verfahren der Absatzprognose gleichwertige Prognosegenauigkeit.

Ein Vergleich der Prognosegenauigkeit des mv.MLP mit dem gleitenden Mittelwert MA, welcher aufgrund seiner verbreiteten Anwendung in der Disposition an Warenautomaten als empirischer Benchmark gilt, zeigt eine hochsignifikante Verbesserung ($p < 0,01$) der Prognosegenauigkeit des *mwMAE* von 3,6 auf 3,1 um 0,5 Verpackungseinheiten und 16,1% pro Prognosevorgang und Marke. Aggregiert über die durchschnittlich 18 Sorten einer Absatzstelle (siehe Tabelle 6.3) und über die 52 Dispositionszeitpunkte pro Jahr entspricht dies einer verbesserten Prognosegenauigkeit von 468 Verpackungseinheiten pro Absatzstelle und damit einer erheblichen Kapitalfreisetzung.¹

Auch im Vergleich zum zweitbesten Verfahren S.ES erzielt das mv.MLP eine signifikant ($p = 0,040^*$) erhöhte Prognosegenauigkeit von 0,23 Verpackungseinheiten oder 7,3% des *mwMAE*. Dies kann sowohl auf die erweiterte Datenbasis eines multivariaten Verfahrens als auch auf die Eigenschaften der NN zurückgeführt werden und ist genauer zu analysieren. Ein unmittelbarer Vergleich des multivariaten NN mv.MLP mit dem multivariaten statistischen Expertensystem mv.EXP.ES auf der gleichen Datenbasis belegt die signifikant ($p = 0,039^*$) gesteigerte Prognosegüte der NN um 0,5 Verpackungseinheiten und 16,1%. Ein Vergleich der multivariaten Verfahren mv.MLP und mv.EXP.ES mit einem mittleren *mwMAE* von 3,39 mit ihren univariaten Varianten MLP und EXP.ES mit 3,62 zeigt einen geringen, aber konsistenten Zuwachs der Prognosegenauigkeit durch multivariate Prognoseverfahren um 6,4%. Dagegen kann durch die Verfahrensklasse der NN von MLP und mv.MLP eine durchschnittliche Verbesserung der Prognosegenauigkeit um 0,49 Verpackungseinheiten oder 13,1% im Vergleich zu Expertensystemen der Exponentiellen Glättung aus EXP.ES und mv.EXP.ES erzielt werden.² Der Zuwachs an Prognosegenauigkeit ist somit überwiegend den Attributen der Verfahrensklasse der NN zuzuweisen, erst nachrangig der multivariaten Modellierung.

Die Naiven Verfahren NF1 und NF2 für stationäre oder saisonale Zeitreihen zeigen wie erwartet mit $p \leq 0,05$ eine signifikant schlechtere Prognosegüte als alle anderen Verfahren. Sie schneiden mit den höchsten *sumMAE*, *mwMAE* und *mdMAE* sowie dem niedrigsten Rang (11 / 11) 10 bzw. (10 / 10) 11 konsistent als schlechteste Verfahren ab. Dieses Bild wird auch bei Betrachtung der Gütemaße des summierten Rangs bestätigt, wobei sich durch Verwendung der ordinalen Gütemaße ein konsistenteres Bild von (11 / 11) 11 für NF1 und (10 / 10) 10 für NF2 ergibt.³ Dies belegt, dass die Zeitreihen globale oder lokale Informationen enthalten, die durch die statistischen und neuronalen Prognoseverfahren erfolgreich identifiziert und genutzt werden konnten, um bessere Prognosen als die der Naiven Verfahren zu liefern. Die gesteigerte Prognosegüte der statistischen Verfahren belegt zudem ihre Position als fortgeschrittene Benchmark-Verfahren und bestätigt die Rolle der Naiven Verfahren als einfache Benchmarks für eine kritische untere Grenze der Prognosegüte.

¹ Bezogen auf alle Absatzstellen des Unternehmens entspräche dies einer erhöhten Prognosegenauigkeit von über 100 Millionen Verpackungseinheiten pro Jahr. Allerdings verbleibt nicht die prognostizierte Menge als physischer Bestand in den Abverkaufsstellen, sondern nur die anschließend zu bestimmende Bestellmenge. Anhand der Prognosegenauigkeit können daher keine Aussagen über die Höhe möglicher Einsparungen getroffen werden.

² Sie entspricht der Verbesserung von 3,75 des *mwMAE* auf 3,26 als mittlerer Fehler der Verfahren der NN (vgl. Tabelle 6.14).

³ Aufgrund der insignifikanten Unterschiede kann aus dem leichten Vorteil des saisonalen NF2 jedoch nicht auf eine saisonale Struktur der Zeitreihen geschlossen werden.

Die Gruppe der einfachen statistischen Verfahren S.ES, DT.ES und der Expertensysteme EXP.ES, EXP.ARIMA, EXP.ES&AR und mv.EXP.ES zeigt hingegen keine statistisch signifikanten Abweichungen der Fehlerwerte. Sie erlauben somit keine eindeutige Rangordnung von Verfahrensgruppen, weder der multivariaten gegenüber den univariaten Verfahren noch der Expertensysteme gegenüber Prognoseverfahren mit aggregierter Modellselektion. Die Ergebnisse bestätigen aber zumindest für die statistischen Verfahren und insbesondere für S.ES die Nullhypothese H_1 aus vorangegangenen Vergleichsverfahren durch MAKRIDAKIS und HIBON [2000], dass komplexe Verfahren keine höhere Prognosegenauigkeit aufweisen als einfache.¹ Diese Erkenntnis kann auf multivariate Verfahren erweitert werden. Unter Vernachlässigung der Modelle der NN steht dem *mwMAE* des multivariaten Expertensystems von 3,6 ein nur insignifikant höherer durchschnittlicher Prognosefehler aller univariaten Verfahren und Expertensysteme von 3,67 entgegen. Werden MLPs und mv.MLP ebenfalls zu den komplexen Verfahren gezählt, kann H_1 jedoch verworfen werden, da die Verfahren der NN signifikant bessere Ergebnisse als einfache und komplexe statistische Verfahren und Expertensysteme erzielt haben. Diese Erkenntnisse der Mittelwerte, Mediane und Schwankungsbreiten der Ergebnisse werden durch die entsprechenden Box-Whisker-Diagramme auf den Testdaten über alle Zeitreihen und Zeitreihenursprüngen in Abbildung 6.15 verdeutlicht.

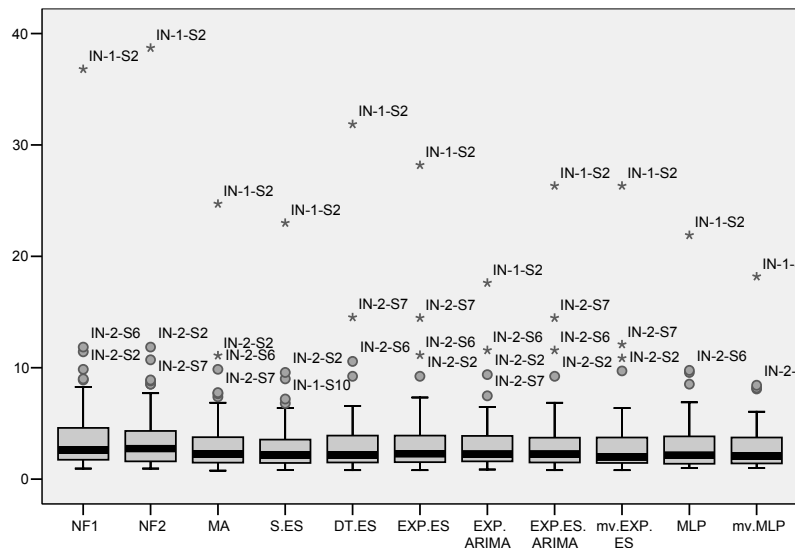


Abbildung 6.15: Box-Whisker-Grafiken des *mwMAE* auf den Testdaten über alle Zeitreihen

Die Verteilung der Prognosefehler auf der Testmenge in Abbildung 6.15 verdeutlicht die Unterschiede nach der Betrachtung von Mittelwerten, Medianen und resultierenden Rangfolgen sowie die Indifferenz der Ergebnisse für einen Großteil der Verfahren.²

Die unterschiedliche Prognosegüte der Varianten des Expertensystems weist auf Probleme des Modellansatzes hin. Das Expertensystem zur reinen ARIMA-Modellierung EXP.ARIMA zeigt eine gesteigerte, wenn auch insignifikante, Prognosegenauigkeit gegenüber dem Expertensystem EXP.ES&ARIMA. Theoretisch hätte die Expertenroutine EXP.ES&ARIMA einen niedrigeren Prognosefehler durch alleinige Anwendung von ARIMA-Modellen erzielen können. Zudem erreichen die univariaten sowie die multivariaten Expertensysteme nicht die Prognosegüte des einfachen univariaten Prognoseverfahrens S.ES ohne Modellselektion und

¹ Die Berechnung des Mittelwertes erfolgt ohne Berücksichtigung der Naiven Verfahren NF1, NF2 und des gleitenden Mittelwertes MA.

² Ausreißer mit einem 1,5- bis dreifachen Interquartil-Abstand werden als Kreise dargestellt, extreme Werte mit mehr als dreifachem Interquartil-Abstand als Sterne. Ohne diese Korrektur würden die Box-Plots aller Verfahren entsprechend vergrößert.

mit konstantem α -Parameter, welches ebenfalls als Spezialfall aller Expertensysteme hätte ausgewählt werden können. Dies verdeutlicht nachhaltig die Probleme einer validen Modellselektion auf verrauschten Beobachtungswerten einer Kalibrationsstichprobe.¹

Zum anderen weist die vergleichsweise hohe Prognosegenauigkeit der einfachen Verfahren auf die Robustheit des Verfahrens der S.ES auf den stark verrauschten Zeitreihen der Stichprobe mit inkonsistenter Struktur und Niveauverschiebungen hin. Dies kann durch die weite Verbreitung der S.ES als Benchmark im Handel auch als Indiz für die externe Validität der Untersuchung gedeutet werden.² Die vergleichsweise niedrige Prognosegüte des Benchmarks des gedämpften Trends DT.ES kann hingegen auf das geringe Auftreten lokaler Trends in den Zeitreihendaten zurückgeführt werden. Die nur insignifikanten Unterschiede zwischen den Verfahren bestätigen auch die Nullhypothese H_5 nach FILDES, nach der Expertensysteme mit individueller Modellselektion bei einem kurzen Prognosehorizont keine signifikanten Vorteile gegenüber Ansätzen mit aggregierter Modellselektion und ohne Anpassung an die Zeitreihe besitzen. Der mittlere *mwMAE* der Expertensysteme mit individueller Modellselektion von 3,7 ist nur insignifikant höher als der mittlere *mwMAE* einfacher statistischer Prognoseverfahren von 3,6.³ Dies liefert auch eine mögliche Erklärung, warum die Nullhypothese H_6 verworfen werden muss. Aufgrund der insignifikanten Unterschiede kann nicht bestätigt werden, dass Verfahren mit aggregierter Modellselektion und Parameteroptimierung wie EXP.ES geringere Prognosefehler aufweisen als Verfahren ohne Parameteroptimierung wie S.ES.

Somit wird zunächst festgestellt, dass multivariate Verfahren sowie Expertensysteme mit individueller Modellselektion je Zeitreihe keine signifikant höhere Prognosegenauigkeit bieten als univariate statistische Prognoseverfahren ohne Modellselektion oder NN. Der Verfahrenvergleich etabliert somit NN als neues Benchmark-Verfahren zur Prognose, unter Bestätigung bestehender Erkenntnisse aus vorangegangenen Vergleichsstudien.

Die Analyse der Fehlermaße und Rangordnungen verdeutlicht, dass neben den besten und schlechtesten Verfahren keine konsistente und homogene Rangordnung der Verfahren auf der Testmenge möglich ist. Dies würde eine tiefere Analyse der Konsistenz und Robustheit der Ergebnisse und möglicher Einflussfaktoren erfordern. Zudem entsteht eine abweichende Rangordnung der Verfahren nach ihren Trainings- oder Validierungsfehlern, was in Frage stellt, ob ein einzelnes Verfahren ex ante anhand des Fehlers auf der Validierungsmenge als überlegen hätte identifiziert werden können. Diese Annahme wird auch durch die geringe Prognosegüte des S.ES auf Trainings- und Validierungsmenge im Vergleich zur hohen Genauigkeit auf der Testmenge sowie die inkonsistente Modellselektion der Expertensysteme unterstützt. Hierzu erfolgt eine vergleichende Analyse der Korrelationskoeffizienten der Prognosefehler *mwMAE* auf Trainings-, Validierungs- und Testmenge, basierend auf den metrischen Distanzen nach PEARSON sowie den ordinalen Rangfolgen nach SPEARMAN, die in Tabelle 6.15 dargestellt sind.

¹ Diese Ergebnisse sind robust über mehrere Zeitreihenursprünge und Zeitreihen, was jedoch nicht zur Falsifikation der Nullhypothese H_4 interpretiert werden kann, nach der die Variabilität der Beurteilung von Verfahren ihre Untersuchung über mehrere Zeitreihen-Ursprünge erfordert.

² Die überraschend höhere Prognosegüte einfacher statistischer Verfahren des S.ES und MA ist jedoch nicht statistisch signifikant und sollte nicht überbewertet werden.

³ Die Berechnung erfolgt ohne Berücksichtigung der NN oder der Naiven Benchmarks.

Tabelle 6.15: Korrelationen von Trainings-, Validierungs- und Testfehler der *mwMAE*

	PEARSON-Korrelation des <i>mwMAE</i>			SPEARMANS Korrelation des Rangs nach <i>mwMAE</i>		
	Training	Validierung	Test	Training	Validierung	Test
Train	1,000	0,980**	0,859**	1,000	0,964**	0,545
Valid	0,980**	1,000	0,886**	0,964**	1,000	0,555
Test	0,859**	0,886**	1,000	0,545	0,555	1,000

* Korrelation ist signifikant bei einem 0.05 Konfidenzniveau (2-seitiger Test, Berechnung mit SPSS 13.0).

** Korrelation ist hochsignifikant bei einem 0.01 Konfidenzniveau (2-seitiger Test, Berechnung mit SPSS 13.0).

Die parametrischen Korrelationskoeffizienten nach PEARSON zeigen eine hochsignifikante positive Korrelation zwischen Trainings- und Validierungsfehler von 0,980 sowie zwischen Validierungs- und Testfehler von 0,886. Somit könnte ein Prognoseverfahren mit geringem Testfehler anhand eines niedrigen Validierungsfehlers identifiziert und angewendet werden.¹ Dagegen berechnet der nichtparametrische Korrelationskoeffizient nach SPEARMANS rho die Korrelation der ordinalen Rangordnung der Verfahren nach dem *mwMAE*.² Während die Korrelation von Trainings- und Validierungsfehler mit 0,964 weiterhin hochsignifikant ist, sind die positiven Korrelationen von Trainings- und Testfehler mit 0,545 sowie von Validierungs- und Testfehler mit 0,555 nicht statistisch signifikant. Dies bestätigt die Probleme in der Auswahl adäquater Verfahren anhand ihrer Prognosegüte auf der Kalibrierungstichprobe. Die Diskrepanz der metrischen und ordinalen Gütekriterien kann durch einen Vergleich der Korrelationsdiagramme der Prognoseverfahren über die Teildatenmengen visualisiert werden. Diese sind in Abbildung 6.16 a.) für die Korrelation nach dem metrischen *mwMAE* und in b.) nach dem ordinalen *Rang mwMAE* dargestellt.

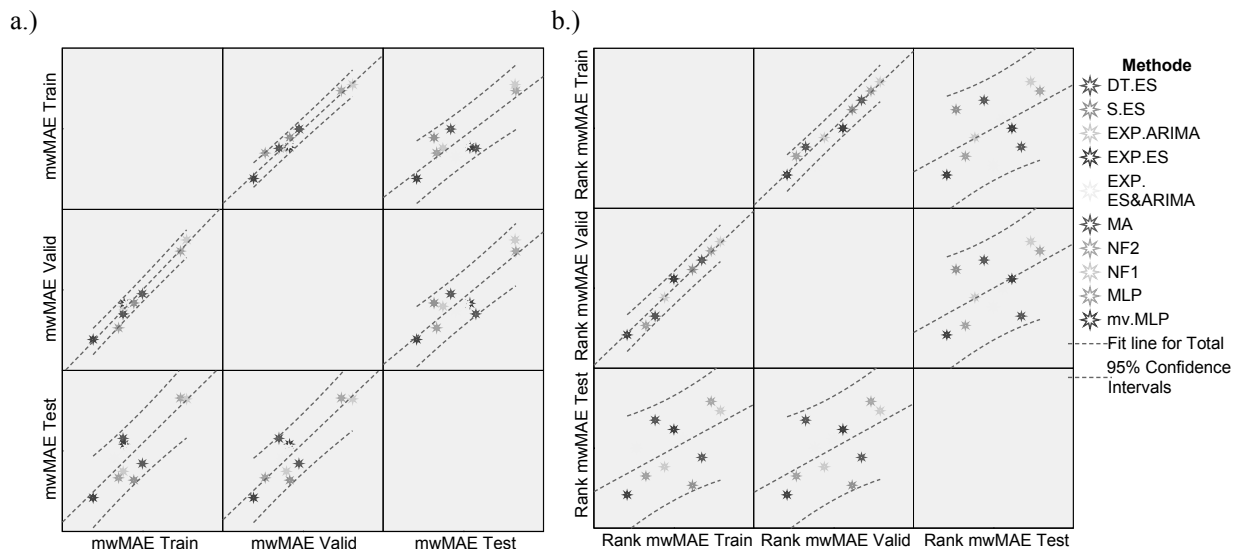


Abbildung 6.16: Korrelationsdiagramm der Trainings-, Validierungs- und Testfehler des *mwMAE* (a.) und des *Rang mwMAE* (b.) über alle Zeitreihen.³

¹ Dies belegt auch die Konsistenz der Prognoseergebnisse über alle Datensätze: Ein überlegenes Verfahren wie mv.MLP oder MLP mit niedrigem Fehler auf der Trainingsmenge zeigt auch niedrige Prognosefehler in der tatsächlichen Anwendung auf der Testmenge und umgekehrt.

² Die Normalverteilungsprämisse ist bei den Fehlerwerten des *mwMAE* nicht gegeben. Somit werden die Korrelationskoeffizienten nach PEARSON durch Ausreißer dominiert und überhöhte Korrelationen ausgewiesen. Sie korrespondiert jedoch mit der Analyse der ursprünglichen Fehlerwerte nach *mwMAE* und wird daher hier zum Vergleich dargestellt. Die Berechnung von SPEARMANS rho erfolgt auf ordinal geordneten Beobachtungen und somit rangbasiert. Daher sind seine Ergebnisse auf den metrisch skalierten *mwMAE* identisch mit denen auf den ordinal skalierten *Rang mwMAE*. Die Ergebnisse gelten auch bei Gültigkeit der parametrischen Prämisse, zum Beispiel der Normalverteilungshypothese, und liefern robuste Ergebnisse. Alle Ergebnisse nach SPEARMANS rho werden durch die Ergebnisse nach KENDALL's tau-b bestätigt, auf deren Abbildung hier verzichtet wird.

³ Die Linien repräsentieren den Mittelwert und die 95%-Konfidenzintervalle auf den jeweiligen Datenmengen. Die Verfahren mit den niedrigsten Fehlern auf der Validierungs- und Testmenge werden im unteren linken

Die Darstellung verdeutlicht neben einer positiven Korrelation auf allen Mengen die Dominanz der Maxima und Minima auf die metrischen Korrelationen nach $mwMAE$ in Abbildung 6.16.a sowie die zunehmende Streuung der rangbasierten Korrelationen durch die vergrößerten 95%-Konfidenzintervalle auf den Testfehlern in Abbildung 6.16.b. Die für die Modellauswahl relevante Korrelation der Fehlermaße von Validierungs- und Testdaten wird in Abbildung 6.17 detailliert dargestellt, indem die Beobachtungen nach aufsteigendem Validierungsfehler geordnet und durch eine Linie verbunden werden.¹

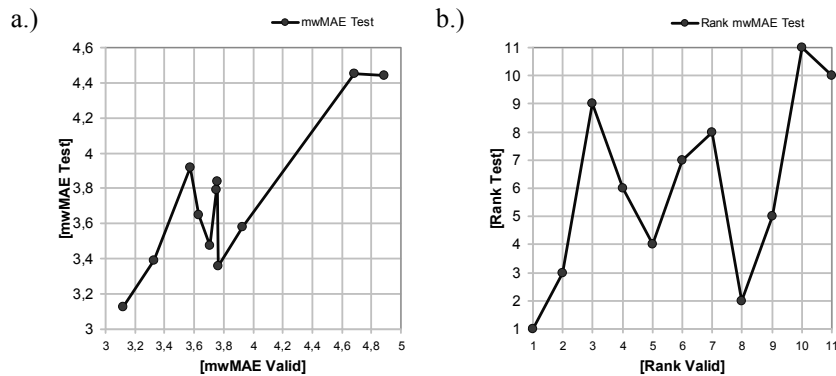


Abbildung 6.17: Korrelationsdiagramm der Validierungs- und Testfehler des $mwMAE$ (a.) und des Rang $mwMAE$ (b.) über alle Zeitreihen.

Die Verfahren des mv.MLP und MLP mit den geringsten Fehlern auf Validierungs- und Testmenge im unteren linken Bereich der Korrelationsdiagramme sowie die Verfahren NF1 und NF2 mit den höchsten Prognosefehlern im oberen rechten Bereich zeigen eine klare Korrelation und erlauben eine problemlose Rangreihung. Dagegen erscheint eine Rangbildung für Verfahren mit mittleren Prognosefehlern kaum realisierbar, da Verfahren mit höherem Validierungsfehler vielfach niedrigere Testfehler aufweisen.² Somit kann lediglich die überlegene Leistung der Verfahren des mv.MLP und MLP sowie die schwache Leistung von NF1 und NF2 mit statistischer Signifikanz festgestellt werden. Die relative Leistung der verbleibenden Verfahren liefert hingegen inkonsistente Ergebnisse, die nur schwer einer Verallgemeinerung Genüge tun. Dies belegt die Probleme einer validen Bestimmung der Prognosegüte konkurrierender Verfahren durch unterschiedliche Fehlermaße sowie die Herausforderung einer Ex-ante-Verfahrensauswahl auf verrauschten Daten. Ferner wird durch die grafische Darstellung deutlich, dass der metrisch skalierte $mwMAE$ in Abbildung 6.17.a im Vergleich zum ordinalen Rang $mwMAE$ in Abbildung 6.17.b durch die metrische Distanz der Fehler zusätzliche Informationen in der Beurteilung der Verfahrensgüte berücksichtigt, im Gegensatz zur robusteren Verwendung der ordinalen Abstände der geordneten Rangfolge. Dies bestätigt seine Stellung als das relevante Fehlermaß zur Beurteilung der Verfahren sowie die Notwendigkeit zur Berücksichtigung anderer, robuster Gütekriterien.

Bereich dargestellt, die Verfahren mit den höchsten Validierungs- und Testfehlern im oberen rechten Bereich des Korrelationsdiagramms

¹ Die Darstellungen entsprechen den Abbildungen der mittleren unteren Quadranten der Abbildung 6.16 a.) und b.), die dort ohne Linien als Korrelationsdiagramm dargestellt sind.

² Bei einer Anordnung nach aufsteigendem Validierungsfehler entlang der Abszisse ist kein konsistentes Muster der Generalisierungsfehler auf der Ordinate erkennbar. So führt beispielsweise die Auswahl des drittbesten Verfahrens auf der Validierungsmenge DT.ES zu einer suboptimalen Leistung auf den Testdaten mit einem neunten Rang, mit Rang $mwMAE$ (4 / 3) 9. Im Vergleich würde die Wahl der sechs nachfolgenden Verfahren nach dem Validierungsfehler zu einem niedrigeren Testfehler führen. Das viertbeste Verfahren des EXP.ARIMA zeigt beispielsweise einen niedrigeren Testfehler und zudem einen konsistenteren Rang der Fehlermaße, mit Rang $mwMAE$ (5 / 5) 4. Das achtbeste Verfahren des S.ES besitzt dagegen einen niedrigen Testfehler, der nicht aus dem Validierungsfehler erkennbar ist, mit Rang $mwMAE$ (8 / 8) 2.

Tabelle 6.16: Fehlermaße und Rang der Fehlermaße des MAPE über alle Absatzstellen

Verfahren	MAPE		
	<i>sumMAPE</i>	<i>mwMAPE</i>	<i>mdMAPE</i>
NF1	-	(80,8% / 79,3%) 75,0%	(83,1% / 80,5%) 74,6%
NF2	-	(83,0% / 83,4%) 76,7%	(86,9% / 82,7%) 83,0%
MA	-	(74,7% / 73,1%) 70,5%	(78,9% / 75,3%) 71,6%
S.ES	-	(75,1% / 75,9%) 70,2%	(79,4% / 77,2%) 71,1%
DT.ES	-	(74,3% / 73,9%) 79,0%	(77,4% / 72,1%) 72,4%
EXP.ES	-	(76,1% / 76,6%) 79,4%	(79,0% / 77,4%) 73,4%
EXP.ARIMA	-	(75,3% / 77,3%) 80,2%	(76,9% / 80,8%) 83,6%
EXP.ES&ARIMA	-	(76,3% / 77,4%) 79,7%	(80,7% / 77,9%) 75,9%
mv.EXP.ES	-	(72,7% / 73,2%) 77,2%	(77,7% / 75,3%) 69,9%
MLP	-	(72,5% / 66,7%) 69,8%	(71,4% / 67,8%) 67,6%
mv.MLP	-	(67,8% / 68,8%) 70,7%	(68,6% / 67,2%) 67,2%

Verfahren	Rang MAPE		
	<i>Rang sumMAPE</i>	<i>Rang mwMAPE</i>	<i>Rang mdMAPE</i>
NF1	-	(10 / 10) 5	(10 / 9) 8
NF2	-	(11 / 11) 6	(11 / 11) 10
MA	-	(5 / 3) 3	(6 / 4) 5
S.ES	-	(6 / 6) 2	(8 / 6) 4
DT.ES	-	(4 / 5) 8	(4 / 3) 6
EXP.ES	-	(8 / 7) 9	(7 / 7) 7
EXP.ARIMA	-	(7 / 8) 11	(3 / 10) 11
EXP.ES&ARIMA	-	(9 / 9) 10	(9 / 8) 9
mv.EXP.ES	-	(3 / 4) 7	(5 / 5) 3
MLP	-	(2 / 1) 1	(2 / 2) 2
mv.MLP	-	(1 / 2) 4	(1 / 1) 1

Verfahren	<i>sumRang MAPE</i>		
	<i>sumRang MAPE</i>	<i>mwRang MAPE</i>	<i>MdRang MAPE</i>
NF1	-	(8,7 / 8,3) 7,0	(10 / 10) 10
NF2	-	(9,0 / 8,4) 7,1	(11 / 11) 11
MA	-	(5,9 / 6,2) 5,4	(9 / 7) 6
S.ES	-	(5,8 / 6,6) 5,1	(8 / 9) 3
DT.ES	-	(4,5 / 5,2) 5,3	(3 / 4) 5
EXP.ES	-	(5,4 / 5,5) 5,8	(5 / 6) 8
EXP.ARIMA	-	(5,5 / 6,2) 5,8	(6 / 7) 8
EXP.ES&ARIMA	-	(5,5 / 5,4) 5,7	(7 / 5) 7
mv.EXP.ES	-	(4,3 / 4,4) 5,1	(2 / 3) 4
MLP	-	(4,9 / 2,9) 4,3	(4 / 1) 2
mv.MLP	-	(3,9 / 3,4) 4,3	(1 / 2) 1

Verfahren	<i>Rang sumRang MAPE</i>		
	<i>Rang sumRang MAPE</i>	<i>Rang mwRang MAPE</i>	<i>Rang mdRang MAPE</i>
NF1	-	(10 / 10) 10	(10 / 10) 10
NF2	-	(11 / 11) 11	(11 / 10) 10
MA	-	(9 / 7) 6	(9 / 9) 8
S.ES	-	(8 / 9) 3	(8 / 8) 2
DT.ES	-	(3 / 4) 5	(4 / 3) 7
EXP.ES	-	(5 / 6) 8	(4 / 3) 5
EXP.ARIMA	-	(6 / 7) 8	(4 / 7) 8
EXP.ES&ARIMA	-	(7 / 5) 7	(7 / 3) 5
mv.EXP.ES	-	(2 / 3) 4	(3 / 3) 2
MLP	-	(4 / 1) 2	(2 / 2) 2
mv.MLP	-	(1 / 2) 1	(1 / 1) 1

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

Als weiteres Fehlermaß wird der *MAPE* berücksichtigt, und seine Ergebnisse werden in Tabelle 6.16 präsentiert, um die relative Genauigkeit der Prognoseverfahren über die Fehlermaße hinaus auf Konsistenz überprüfen zu können.¹

Die Analyse der Prognosegüte nach *mwMAPE* zeigt wie erwartet leichte Abweichungen zur Güte nach *mwMAE*. Auch wenn mv.MLP, MLP und S.ES weiter zu den besten Verfahren gehören, liefert nun das MLP, gefolgt von S.ES, MA und erst anschließend vom mv.MLP, die besten Prognosen. Darüber hinaus gewinnen die Naiven Verfahren NF1 und NF2 deutlich an Genauigkeit, während Varianten der univariaten Expertensysteme die niedrigste Prognosegüte nach *MAPE* aufweisen. Die Analyse der Prognosegüte nach dem robusteren *mdMAPE* bestätigt hingegen erneut das mv.MLP als bestes Verfahren, gefolgt von MLP und S.ES. Auch die Beurteilung nach den rangbasierten Fehlermaßen weist das mv.MLP als bestes multivariates Verfahren und das MLP als bestes univariates Verfahren aus. Dies kann als Indiz für die besondere Reagibilität des *MAPE* auf einzelne Ausreißer der Prognosegüte aufgefasst werden, während robuste Varianten des Medians oder der Rangliste nach *MAPE* die Erkenntnisse des *MAE* und *Rang MAE* bestätigen.

Die abweichende Güte und Rangordnung der Verfahren bestätigt Erkenntnisse der M3-Competition, wonach prozentuale Fehlermaße trotz weiter Verbreitung in der Praxis zu erheblichen Verzerrungen in der Beurteilung der Prognosegenauigkeit konkurrierender Verfahren führen können. Somit wird auch die Nullhypothese H_2 aus Abschnitt 6.1.1, nach der unterschiedliche Fehlermaße zu unterschiedlichen Rangordnungen von Prognoseverfahren führen, bestätigt. Unterschiedliche Fehlermaße begünstigen somit unterschiedliche Prognoseverfahren, so dass die Ex-ante-Bestimmung eines robusten Fehlermaßes gemäß der Zielsetzung des Verfahrensvergleichs aus der Anwendungsdomäne eine herausragende Bedeutung erhält.

Die Höhe des prozentualen Prognosefehlers von durchschnittlich über 70% verdeutlicht auch die relative Stärke der Zufallsschwankungen in den Absatzzahlen, bedingt durch eine Vielzahl von Zeitreihen mit geringem Artikelumschlag. Gemäß dem Versuchsaufbau wird daher nachfolgend untersucht, ob einzelne Absatzstellen, Zeitreihen, Absatzstrukturen, Umschlaggeschwindigkeiten oder Artikelklassifikationen identifiziert werden können, die eine Anwendung von NN oder eines anderen Prognoseverfahrens begünstigen.

6.3.1.2 Ergebnisse nach Absatzstelle

Um Robustheit und Konsistenz der Verfahrensgüte zu überprüfen, wird die Analyse von einer Betrachtung über alle Zeitreihen auf die Genauigkeit der Prognoseverfahren je Abverkaufsstelle ausgeweitet, deren Ergebnisse in Tabelle 6.17 präsentiert sind.²

¹ Auf eine Berechnung der Summe der *MAPE* über alle Zeitreihen wird mangels Aussagekraft verzichtet. Wie unter 3.4.2 dargestellt neigen prozentuale Fehlermaße bei niedrigen Beobachtungswerten zu Verzerrungen. Dies kann im vorliegenden Experiment insbesondere bei Artikeln mit geringem durchschnittlichem Absatz pro Beobachtungsperiode zu abweichenden Ergebnissen führen.

² Auf eine separate Abbildung des Rangs des *mwMAE* wird verzichtet, da durch die konstante Anzahl von Zeitreihen je Verfahren und Absatzstelle keine abweichenden Ergebnisse zur Rangbildung nach dem *sum-MAE* entstehen, so dass beide rangbasierten Gütemaße identisch sind.

Tabelle 6.17: Summe, Mittelwert und Rang des *MAE* je Absatzstelle

<i>sumMAE</i> je Absatzstelle				
Verfahren	AU-1	AU-2	IN-1	IN-2
NF1	(28,76 / 32,81) 26,38	(31,48 / 35,97) 31,86	(122,23 81,78) 86,23	(67,26 / 64,39) 50,97
NF2	(28,74 / 31,36) 26,95	(31,75 / 37,39) 30,24	(111,32 / 72,53) 89,11	(70,81 / 64,89) 49,61
MA	(24,20 / 27,61) 23,86	(27,09 / 30,08) 25,90	(90,35 / 64,11) 60,89	(56,08 / 51,11) 46,96
S.ES	(23,80 / 26,39) 23,10	(25,62 / 29,14) 24,38	(84,42 / 61,89) 58,31	(53,75 / 48,25) 42,02
DT.ES	(21,97 / 25,42) 25,33	(24,49 / 27,19) 25,24	(77,74 / 57,44) 70,30	(50,86 / 47,14) 51,46
EXP.ES	(22,76 / 25,14) 26,00	(24,41 / 27,75) 24,43	(78,43 / 65,00) 67,51	(50,78 / 47,56) 50,95
EXP.ARIMA	(22,85 / 25,03) 27,24	(24,58 / 27,97) 24,33	(75,41 / 60,83) 56,04	(52,38 / 49,22) 45,41
EXP.ES&ARIMA	(22,98 / 24,97) 25,62	(24,63 / 27,89) 23,43	(79,31 / 63,92) 65,99	(52,24 / 48,14) 51,90
mv.EXP.ES	(21,82 / 24,64) 22,43	(23,63 / 26,33) 23,76	(78,50 / 63,33) 65,13	(41,13 / 45,56) 49,19
MLP	(21,20 / 22,50) 22,81	(22,87 / 25,81) 22,71	(73,11 / 53,89) 58,10	(52,49 / 44,17) 45,68
mv.MLP	(20,72 / 22,42) 22,57	(21,98 / 24,94) 23,10	(58,09 / 45,75) 50,43	(38,91 / 44,17) 41,52
<i>mwMAE</i> je Absatzstelle				
Verfahren	AU-1	AU-2	IN-1	IN-2
NF1	(2,40 / 2,73) 2,20	(2,42 / 2,77) 2,45	(12,22 / 8,18) 8,62	(7,47 / 7,15) 5,66
NF2	(2,40 / 2,61) 2,25	(2,44 / 2,88) 2,33	(11,13 / 7,25) 8,91	(7,87 / 7,21) 5,51
MA	(2,02 / 2,30) 1,99	(2,08 / 2,31) 1,99	(9,03 / 6,41) 6,09	(6,23 / 5,68) 5,22
S.ES	(1,98 / 2,20) 1,92	(1,97 / 2,24) 1,88	(8,44 / 6,19) 5,83	(5,97 / 5,36) 4,67
DT.ES	(1,83 / 2,12) 2,11	(1,88 / 2,09) 1,94	(7,77 / 5,74) 7,03	(5,65 / 5,24) 5,72
EXP.ES	(1,90 / 2,09) 2,17	(1,88 / 2,13) 1,88	(7,84 / 6,50) 6,75	(5,64 / 5,28) 5,66
EXP.ARIMA	(1,90 / 2,09) 2,27	(1,89 / 2,15) 1,87	(7,54 / 6,08) 5,60	(5,82 / 5,47) 5,05
EXP.ES&ARIMA	(1,91 / 2,08) 2,13	(1,89 / 2,15) 1,80	(7,93 / 6,39) 6,60	(5,80 / 5,35) 5,77
mv.EXP.ES	(1,82 / 2,05) 1,87	(1,82 / 2,03) 1,83	(7,85 / 6,33) 6,51	(4,57 / 5,06) 5,47
MLP	(1,77 / 1,88) 1,90	(1,76 / 1,99) 1,75	(7,31 / 5,39) 5,81	(5,83 / 4,91) 5,08
mv.MLP	(1,73 / 1,87) 1,88	(1,69 / 1,92) 1,78	(5,81 / 4,58) 5,04	(4,32 / 4,91) 4,61
<i>Rang sumMAE</i> und <i>Rang mwMAE</i> je Absatzstelle				
Verfahren	AU-1	AU-2	IN-1	IN-2
NF1	(11 / 11) 9	(10 / 10) 11	(11 / 11) 10	(10 / 10) 9
NF2	(10 / 10) 10	(11 / 11) 10	(10 / 10) 11	(11 / 11) 7
MA	(9 / 9) 5	(9 / 9) 9	(9 / 8) 5	(9 / 9) 5
S.ES	(8 / 8) 4	(8 / 8) 6	(8 / 5) 4	(8 / 7) 2
DT.ES	(4 / 7) 6	(5 / 4) 8	(4 / 3) 9	(4 / 4) 10
EXP.ES	(5 / 6) 8	(4 / 5) 7	(5 / 9) 8	(3 / 5) 8
EXP.ARIMA	(6 / 5) 11	(6 / 7) 5	(3 / 4) 2	(6 / 8) 3
EXP.ES&ARIMA	(7 / 4) 7	(7 / 6) 3	(7 / 7) 7	(5 / 6) 11
mv.EXP.ES	(3 / 3) 1	(3 / 3) 4	(6 / 6) 6	(2 / 3) 6
MLP	(2 / 2) 3	(2 / 2) 1	(2 / 2) 3	(7 / 1) 4
mv.MLP	(1 / 1) 2	(1 / 1) 2	(1 / 1) 1	(1 / 1) 1

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

In Abbildung 6.18 erfolgt eine grafische Darstellung der kumulierten *mwMAE* je Prognoseverfahren und Absatzstelle, sortiert nach absteigendem Rang der Verfahren nach *mwMAE* über alle Zeitreihen und Absatzstellen der Testmenge.¹

¹ Die Darstellung entspricht Abbildung 6.14 unter zusätzlicher Abbildung der Fehleraufteilung auf die einzelnen Absatzstellen. Die Anordnung nach absteigendem Rang oder aufsteigendem Prognosefehler *mwMAE* zeigt ebenfalls das mv.MLP als bestes Prognoseverfahren.

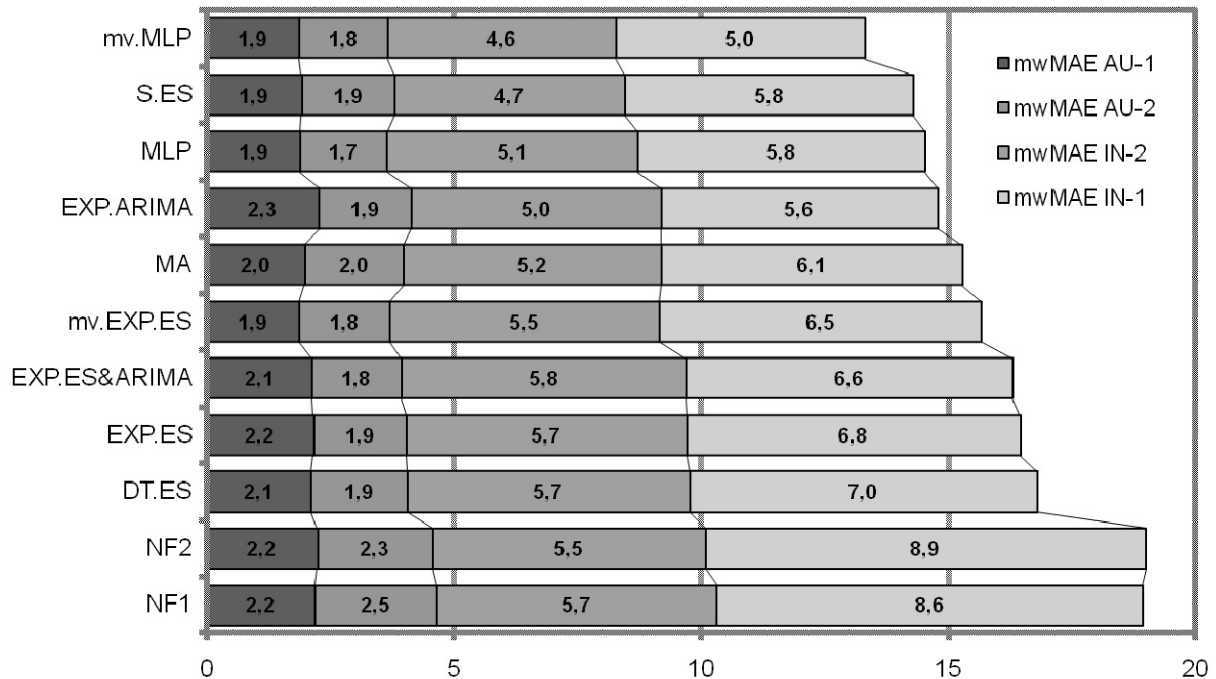


Abbildung 6.18: Kumulierte $mwMAE$ je Absatzstelle AU-1, AU-2, IN-1 und IN-2 auf der Testmenge

Über alle Absatzstellen entsprechen die Rangordnung und die kumulierten $mwMAE$ in Abbildung 6.18 der Analyse über alle Zeitreihen. Wie dort zeigt das mv.MLP auch auf den Absatzstellen der Innenautomaten IN-1 und IN-2 mit höherem Umsatz konsistent die höchste Prognosegenauigkeit auf Trainings-, Validierungs- und Testmenge. Dagegen erzielt mav.MLP auf den beiden Absatzstellen der Außenautomaten AU-1 und AU-2 lediglich einen zweiten Rang, nach dem konkurrierenden Verfahren mv.EXP.ES und vor dem univariaten MLP. Darüber hinaus lässt sich wie zuvor aufgrund der Heterogenität der Zeitreihen und Absatzstellen keine eindeutige Überlegenheit einzelner Verfahren oder Verfahrensgruppen je Absatzstelle oder Klassen von Absatzstellen erkennen. Die neuronalen Verfahren MLP und mv.MLP gehören konstant über alle Absatzstellen zu den vier genauesten Verfahren, während die Naiven Benchmarks NF1 und NF2 erneut die geringste Prognosegüte liefern.

Sowohl die Summe als auch die Mittelwerte der Prognosefehler des MAE in Tabelle 6.17 zeigen deutliche Unterschiede zwischen Innen- und Außenautomaten.¹ Der $mwMAE$ aller Prognoseverfahren an den Außenautomaten beträgt nur 2,01 Verkaufseinheiten im Vergleich zum Mittelwert an Innenautomaten von 5,96 Verkaufseinheiten. Dies beruht auf der unterschiedlichen Verteilung von schnelldrehenden A-Artikeln und langsamdrehenden B- und C-Artikeln mit ihren unterschiedlichen Umsatzmengen an Innen- und Außenautomaten (siehe Tabelle 6.9 in Abschnitt 6.2.2.2.1). Aufgrund des höheren Artikelumsatzes der Innenautomaten ist ein höherer absoluter Prognosefehler zu erwarten. Wird der $mwMAE$ von 2,0 an Außenautomaten in Relation zum durchschnittlichen Absatz von 4,3 Schachteln gesetzt, ergibt sich ein relativer Fehler von 46,7%. Im Gegensatz dazu führt der $mwMAE$ von 5,96 in Relation zum durchschnittlichen Umsatz von 17,6 Schachteln an Innenautomaten zu einem relativen Fehler von nur 33,8%. Unter Berücksichtigung der um 67% nahezu konstanten Variationskoeffizienten (siehe Tabelle 6.8) deutet die verringerte Prognosegenauigkeit der

¹ Während die Summe eine direkte Einschätzung der Höhe der Prognosefehler je Abverkaufsstelle zulässt, zeigen die Mittelwerte die relative Güte der Prognoseverfahren je Zeitreihe einer jeden Absatzstelle.

Verfahren einschließlich mv.MLP und MLP an Außenautomaten auf die Herausforderungen in der Prognose von Waren mit langsamem Artikelumschlag hin.¹

Während die Analyse der Mittelwerte und Rangfolgen nach Absatzstellen trotz geringfügiger Schwankungen des jeweils besten Verfahrens eine konsistent hohe Prognosegenauigkeit und einen hohen Rang der mv.MLP und des MLP zeigt, wird dies nicht durch die medianbasierten Mittelwerte und Rangordnungen des *MAE* in Tabelle 6.18 bestätigt.²

Tabelle 6.18: *mdMAE* und *Rang mdMAE* nach Absatzstellen

Verfahren	<i>mdMAE</i> je Absatzstelle			
	AU-1	AU-2	IN-1	IN-2
NF1	(2,31 / 2,38) 2,07	(2,63 / 2,89) 2,29	(6,82 / 4,56) 5,90	(3,43 / 7,92) 5,14
NF2	(2,22 / 2,32) 2,10	(2,56 / 3,08) 2,52	(6,10 / 5,07) 6,55	(3,56 / 3,53) 4,90
MA	(1,94 / 2,24) 1,83	(2,19 / 2,53) 2,05	(5,28 / 4,01) 4,67	(2,69 / 2,83) 3,86
S.ES	(1,86 / 2,03) 1,81	(2,07 / 2,36) 1,76	(5,10 / 3,97) 4,48	(2,67 / 2,86) 3,67
DT.ES	(1,72 / 1,93) 1,74	(1,99 / 2,14) 1,90	(4,92 / 4,13) 5,45	(2,70 / 2,75) 3,48
EXP.ES	(1,75 / 1,94) 1,88	(1,99 / 2,14) 1,90	(4,94 / 4,01) 5,24	(2,59 / 2,89) 3,24
EXP.ARIMA	(1,74 / 1,92) 2,43	(1,97 / 2,36) 1,90	(4,58 / 3,79) 5,17	(2,60 / 2,89) 3,24
EXP.ES&ARIMA	(1,75 / 1,94) 1,88	(1,99 / 2,14) 1,86	(5,01 / 4,17) 5,69	(2,59 / 2,89) 3,24
mv.EXP.ES	(1,75 / 1,89) 1,55	(1,84 / 2,14) 1,90	(4,68 / 4,17) 5,69	(2,49 / 2,78) 3,29
MLP	(1,78 / 1,88) 1,67	(1,82 / 2,22) 1,86	(4,03 / 3,81) 4,31	(2,72 / 2,81) 4,29
mv.MLP	(1,80 / 1,82) 1,71	(1,70 / 1,97) 1,81	(3,26 / 3,61) 4,45	(2,77 / 2,78) 3,76
Verfahren	<i>Rang mdMAE</i> je Absatzstelle			
	AU-1	AU-2	IN-1	IN-2
NF1	(11 / 11) 9	(11 / 10) 10	(11 / 10) 10	(10 / 11) 11
NF2	(10 / 10) 10	(10 / 11) 11	(10 / 11) 11	(11 / 10) 10
MA	(9 / 9) 6	(9 / 9) 9	(9 / 6) 4	(6 / 5) 8
S.ES	(8 / 8) 5	(8 / 7) 1	(8 / 4) 3	(5 / 6) 6
DT.ES	(1 / 5) 4	(5 / 2) 5	(5 / 7) 7	(7 / 1) 5
EXP.ES	(4 / 6) 7	(5 / 2) 5	(6 / 5) 6	(2 / 7) 1
EXP.ARIMA	(2 / 4) 11	(4 / 7) 5	(3 / 2) 5	(4 / 7) 1
EXP.ES&ARIMA	(4 / 6) 7	(5 / 2) 3	(7 / 8) 8	(2 / 7) 1
mv.EXP.ES	(3 / 3) 1	(3 / 2) 5	(4 / 8) 8	(1 / 2) 4
MLP	(6 / 2) 2	(2 / 6) 3	(2 / 3) 1	(8 / 4) 9
mv.MLP	(7 / 1) 3	(1 / 1) 2	(1 / 1) 2	(9 / 2) 7

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

Die Verfahren der NN verlieren unter Beurteilung durch *mdMAE* an Prognosegenauigkeit zugunsten anderer Prognoseverfahren aus unterschiedlichen Verfahrensklassen. Insbesondere für die Absatzstelle der Innenautomaten IN-2 mit überwiegend schnelldrehenden Waren und vielen Nullmengen (vgl. Tabelle 6.8) zeigen die NN eine geringe Prognosegüte auf den unteren Rängen aller Verfahren. Dagegen zeigen mv.MLP und MLP auf den anderen Absatzstellen AU-1, AU-2 und IN-1 konsistent eine führende Prognosegüte als eines der besten drei Verfahren. Die sonst robuste Prognosegüte der NN über unterschiedliche Absatzstellen wird ebenfalls durch die ordinalen Gütemaße des summierten Rangs in Tabelle 6.19 bestätigt.³

¹ Die Probleme in der Prognose von Langsamdreherern wird nachfolgend in der Untersuchung der unterschiedlichen Zeitreihenstrukturen in Abschnitt 6.3.1.4 erneut aufgegriffen.

² Die Ergebnisse je Absatzstelle nach dem Fehlermaß des *MAPE* bestätigen die überlegene Leistung von neuronalen Verfahren. Sie werden aufgrund ihrer geringeren Aussagekraft im Vergleich zum *MAE* und aus Platzgründen nachfolgend nicht weiter analysiert und diskutiert.

³ Da die Mittelwerte der Ränge eine identische Rangreihung liefern und die Mediane der Summe der Ränge nach *MAE* oder *MAPE* diese Ergebnisse bestätigen, wird auf ihre Darstellung aus Platzgründen verzichtet.

Tabelle 6.19: Summen des Rang $mwMAE$ und entsprechende Rangordnung je Absatzstelle

Verfahren	Summe des Rang $mwMAE$ je Absatzstelle			
	AU-1	AU-2	IN-1	IN-2
NF1	(121 / 126) 84	(136 / 121) 115	(110 / 98) 89	(86 / 91) 80
NF2	(126 / 111) 89	(135 / 135) 113	(91 / 94) 95	(97 / 94) 58
MA	(89 / 102) 53	(106 / 103) 98	(79 / 57) 52	(69 / 67) 52
S.ES	(85 / 85) 44	(80 / 89) 73	(75 / 50) 47	(59 / 56) 37
DT.ES	(46 / 63) 68	(51 / 54) 49	(39 / 45) 50	(46 / 36) 53
EXP.ES	(55 / 54) 71	(54 / 55) 52	(44 / 61) 48	(41 / 42) 38
EXP.ARIMA	(54 / 50) 86	(55 / 65) 59	(46 / 55) 51	(39 / 55) 45
EXP.ES&ARIMA	(58 / 47) 65	(57 / 60) 40	(50 / 63) 51	(44 / 46) 43
mv.EXP.ES	(44 / 45) 43	(49 / 45) 48	(46 / 58) 49	(28 / 30) 48
MLP	(37 / 19) 43	(56 / 37) 45	(46 / 34) 49	(43 / 17) 51
mv.MLP	(33 / 23) 43	(30 / 23) 46	(17 / 20) 32	(20 / 21) 36
Verfahren	Rang der Summe des Rang $mwMAE$ je Absatzstelle			
	AU-1	AU-2	IN-1	IN-2
NF1	(10 / 11) 9	(11 / 10) 11	(11 / 11) 10	(10 / 10) 11
NF2	(11 / 10) 11	(10 / 11) 10	(10 / 10) 11	(11 / 11) 10
MA	(9 / 9) 5	(9 / 9) 9	(9 / 6) 9	(9 / 9) 8
S.ES	(8 / 8) 4	(8 / 8) 8	(8 / 4) 2	(8 / 8) 2
DT.ES	(4 / 7) 7	(3 / 4) 5	(2 / 3) 6	(7 / 4) 9
EXP.ES	(6 / 6) 8	(4 / 5) 6	(3 / 8) 3	(4 / 5) 3
EXP.ARIMA	(5 / 5) 10	(5 / 7) 7	(4 / 5) 7	(3 / 7) 5
EXP.ES&ARIMA	(7 / 4) 6	(7 / 6) 1	(7 / 9) 7	(6 / 6) 4
mv.EXP.ES	(3 / 3) 1	(2 / 3) 4	(4 / 7) 4	(2 / 3) 6
MLP	(2 / 1) 1	(6 / 2) 2	(4 / 2) 4	(5 / 1) 7
mv.MLP	(1 / 2) 1	(1 / 1) 3	(1 / 1) 1	(1 / 2) 1

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

Aus Tabelle 6.19 geht hervor, dass lediglich das mv.MLP konsistent eine hohe Prognosegenauigkeit und einen niedrigen Rang der Verfahren erreicht und die Naiven Verfahren NF1 und NF2 einen konstant hohen Rang erreichen. Dagegen zeigen unterschiedliche Verfahren der MLP, mv.EXP.ES, S.ES und EXP.ES&ARIMA auf unterschiedlichen Absatzstellen eine hohe Genauigkeit und einen entsprechend niedrigen Rang. Trotz der insgesamt führenden Genauigkeit des Prognoseverfahrens mv.MLP kann jedoch keine reine Zustandsdominanz als bestes Verfahren über alle Absatzstellen festgestellt werden.¹

Wie zuvor stellt sich daher je Abverkaufsstelle die Frage, ob ein Verfahren mit hoher Prognosegenauigkeit ex ante anhand der Prognosegüte auf der Kalibrationsstichprobe identifiziert werden kann, um somit durch eine absatzstellenspezifische Verfahrensselektion eine gesteigerte Prognosegüte zu erzielen. Hierfür wird erneut die Korrelation zwischen den Fehlermaßen auf Trainings-, Validierungs- und Testmenge analysiert, die ebenfalls die Robustheit der Prognosegenauigkeit je Abverkaufsstelle nach $mwMAE$ und Rang $mwMAE$ verifiziert, und in Tabelle 6.20 dargestellt.²

¹ Es ist anzumerken, dass dies nach Erkenntnissen der M-Competitions und anderer Verfahrensvergleiche auch nicht zu erwarten war. Vielmehr ist anzunehmen, dass einzelne Verfahren aufgrund der Schwankungen der Zeitreihen „zufällig“ auf einzelnen Zeitreihen oder Absatzstellen besser abschneiden können.

² Liegt eine hohe, statistisch signifikante Korrelation zwischen Validierungs- und Testfehlern vor, so ließe dies eine valide und reliable Auswahl eines guten Prognoseverfahrens je Absatzstelle ex ante zu, somit auch eine anschließende Verringerung des Gesamtfehlers.

Tabelle 6.20: Korrelationen der Trainings-, Validierungs- und Testfehler der *mwMAE* nach Absatzstelle

POS	Datensatz	PEARSON-Korrelation des <i>mwMAE</i>			SPEARMAN-Korrelation des <i>Rang mwMAE</i>		
		Train	Valid	Test	Train	Valid	Test
AU-1	Train	1,000	0,979**	0,587	1,000	0,909**	0,636*
	Valid	0,979**	1,000	0,561	0,909**	1,000	0,582
	Test	0,587	0,561	1,000	0,636*	0,582	1,000
AU-2	Train	1,000	0,994**	0,962**	1,000	0,982**	0,800**
	Valid	0,994**	1,000	0,953**	0,982**	1,000	0,809**
	Test	0,962**	0,953**	1,000	0,800**	0,809**	1,000
IN-1	Train	1,000	0,937**	0,882**	1,000	0,873**	0,682*
	Valid	0,937**	1,000	0,841**	0,873**	1,000	0,755**
	Test	0,882**	0,841**	1,000	0,682*	0,755**	1,000
IN-2	Train	1,000	0,905**	0,363	1,000	0,811**	0,091
	Valid	0,905**	1,000	0,368	0,811**	1,000	0,232
	Test	0,363	0,368	1,000	0,091	0,232	1,000

* Korrelation ist signifikant bei einem 0.05 Konfidenzniveau (2-seitiger Test, Berechnung mit SPSS 13.0).

** Korrelation ist hoch-signifikant bei einem 0.01 Konfidenzniveau (2-seitiger Test, Berechnung mit SPSS 13.0).

Die Analyse der Korrelationskoeffizienten in Tabelle 6.20 zeigt insignifikante Korrelationen nach PEARSON sowie SPEARMAN für die Absatzstellen AU-1 und IN-2. Somit kann für diese Absatzstellen keine valide Ex-ante-Auswahl eines Prognoseverfahrens nach ihrem *mwMAE* oder *Rang mwMAE* auf der Validierungsmenge erfolgen. Somit ist auch durch eine absatzstellenspezifische Selektion von Prognoseverfahren nicht immer eine systematische Verbesserung der Prognosegüte zu erwarten. Da aus einer Unterscheidung nach Absatzstellen keine verlässlichen Hinweise auf die Bedingungen abzuleiten sind, in denen NN anderen Prognoseverfahren unterlegen oder überlegen sind, wird die Analyse auf einzelne Zeitreihen und homogene Gruppen von Zeitreihen erweitert.

6.3.1.3 Ergebnisse nach Zeitreihe

Nachfolgend wird die relative Prognosegüte der Prognoseverfahren je Zeitreihe analysiert. Da eine Ausgabe aller Fehlermaße des *MAE*, *MAPE* und ihrer Rangordnung zu umfangreich ist, erfolgt lediglich eine grafische Abbildung der Prognosefehler der *MAE* je Zeitreihe in Kivita-graphen der univariaten und der multivariaten Prognoseverfahren. Eine vollständige Darstellung der numerischen Prognosefehler des *MAE*, *MAPE*, *Rang MAE* und *Rang MAPE* je Zeitreihe findet sich im Anhang IX.E.i.

Abbildung 6.20 vergleicht zunächst die multivariaten Prognoseverfahren des mv.MLP und des mv.EXP.ES für alle Zeitreihen in zwei Kivita-graphen.¹ Kivita-graph (a.) zeigt die Prognosefehler des *MAE* für alle Zeitreihen der Außenautomaten in Absatzstelle AU-1 und AU-2, Kivita-graph (b.) die *MAE* für alle Zeitreihen der Innenautomaten IN-1 und IN-2.

¹ Ein Kivita-graph zeigt die Fehlerprofile der Prognoseverfahren als Linien an, mit den Speichen des Rades als Achsen der einzelnen Zeitreihen und einem Fehler von Null im Zentrum des Kivita-graphen. Näher am Ursprung liegende Linien zeigen somit Verfahren mit einem geringeren *MAE* auf der Testmenge an. Ein Verfahren mit einem konsistenten Verlauf innerhalb eines anderen Profils über alle Zeitreihen dominiert somit ein anderes Verfahren durch niedrigere Fehler.

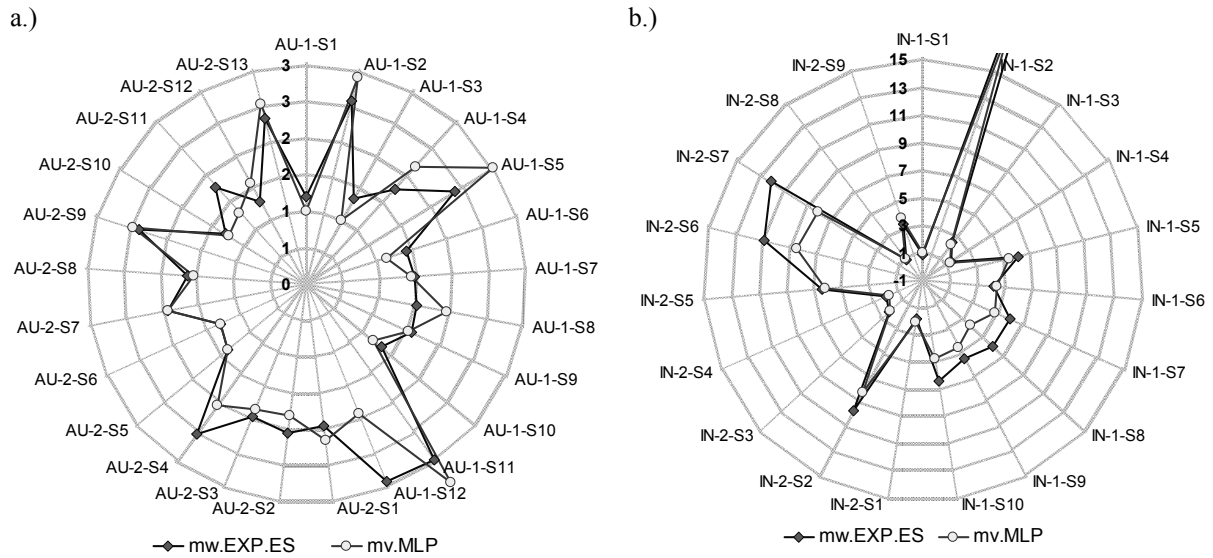


Abbildung 6.19: Prognosegüte der multivariaten Verfahren mv.MLP und mv.EXP.ES nach MAE je Zeitreihe auf der Testmenge der Außenautomaten (a.) und der Innenautomaten (b.)

Abbildung 6.19 verdeutlicht die homogene Leistung der zwei multivariaten Verfahren, wobei das mv.MLP auf der Mehrheit der 44 Zeitreihen sowohl für Innenautomaten als auch für Außenautomaten einen geringeren Prognosefehler aufweist als das mv.EXP.ES. Zudem ist anzumerken, dass das mv.EXP.ES für jede Zeitreihe und für jeden Prognoseursprung ein anderes Prognoseverfahren selektiert und parametrisiert, während das mv.MLP eine konstante Architektur ohne Modellselektion für jeden Prognoseursprung verwendet.

Abbildung 6.20 zeigt den korrespondierenden Vergleich der univariaten Verfahren des MLP gegen eine relevante Auswahl der statistischen Prognoseverfahren in zwei Kivitagraphen getrennt nach Innen- und Außenautomaten.

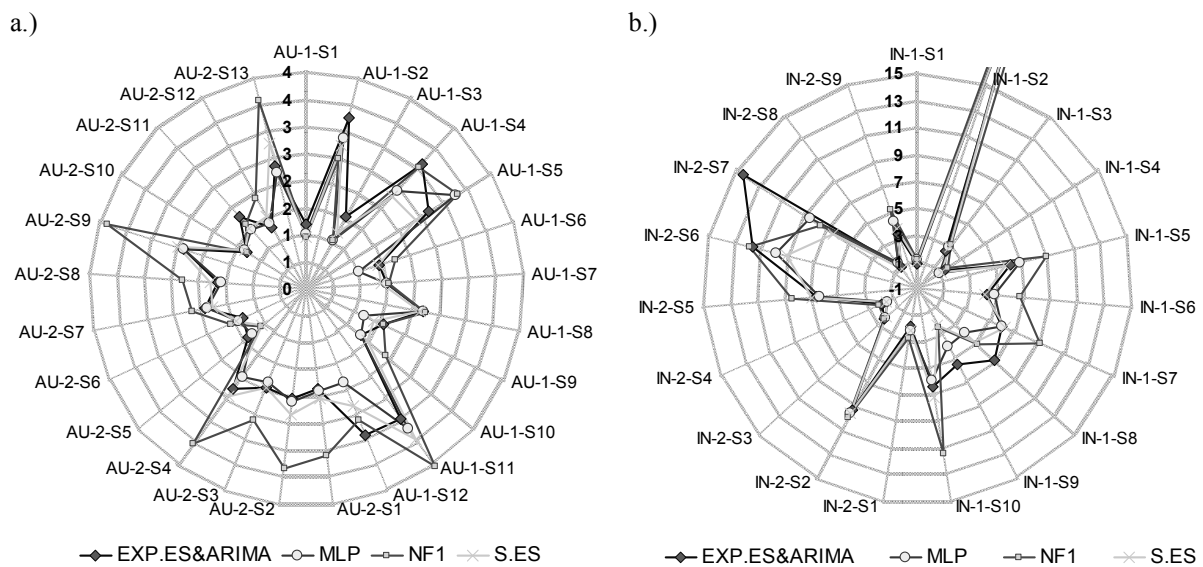


Abbildung 6.20: Prognosegüte der univariaten Verfahren MLP und ausgewählter statistischer Verfahren nach MAE je Zeitreihe an Außenautomaten (a.) und Innenautomaten (b.)

Die Abbildung zeigt die niedrigen Prognosefehler des univariaten MLP auf der Mehrheit der Zeitreihen im Vergleich zu anderen Prognoseverfahren. Die Abstände der Fehlerwerte sind jedoch geringer als im Vergleich multivariater Verfahren, was die Probleme in der validen

Unterscheidung der Prognoseverfahren und ihrer Rangreihung erklärt. Die Abbildung verdeutlicht ferner die Komplexität und Heterogenität eines Verfahrensvergleiches über mehrere Zeitreihen und den beschränkten Erkenntniszuwachs in der Analyse der Fehlerwerte in diesem Grad der Detaillierung. Für einzelne Zeitreihen ergibt sich jeweils eine klare Dominanz einzelner, jedoch wechselnder Verfahren, während für andere keine valide Rangreihung erfolgen kann. Die Ergebnisse bestätigen somit den vorliegenden Versuchsaufbau der Prognose unter Verwendung einer Vielzahl von Zeitreihen und Zeitreihenursprüngen um die Ermittlung generalisierbarer Erkenntnisse sicherzustellen. Die Untersuchung wird nachfolgend auf Kategorien der Zeitreihen erweitert.¹

6.3.1.4 Ergebnisse nach Zeitreihencharakteristika

In Ergänzung zur Analyse über alle Zeitreihen und Absatzstellen wird die Prognosegüte der Verfahren geordnet nach den warenlogistischen Kriterien der Anwendungsdomäne hinsichtlich der Struktur des Zeitreihenverlaufs, der Geschwindigkeit des Warenumschlages und der Bedeutung der Ware im Sortiment analysiert.²

Tabelle 6.21 zeigt die Ergebnisse aller 44 Zeitreihen kategorisiert nach der Struktur des Absatzverlaufs (siehe Abschnitt 6.2.2) für Zeitreihen mit regelmäßigen, unregelmäßigen und sporadischen Verlaufsformen. Die Ergebnisse bestätigen die konstant hohe Prognosegenauigkeit der NN über alle Verlaufsformen der Zeitreihen. So erzielt das mv.MLP als bestes Prognoseverfahren erneut den niedrigsten Prognosefehler auf den Testdaten aller sporadischen sowie der regelmäßigen Zeitreihen und einen um 0,06 Verpackungseinheiten höheren Prognosefehler als das S.ES mit einem zweiten Rang auf den unregelmäßigen Zeitreihen. Das univariate MLP zeigt ebenfalls eine hohe Genauigkeit über unterschiedliche Zeitreihenformen hinweg, mit einem zweiten Rang hinter dem mv.MLP auf sporadischen und regelmäßigen Zeitreihen und einem vierten Platz auf den unregelmäßigen Zeitreihen. Die Analyse der median basierten Gütemaße bestätigt ihre robuste Leistung.

¹ Eine Korrelationsanalyse der Verfahren je Zeitreihe zeigt eine weitere Verringerung der statistischen Signifikanz zwischen den Fehlern der Validierungs- und Testmengen als in Tabelle 6.15 und Tabelle 6.20. Ihre Abbildung entfällt hier, da die Abbildung über alle Zeitreihen nur wenig Informationszugewinn bietet.

² Aufgrund möglicher Verzerrungen der Analyse durch das Fehlermaß des *MAPE* bei Artikeln mit langsamem Warenumschlag sowie mit sporadischem Verlauf wird die Analyse auf die Fehlermaße des *mwMAE* und des *mdMAE* beschränkt. Die Zunahme der absoluten Fehlerwerte *mwMAE* sowie *mdMAE* von den sporadischen bis hin zu den regelmäßigen Zeitreihen ist dennoch im Zusammenhang mit den steigenden Absatzzahlen zu betrachten. Der durchschnittliche *mwMAE* von 2,22 Verpackungseinheiten über alle Prognoseverfahren auf sporadischen Zeitreihen ist geringer als der *mwMAE* für Zeitreihen mit unregelmäßigen Verläufen von 2,89 und von regelmäßigen Zeitreihenverläufen mit 4,97. Durch die steigende Absatzhöhe steigen die relativen Fehler je Zeitreihenverlauf tatsächlich nur unterproportional an. So entspricht der absolute Fehler auf sporadischen Zeitreihen einem relativen Fehler von 88,8%, der unregelmäßiger Verläufe 56,6% und der regelmäßiger Verläufe nur 27,4%.

Tabelle 6.21: *mwMAE* und *mdMAE* über alle Zeitreihen nach Absatzverlauf

<i>mwMAE</i>			
Verfahren	Sporadisch	Unregelmäßig	Regelmäßig
NF1	(2,67 / 1,82) 2,12	(4,74 / 3,93) 3,30	(7,34 / 6,58) 6,19
NF2	(2,41 / 1,72) 2,40	(4,57 / 3,87) 3,34	(7,21 / 6,21) 6,11
MA	(2,11 / 1,56) 2,14	(3,78 / 3,17) 2,61	(5,78 / 5,27) 4,96
S.ES	(2,16 / 1,66) 2,04	(3,60 / 3,08) 2,50	(5,43 / 4,97) 4,58
DT.ES	(2,06 / 1,42) 2,39	(3,36 / 2,95) 2,99	(5,06 / 4,71) 5,26
EXP.ES	(2,10 / 1,56) 2,60	(3,45 / 3,03) 3,03	(5,03 / 5,04) 4,99
EXP.ARIMA	(1,99 / 1,31) 2,26	(3,41 / 3,16) 2,80	(5,04 / 4,82) 4,48
EXP.ES&ARIMA	(2,16 / 1,57) 2,54	(3,49 / 3,06) 3,06	(5,12 / 4,96) 4,88
mv.EXP.ES	(1,95 / 1,42) 2,32	(3,35 / 3,05) 2,87	(4,57 / 4,74) 4,79
MLP	(1,68 / 1,51) 1,82	(2,79 / 2,79) 2,76	(5,49 / 4,30) 4,43
mv.MLP	(1,57 / 1,67) 1,74	(2,42 / 2,53) 2,56	(4,35 / 4,08) 4,05
<i>Rang nach mwMAE</i>			
Verfahren	Rang sporadisch	Rang unregelmäßig	Rang regelmäßig
NF1	(11 / 11) 4	(11 / 11) 10	(11 / 11) 11
NF2	(10 / 10) 9	(10 / 10) 11	(10 / 10) 10
MA	(7 / 6) 5	(9 / 9) 3	(9 / 9) 7
S.ES	(9 / 8) 3	(8 / 7) 1	(7 / 7) 4
DT.ES	(5 / 2) 8	(4 / 3) 7	(5 / 3) 9
EXP.ES	(6 / 5) 11	(6 / 4) 8	(3 / 8) 8
EXP.ARIMA	(4 / 1) 6	(5 / 8) 5	(4 / 5) 3
EXP.ES&ARIMA	(8 / 7) 10	(7 / 6) 9	(6 / 6) 6
mv.EXP.ES	(3 / 3) 7	(3 / 5) 6	(2 / 4) 5
MLP	(2 / 4) 2	(2 / 2) 4	(8 / 2) 2
mv.MLP	(1 / 9) 1	(1 / 1) 2	(1 / 1) 1
<i>mdMAE</i>			
Verfahren	Sporadisch	Unregelmäßig	Regelmäßig
NF1	(1,41 / 1,75) 1,19	(2,57 / 2,67) 2,29	(3,01 / 3,47) 3,29
NF2	(1,50 / 1,50) 1,26	(2,40 / 2,89) 2,24	(3,04 / 3,81) 3,19
MA	(1,21 / 1,47) 1,26	(2,12 / 2,28) 1,76	(2,50 / 2,97) 2,48
S.ES	(1,22 / 1,40) 1,14	(2,03 / 2,19) 1,76	(2,38 / 2,81) 2,43
DT.ES	(1,16 / 1,21) 1,24	(1,94 / 2,11) 1,62	(2,28 / 2,83) 2,95
EXP.ES	(1,13 / 1,21) 1,36	(1,97 / 2,08) 1,62	(2,32 / 2,81) 2,57
EXP.ARIMA	(1,17 / 1,17) 1,24	(1,80 / 2,08) 1,90	(2,38 / 2,83) 2,71
EXP.ES&ARIMA	(1,25 / 1,21) 1,24	(1,97 / 2,08) 1,62	(2,35 / 2,81) 2,38
mv.EXP.ES	(1,25 / 1,21) 1,24	(1,97 / 2,08) 1,62	(2,09 / 2,64) 2,38
MLP	(0,91 / 1,07) 1,24	(1,79 / 1,89) 1,57	(2,44 / 2,53) 2,24
mv.MLP	(0,95 / 1,03) 1,12	(1,70 / 1,81) 1,57	(2,28 / 2,47) 2,48
<i>Rang nach mdMAE</i>			
Verfahren	Rang sporadisch	Rang unregelmäßig	Rang regelmäßig
NF1	(10 / 11) 3	(11 / 10) 11	(10 / 10) 11
NF2	(11 / 10) 9	(10 / 11) 10	(11 / 11) 10
MA	(6 / 9) 9	(9 / 9) 7	(9 / 9) 5
S.ES	(7 / 8) 2	(8 / 8) 7	(7 / 4) 4
DT.ES	(4 / 4) 4	(4 / 7) 3	(2 / 7) 9
EXP.ES	(3 / 4) 11	(5 / 3) 3	(4 / 4) 7
EXP.ARIMA	(5 / 3) 4	(3 / 3) 9	(6 / 7) 8
EXP.ES&ARIMA	(8 / 4) 4	(5 / 3) 3	(5 / 4) 2
mv.EXP.ES	(8 / 4) 4	(5 / 3) 3	(1 / 3) 2
MLP	(1 / 2) 4	(2 / 2) 1	(8 / 2) 1
mv.MLP	(2 / 1) 1	(1 / 1) 1	(3 / 1) 5

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

Die Naiven Verfahren NF1 und NF2 zeigen bis auf die Testdaten der sporadischen Zeitreihen erneut die niedrigste Prognosegenauigkeit über alle Teildatenmengen. Dagegen kann für die nachfolgenden Prognoseverfahren erneut kein Verfahren und auch keine Verfahrensklasse eindeutig eine konsistent bessere Leistung zeigen. Erneut liefert die konstante Anwendung des Expertensystems EXP.ARIMA bessere Ergebnisse als EXP.ES&ARIMA, wenngleich S.ES und DT.ES noch bessere Ergebnisse erzielen. Dennoch zeigt sich kein konsistentes Bild, und zwar auf keinem der Zeitreihenverläufe. Selbst für sporadische Zeitreihen zeigen offensichtlich fehlspezifizierte Verfahren des MA oder DT.ES eine robuste Prognosegenauigkeit, während weder die Expertensysteme mit dedizierten Routinen zur Erkennung und Prognose sporadischer Zeitreihenverläufe durch CROSTONS Exponentielle Glättung noch ihre multivariaten Varianten eine signifikante Überlegenheit dokumentieren.¹

Alternativ werden in Tabelle 6.22 die Zeitreihen nach der Umschlagsgeschwindigkeit der Waren kategorisiert und in den zwei Klassen der Langsamdreher und Schnelldreher analysiert.

Tabelle 6.22: *mwMAE*, *mdMAE* und Rangordnung über alle Absatzstellen nach Umschlagsgeschwindigkeit

Verfahren	<i>mwMAE</i>		<i>mdMAE</i>	
	Langsamdreher	Schnelldreher	Langsamdreher	Schnelldreher
NF1	(3,21 / 3,08) 2,70	(16,76 / 12,99) 12,28	(2,68 / 2,85) 2,29	(11,95 / 10,76) 9,43
NF2	(3,19 / 3,05) 2,76	(15,98 / 12,06) 12,07	(2,63 / 2,92) 2,50	(12,01 / 10,99) 8,67
MA	(2,67 / 2,51) 2,29	(12,70 / 10,31) 9,38	(2,24 / 2,38) 2,19	(9,57 / 8,86) 7,52
S.ES	(2,58 / 2,43) 2,19	(11,85 / 9,79) 8,63	(2,14 / 2,24) 2,05	(9,20 / 8,47) 7,00
DT.ES	(2,44 / 2,30) 2,62	(10,89 / 9,32) 9,74	(2,03 / 2,15) 1,93	(8,33 / 8,24) 6,48
EXP.ES	(2,45 / 2,33) 2,59	(11,01 / 10,18) 9,45	(2,07 / 2,14) 2,00	(8,30 / 8,53) 6,64
EXP.ARIMA	(2,45 / 2,38) 2,44	(10,89 / 9,67) 8,14	(2,01 / 2,08) 2,00	(9,11 / 8,92) 6,36
EXP.ES&ARIMA	(2,49 / 2,34) 2,58	(11,20 / 10,07) 9,25	(2,07 / 2,14) 1,93	(9,11 / 8,89) 6,43
mv.EXP.ES	(2,34 / 2,28) 2,41	(10,11 / 9,74) 9,20	(1,99 / 2,14) 1,86	(6,53 / 8,40) 6,36
MLP	(2,20 / 2,16) 2,27	(11,31 / 8,59) 8,46	(2,00 / 2,03) 1,88	(9,24 / 7,92) 6,57
mv.MLP	(2,03 / 2,11) 2,23	(8,35 / 7,67) 7,18	(1,97 / 1,96) 1,90	(6,31 / 7,67) 5,76
	<i>Rang nach mwMAE</i>		<i>Rang nach mdMAE</i>	
	Langsamdreher	Schnelldreher	Langsamdreher	Schnelldreher
NF1	(11 / 11) 10	(11 / 11) 11	(11 / 10) 10	(10 / 10) 11
NF2	(10 / 10) 11	(10 / 10) 10	(10 / 11) 11	(11 / 11) 10
MA	(9 / 9) 4	(9 / 9) 7	(9 / 9) 9	(9 / 7) 9
S.ES	(8 / 8) 1	(8 / 6) 4	(8 / 8) 8	(7 / 5) 8
DT.ES	(4 / 4) 9	(4 / 3) 9	(5 / 7) 4	(4 / 3) 5
EXP.ES	(6 / 5) 8	(5 / 8) 8	(6 / 4) 6	(3 / 6) 7
EXP.ARIMA	(5 / 7) 6	(3 / 4) 2	(4 / 3) 6	(5 / 9) 2
EXP.ES&ARIMA	(7 / 6) 7	(6 / 7) 6	(6 / 4) 4	(5 / 8) 4
mv.EXP.ES	(3 / 3) 5	(2 / 5) 5	(2 / 4) 1	(2 / 4) 2
MLP	(2 / 2) 3	(7 / 2) 3	(3 / 2) 2	(8 / 2) 6
mv.MLP	(1 / 1) 2	(1 / 1) 1	(1 / 1) 3	(1 / 1) 1

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

Auch für unterschiedliche Umschlagsgeschwindigkeiten zeigen die Verfahren der NN erneut eine robuste Leistung über alle Warenkategorien als eines der drei besten Verfahren. Naive Verfahren weisen konsistent die niedrigste Prognosegüte auf. Innerhalb der Klasse der Mittelwerte zeigt die Exponentielle Glättung S.ES eine konsistent bessere Leistung als der

¹ Die verringerte Prognosegüte der Expertensysteme kann erneut als Indiz für mögliche Probleme in der validen und reliablen Identifikation von Zeitreihenverläufen und einer adäquaten Modellselektion dienen, obwohl die Erkennung sporadischer Zeitreihenverläufe aufgrund der klaren Definition von Nullperioden unproblematisch ist. Die Probleme in der Modellselektion erfordert somit weiterführende Untersuchungen für die Erstellung genauer Prognosen und somit der Bestimmung wirtschaftlicher Bestellmengen..

einfache gleitende Mittelwert MA sowie DT.ES. Multivariate Varianten der Verfahren zeigen eine bessere Prognosegüte als univariate, sowohl für NN durch mv.MLP als auch durch mv.EXP.ES im Vergleich zu den univariaten Expertensystemen EXP.ES, ARIMA und ES&ARIMA. Die Ergebnisse bestätigen weitgehend die bisherigen Erkenntnisse, wonach NN andere Verfahren auch bei den heterogenen Artikelklassen der Schnell- bzw. Langsamdreher dominieren, ohne relative Rangreihung der statistischen Verfahren oder Verfahrensklassen.

Abschließend werden die Zeitreihen nach der materialwirtschaftlichen ABC-Klassifikation gemäß ihrer Umsatzbedeutung für das Unternehmen kategorisiert und die Prognosefehler in Tabelle 6.23 dargestellt.¹ Auch hier zeigt sich das konstante Bild einer hohen Prognosegenauigkeit der Verfahren der NN. Sowohl mv.MLP als auch MLP zeigen als bestes sowie als zweit- bzw. drittbestes Verfahren eine führende Prognosegüte. Das Verfahren des mvMLP ist das beste multivariate Verfahren, während das MLP in zwei der drei Artikelklassen jeweils unterschiedlichen Prognoseverfahren des EXP.ARIMA und S.ES unterliegt. Erneut zeigt das S.ES eine konstant bessere Leistung als MA und komplexe Expertensysteme.

Als Konsequenz kann festgehalten werden, dass NN als univariate oder multivariate Prognoseverfahren unabhängig von Zeitreihenstruktur, Umschlagsgeschwindigkeit oder materialwirtschaftlichen Gesichtspunkten konsistent zu den genauesten Prognoseverfahren zählen. Sie zeigen eine höhere und robustere Prognosegüte als alle statistischen Verfahren ohne Modellselektion und als alle Expertensysteme mit individueller Verfahrensselektion, mit Ausnahme der einfachen Exponentiellen Glättung S.ES. Auch wenn keine grundsätzliche Überlegenheit eines einzelnen Prognoseverfahrens oder einer Verfahrensklasse auf der Grundlage einer Stichprobe von Zeitreihen festgestellt werden kann, dokumentieren die Ergebnisse eine hohe, robuste und wettbewerbsfähige Güte von NN zur Absatzprognose. Die Verfahren der univariaten MLP und der multivariaten mv.MLP zählen zu den genauesten Prognoseverfahren der Vergleichsstudie und befinden sich kontinuierlich auf den obersten drei Rängen über alle Fehlermaße. Es kann somit festgestellt werden, dass NN als Prognoseverfahren – auch ohne die Festlegung einer individuellen Netzwerkarchitektur je Zeitreihe² – nicht signifikant schlechter zur Absatzprognose geeignet sind als statistische Verfahren und Expertensysteme, die als empirische Benchmark-Verfahren etabliert sind. In der vorliegenden Anwendung zeigen NN vielmehr signifikant bessere Ergebnisse als alle konkurrierenden Verfahren. Dies rechtfertigt weiterführende Forschungsanstrengungen zur Entwicklung einer Methodologie der Modellierung von NN zur Absatzprognose. Ferner belegt die Güte von NN die Notwendigkeit weiterführender Verfahrensvergleiche in anderen Anwendungsdomänen, um die Performanz von NN in der Prognose systematisch zu analysieren und NN gegebenenfalls als reliable Prognoseverfahren zu etablieren.

¹ Die Kategorisierung entspricht einer Erweiterung der Unterscheidung in Langsamdreher und Schnelldreher, wobei die A-Artikel den Schnelldrehern entsprechen und B- und C-Artikel den Langsamdrehern (vgl. Abschnitt 6.2.2.2.3).

² Dazu ist festzustellen, dass in der vorliegenden Analyse eine einzelne Netzwerk-Topologie ohne Anpassung zum Beispiel der Inputneuronen an die Lag-Struktur der Zeitreihe bzw. der Anzahl versteckter Neuronen an die Komplexität der Zeitreihenmuster verwendet wurde. Es ist somit möglich, dass eine umfassende Methodologie zur validen und robusten Modellierung von MLPs und mv.MLPs zu einer weiteren Verbesserung der Prognosegenauigkeit führt.

Tabelle 6.23: *mwMAE* und *mdMAE* über alle Zeitreihen nach Artikelklassifikation

<i>mwMAE</i>			
Verfahren	A-Artikel	B-Artikel	C-Artikel
NF1	(16,76 / 12,99) 12,28	(4,26 / 3,99) 3,61	(2,37 / 2,36) 1,97
NF2	(15,98 / 12,06) 12,07	(4,28 / 4,19) 3,81	(2,31 / 2,13) 1,92
MA	(12,70 / 10,31) 9,38	(3,56 / 3,38) 3,10	(1,96 / 1,82) 1,65
S.ES	(11,85 / 9,79) 8,63	(3,40 / 3,26) 2,87	(1,92 / 1,76) 1,64
DT.ES	(10,89 / 9,32) 9,74	(3,20 / 3,11) 3,78	(1,84 / 1,65) 1,70
EXP.ES	(11,01 / 10,18) 9,45	(3,24 / 3,11) 3,73	(1,82 / 1,72) 1,68
EXP.ARIMA	(10,89 / 9,67) 8,14	(3,28 / 3,15) 3,30	(1,78 / 1,77) 1,76
EXP.ES&ARIMA	(11,20 / 10,07) 9,25	(3,27 / 3,12) 3,76	(1,86 / 1,73) 1,64
mv.EXP.ES	(10,11 / 9,74) 9,20	(2,95 / 2,98) 3,39	(1,85 / 1,71) 1,63
MLP	(11,31 / 8,59) 8,46	(2,89 / 2,91) 3,08	(1,65 / 1,56) 1,61
mv.MLP	(8,35 / 7,67) 7,18	(2,59 / 2,86) 3,08	(1,58 / 1,51) 1,54
<i>Rang nach mwMAE</i>			
Verfahren	Rang A-Artikel	Rang B-Artikel	Rang C-Artikel
NF1	(11 / 11) 11	(10 / 10) 7	(11 / 11) 11
NF2	(10 / 10) 10	(11 / 11) 11	(10 / 10) 10
MA	(9 / 9) 7	(9 / 9) 4	(9 / 9) 6
S.ES	(8 / 6) 4	(8 / 8) 1	(8 / 7) 5
DT.ES	(4 / 3) 9	(4 / 4) 10	(5 / 3) 8
EXP.ES	(5 / 8) 8	(5 / 5) 8	(4 / 5) 7
EXP.ARIMA	(3 / 4) 2	(7 / 7) 5	(3 / 8) 9
EXP.ES&ARIMA	(6 / 7) 6	(6 / 6) 9	(7 / 6) 4
mv.EXP.ES	(2 / 5) 5	(3 / 3) 6	(6 / 4) 3
MLP	(7 / 2) 3	(2 / 2) 3	(2 / 2) 2
mv.MLP	(1 / 1) 1	(1 / 1) 2	(1 / 1) 1
<i>mdMAE</i>			
Verfahren	A-Artikel	B-Artikel	C-Artikel
NF1	(11,95 / 10,76) 9,43	(3,07 / 3,71) 3,19	(1,93 / 2,10) 1,67
NF2	(12,01 / 10,99) 8,67	(3,15 / 3,88) 3,29	(2,14 / 2,17) 1,50
MA	(9,57 / 8,86) 7,52	(2,66 / 3,00) 2,55	(1,74 / 1,79) 1,45
S.ES	(9,20 / 8,47) 7,00	(2,49 / 3,04) 2,40	(1,69 / 1,78) 1,40
DT.ES	(8,33 / 8,24) 6,48	(2,39 / 2,83) 2,64	(1,67 / 1,61) 1,45
EXP.ES	(8,30 / 8,53) 6,64	(2,38 / 2,85) 2,45	(1,69 / 1,75) 1,50
EXP.ARIMA	(9,11 / 8,92) 6,36	(2,43 / 2,86) 2,79	(1,67 / 1,81) 1,55
EXP.ES&ARIMA	(9,11 / 8,89) 6,43	(2,42 / 2,85) 2,52	(1,69 / 1,75) 1,45
mv.EXP.ES	(6,53 / 8,40) 6,36	(2,21 / 2,64) 2,38	(1,69 / 1,75) 1,43
MLP	(9,24 / 7,92) 6,57	(2,39 / 2,53) 2,43	(1,48 / 1,51) 1,37
mv.MLP	(6,31 / 7,67) 5,76	(2,25 / 2,58) 2,50	(1,50 / 1,51) 1,36
<i>Rang nach mdMAE</i>			
Verfahren	Rang A-Artikel	Rang B-Artikel	Rang C-Artikel
NF1	(10 / 10) 11	(10 / 10) 10	(10 / 10) 11
NF2	(11 / 11) 10	(11 / 11) 11	(11 / 11) 8
MA	(9 / 7) 9	(9 / 8) 7	(9 / 8) 5
S.ES	(7 / 5) 8	(8 / 9) 2	(8 / 7) 3
DT.ES	(4 / 3) 5	(5 / 4) 8	(4 / 3) 5
EXP.ES	(3 / 6) 7	(3 / 5) 4	(5 / 4) 8
EXP.ARIMA	(5 / 9) 2	(7 / 7) 9	(3 / 9) 10
EXP.ES&ARIMA	(5 / 8) 4	(6 / 5) 6	(5 / 4) 5
mv.EXP.ES	(2 / 4) 2	(1 / 3) 1	(5 / 4) 4
MLP	(8 / 2) 6	(4 / 1) 3	(1 / 1) 2
mv.MLP	(1 / 1) 1	(2 / 2) 5	(2 / 1) 1

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

Darüber hinaus bestärken die Ergebnisse die gewählte Verfahrenskonfiguration der NN mit einer konstanten Architektur und lediglich einer zeitreihenspezifischen Parametrisierung. Die Eine geringe Korrelation zwischen Validierungs- und Testfehler kann die valide und reliable Selektion eines adäquaten Verfahrens bzw. Modellaufbaus erschweren oder sogar verhindern, wie anhand der hohen Variation der jeweils besten Verfahrensform für unterschiedliche Zeitreihen, der geringen Korrelationen zwischen Validierungs- und Testfehlern und der inkonsistenten Ergebnisse der Expertensysteme deutlich wurde. Somit können die Verfahren der NN, welche im Sinne einer simultanen Modellselektion und Modellparametrisierung die Struktur der Zeitreihe ohne explizite Modellselektion unmittelbar aus den Daten erlernen, potenzielle Fehler durch inkorrekte oder ineffiziente Modellselektion vermeiden und konsistent niedrigere Prognosefehler erzielen.¹

Dennoch erlaubt die hohe Prognosegüte kein abschließendes Urteil über die Fähigkeit einzelner Prognoseverfahren zur Unterstützung betriebswirtschaftlicher Fragestellungen wie der Warendisposition. Dies erfordert eine Bewertung der resultierenden Warenbestände durch jedes Prognoseverfahren, bewertet mit den Entscheidungskosten aus Unterbeständen sowie Überbeständen, was im nachfolgenden Abschnitt diskutiert wird.

6.3.2 Ergebnisse der Warendisposition

6.3.2.1 Gesamtergebnisse

Nachfolgend wird die Leistung aller Prognoseverfahren in der Entscheidungsunterstützung zur betrieblichen Warendisposition überprüft. Die Bedarfsprognosen aller Prognoseverfahren aus Abschnitt 6.3.1 bilden dabei die Grundlage zur Berechnung kostenminimaler Bestellmengen.²

Die Berechnung der Bestellmengen wird um ein weiteres Verfahren der NN erweitert, welches mit einer asymmetrischen Zielfunktion der tatsächlichen Entscheidungskosten der Warendisposition trainiert wird. Das so parametrisierte multivariate mv.MLP.ACF sagt ohne eine vorhergehende Bedarfsprognose unmittelbar eine kostenminimale Bestellmenge vorher. Die Ergebnisse des Verfahrensvergleichs zur Warendisposition nach den mittleren linearen Kosten $mwMLLC$ über alle Zeitreihen und Zeitreihenursprünge sind in Abbildung 6.21 dargestellt.

Das multivariate Verfahren mv.MLP.ACF der NN zeigt die höchste Dispositionsgüte aller Verfahren der Vergleichsstudie aufgrund der geringsten Dispositionskosten von 0,0399 $mwMLLC$ über alle Zeitreihen und Dispositionszeitpunkte. Somit beweist das neuartige, mit asymmetrischen Kostenfunktionen trainierte mv.MLP.ACF niedrigere Kosten und somit eine bessere Entscheidungsunterstützung als alle etablierten Prognoseverfahren und eine konventionelle Bestellmengenrechnung im Newsboy-Modell.

¹ Dies stellt die übliche Vorgehensweise von Expertensystemen wie zum Beispiel der Software ForecastPro, SAS Forecast Server oder Autobox – zumindest für die Erkenntnisse aus diesem Datensatz –grundsätzlich in Frage, da diese auf einer Identifikation der Zeitreihenstruktur und entsprechenden Expertenregeln zur Modellselektion beruhen. Entsprechend werfen diese Erkenntnisse weiterführende Forschungsfragen auf, die jedoch eingehende Analysen in separaten Experimenten erfordern.

² Die Berechnungsvorschrift wurde in Abschnitt 6.2.4 hinreichend beschrieben. Für die Verfahren ohne Modellselektion NF1, NF2, MA, S.ES, DT.ES, MLP und mv.MLP werden Sicherheitsbestände auf konventionellem Wege berechnet, unter Ermittlung der Standardabweichung der Prognoseresiduen. Für die Verfahren mit individueller Modellselektion erfolgt die Berechnung der Sicherheitsbestände und Bestellmengen durch interne Berechnungsvorschriften des Expertensystems unter automatischer Identifikation der Verteilungsform der Residuen.

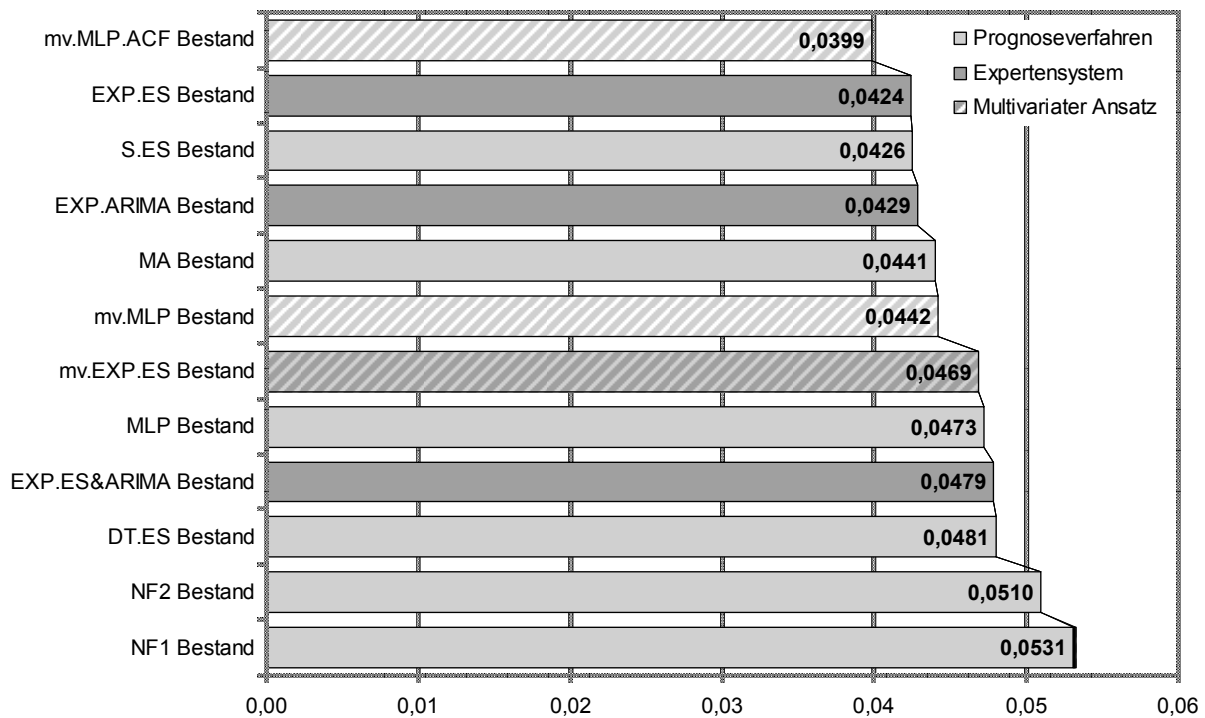


Abbildung 6.21: Dispositionskosten der Prognoseverfahren nach *mwMLLC* auf den Testdaten über alle Zeitreihen¹

Abbildung 6.21 zeigt neben dem mv.MLP.ACF drei homogene Gruppen von Verfahren mit ähnlichen Dispositionskosten: die Verfahren EXP.ES, S.ES, EXP.ARIMA, MA und mv.MLP mit durchschnittlichen Kosten um 0,043, die Gruppe aus mv.EXP.ES, MLP, EXP.ES&ARIMA und DT.ES mit durchschnittlichen Fehlerkosten von 0,048 und die Naiven Benchmark-Verfahren NF1 und NF2 mit durchschnittlichen Fehlerkosten von 0,052. Wie zuvor beim Vergleich der Prognosegenauigkeit zeigen die Naiven Verfahren NF1 und NF2 die geringste Dispositionsgenauigkeit und die höchsten Fehlerkosten aller Verfahren. Sie bestätigen somit ihre Rolle als einfacher Verfahrens-Benchmark und verdeutlichen, dass alle Prognoseverfahren durch die signifikant verbesserten Ergebnisse der Prognose- und Dispositionsgüte Informationen in der Struktur der Zeitreihen erkennen und somit valide und reliable Ansätze zur Warendisposition darstellen.

Die positive und konstante Leistung von NN in der Bedarfsprognose wird in der Beurteilung ihrer Güte zur Bestellmengenrechnung daher nur zum Teil bestätigt. Während das mv.MLP und das MLP im Rahmen der Bedarfsprognose als beste Verfahren abschnitten, erreichen sie nach der Berechnung konventioneller Sicherheitsbestände mit Dispositionskosten der *mwMLLC* von 0,0442 und 0,0473 lediglich Plätze im Mittelfeld aller Verfahren. Auch die Verfahren der Expertensysteme sowie multivariate Ausführungen der Verfahren sind über alle Gruppen der Dispositionskosten verteilt und lassen keine eindeutige Überlegenheit eines spezifischen Verfahrensansatzes erkennen. Dagegen zeigen einfache Prognoseverfahren der S.ES mit einer konventionellen Bestellmengenrechnung im Vergleich zu komplexen Expertensystemen zur integrierten Warendisposition wie auch schon bei der Prognosegenauigkeit

¹ Wie auch in der Untersuchung der Prognosefehler in Abbildung 6.14 sind die Verfahren absteigend sortiert nach den Fehlerkosten des *mwMLLC* bzw. aufsteigend nach dem Rang der Verfahren nach den Fehlerkosten des *mwMLLC*, was als *Rang mwMLLC* bezeichnet wird. Die Darstellung bildet erneut alle univariaten Prognoseverfahren mit aggregierter Modellselektion in heller Schattierung und alle Expertensysteme mit individueller Modellselektion in dunkler Schattierung ab. Multivariate Varianten der Expertensysteme sind schraffiert dargestellt.

eine vergleichsweise gute Dispositionsgüte. Dies ist nachfolgend mittels Analyse der Dispositionsgüte einzelner Absatzstellen und Zeitreihen näher zu analysieren.

Da die Dispositionskosten der *mwMLLC* die tatsächlich entstehenden Entscheidungskosten je Dispositionsvorgang, Ware und Absatzstelle repräsentieren, kann das Einsparpotenzial der *mv.MLP.ACF* im Vergleich zu konkurrierenden Verfahren unmittelbar bestimmt werden. Im Vergleich zur Disposition mit dem Naiven Prognoseverfahren NF1 können Einsparungen an Dispositionskosten in Höhe von € 0,0132 pro Dispositionsvorgang und € 12,36 pro Absatzstelle und Jahr erzielt werden, was einem Einsparpotenzial von über € 2.500.000 ohne Veränderungen der Bevorratungs- und Dispositionsprozesse entspricht. Im Vergleich der Dispositionskosten zum etablierten oder: mit den Dispositionskosten der etablierten Verfahren der gleitenden Mittelwerte MA aus der Praxis der Warenautomaten realisieren die NN dagegen nur eine Einsparung von € 0,0042 je Füllvorgang bzw. € 3,92 pro Absatzstelle und Jahr. Über alle Absatzstellen entspricht dies lediglich einer Einsparung von ca. € 850.000, was in Relation zur Verbesserung der Prognosegüte von 16,1% in Abschnitt 6.3.1 vergleichsweise gering erscheint.

Der Vergleich der konkurrierenden Verfahren erfordert erneut eine Feststellung der statistischen Unterschiede ihrer Bestellmengenkosten über alle Zeitreihen. Aufgrund der nicht normalverteilten Dispositionskosten¹ werden die Ergebnisse erneut durch eine Reihe von Wilcoxon-Vorzeichen-Rang-Tests der ordinalen Rangordnung verglichen. Die Ergebnisse sind paarweise für je zwei Dispositionsansätze in Tabelle 6.25 dargestellt.

Tabelle 6.24: Wilcoxon-Vorzeichen-Rang-Test auf Unterschiede in den Dispositionskosten nach *mwMLLC*²

Verfahren	EXP.		mv.		EXP.		mv.		mv.			
	NF1	NF2	MA	S.ES	DT.ES	ES	ARIMA	ES& ARIMA	ES	MLP	MLP	ACF
NF1	-	,842	,000**	,000**	,001**	,000**	,000**	,000**	,000**	,001**	,002**	,000**
NF2	,842	-	,000**	,000**	,001**	,000**	,000**	,001**	,000**	,008**	,001**	,000**
MA	,000**	,000**	-	,000**	,491	,029*	,104	,653	,131	,304	,129	,223
S.ES	,000**	,000**	,000**	-	,806	,343	,333	,408	,381	,546	,277	,461
DT.ES	,001**	,001**	,491	,806	-	,026*	,376	,376	,035*	,132	,006**	,484
EXP.ES	,000**	,000**	,029*	,343	,026*	-	,471	,042*	,811	,717	,250	,846
EXP.ARIMA	,000**	,000**	,104	,333	,376	,471	-	,352	,771	,677	,415	,425
EXP.ES&AR	,000**	,001**	,653	,408	,376	,042*	,352	-	,025*	,152	,023*	,438
mv.EXP.ES	,000**	,000**	,131	,381	,035*	,811	,771	,025*	-	,912	,657	,608
NN	,001**	,008**	,304	,546	,132	,717	,677	,152	,912	-	,083	,591
mv.MLP	,002**	,001**	,129	,277	,006**	,250	,415	,023*	,657	,083	-	,591
mv.MLP.ACF	,000**	,000**	,223	,461	,484	,846	,425	,438	,608	,591	,767	-
Test mind. *	10	10	4	3	5	5	2	5	4	2	4	2
Test mind. **	10	10	3	3	3	2	2	2	2	2	3	2

* Wilcoxon-Test ist signifikant bei einem $p \leq 0.05$ Konfidenzniveau (2-seitiger Test, Berechnung mit SPSS 13.0).

** Wilcoxon-Test ist hochsignifikant bei einem $p \leq 0.01$ Konfidenzniveau (2-seitiger Test, Berechnung mit SPSS 13.0).

Im Gegensatz zur visuellen Analyse werden die drei Gruppen von Dispositionsverfahren mit unterschiedlichen Fehlerkosten nicht durch die statistischen Signifikanztests bestätigt. Die paarweisen Wilcoxon-Vorzeichen-Rang-Tests belegen nur für die Naiven Verfahren NF1 und NF2 im Vergleich zu allen anderen Verfahren konsistent signifikante Unterschiede nach ihrer

¹ Dies wurde durch eine Reihe von Kolmogorov-Smirnov-Tests auf Normalverteilung sowie andere Verteilungsformen für alle Fehlerwerte bestätigt, die hier aus Platzgründen nicht abgebildet sind. Danach folgen die Fehlerkosten aller Verfahren einer Exponentialverteilung mit $p > 0,01$, DT.ES, EXP.ES&ARIMA, mv.EXP.ES, mv.MLP und MLP sogar mit $p > 0,05$.

² Die Ergebnisse wurden durch den Vorzeichen-Sign-Test und die Signifikanz der Unterschiede sowie die Rangordnung der Verfahren durch einen nichtparametrischen Friedman-Test bestätigt ($p=0,000$), der auf signifikante Unterschiede in wiederholten, ordinalen Messungen im Sinne einer ANOVA der Ränge testet.

ordinalen Rangfolge. Somit sind die mittleren Dispositionskosten des mv.MLP.ACF zwar im Mittel niedriger als die aller anderen Prognoseverfahren, sie unterscheiden sich jedoch nach der Rangordnung statistisch nicht signifikant von denen der MA, S.ES, DT.ES, EXP.ES, EXP.ARIMA, EXP.ES&ARIMA, mv.EXP.ES, MLP oder mv.MLP. Entsprechend bestehen auch keine signifikanten Unterschiede zwischen univariaten und multivariaten Ansätzen der Prognoseverfahren, zwischen Verfahren mit individueller und aggregierter Modellselektion oder den jeweiligen Verfahren der NN hinsichtlich ihrer Dispositionsgüte. Dies steht im direkten Gegensatz zu den Ergebnissen der Prognosegenauigkeit in Abschnitt 6.3.1, wonach die Verfahren der multivariaten mv.MLP und der univariaten MLP sowie der einfachen Exponentiellen Glättung S.ES signifikant niedrigere mittlere Prognosefehler aufwiesen als alle anderen statistischen Verfahren. Die Unterschiede gelten somit lediglich für die Prognosegenauigkeit, nicht aber für die daraus resultierenden Dispositionskosten.

Daher kann nicht mit statistischer Sicherheit bestätigt werden, dass Verfahren der NN eine signifikant bessere Prognosegenauigkeit aufweisen als etablierte statistische Verfahren und Expertensysteme, sondern lediglich gegenüber einfachen Naiven Benchmarks. Allerdings zeigen die Verfahren der NN somit auch keine signifikant schlechtere Dispositionsgüte als die in Theorie und Praxis etablierten Expertensysteme. Darüber hinaus zeigen die NN mit einer einzelnen Netzwerkarchitektur und einer Parametrisierung mit asymmetrischen Kostenfunktionen zur unmittelbaren Vorhersage von kostenminimalen Bestellmengen eine verringerte Komplexität in der Methodologie und der Anwendung im Vergleich zur sequenziellen Berechnung von Sicherheitsbeständen mit vielfachen statistischen Tests und einer Wissensbasis mit umfangreichen Regeln durch die Expertensysteme zur Prognose.

Eine grafische Analyse der mittleren Fehlerkosten $mwMLLC$ und ihrer Streuung über alle Zeitreihen und Zeitreihenursprünge bestätigt diese Erkenntnisse anhand der Box-Whisker-Graphen in Abbildung 6.21.

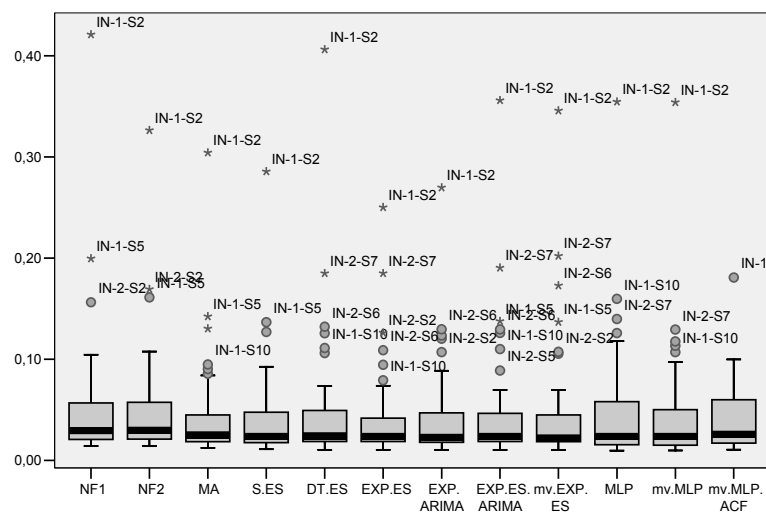


Abbildung 6.22: Box-Whisker-Graphen der Dispositionskosten nach $mwMLLC$ der Testdaten über alle Zeitreihen

Die Abbildung belegt, dass die mv.MLP.ACF eine im Vergleich zu anderen Verfahren weitgehend identische Verteilung der Prognosefehler mit einem leicht höheren Median aufweisen. Allerdings bestimmen sie die Bestellmengen nur mit einem extremen Fehlerwert in IN-1-S2 und zeigen somit eine konsistent hohe Dispositionsgüte ohne Verzerrungen durch extreme Dispositionskosten im Vergleich zu anderen Verfahren. In der Praxis sind entsprechend die gesamten Kosten für die Beurteilung der Wirtschaftlichkeit eines Verfahrens relevant, nicht der Median seiner Kosten unter Vernachlässigung extrem schlechter Vorhersa-

gen. Dies erklärt in der Bildung mittlerer Gütemaße die Entstehung der niedrigeren durchschnittlichen Kosten der *mwMLLC* und einer höheren mittleren Dispositionsgüte der *mv.MLP.ACF*. Andere Verfahren zeigen dagegen einen niedrigeren Median und nur eine insignifikant schlechtere Dispositionsgüte. Eine umfassende Beurteilung der Dispositionsgüte erfordert daher die Berücksichtigung unterschiedlicher Gütekriterien der Momente der Kostenverteilung. Dies bestätigt erneut die Nullhypothese H_2 , nach der unterschiedliche Fehlermaße zu unterschiedlichen Rangordnungen der Performanz der Methoden führen.

Eine Aufstellung aller Fehlerkosten der summierten, mittleren und medianbasierten *MLLC* der Dispositionsverfahren für die Trainings-, Validierungs- und Testdaten sowie ihre korrespondierenden Rangordnung bietet Tabelle 6.25.¹

Die Ergebnisse der Dispositionskosten in Tabelle 6.25 bestätigen das *mv.MLP.ACF* als Verfahren mit den niedrigsten *mwMLLC* und *sumMLLC* über alle Zeitreihen. Diese Ergebnisse sind konsistent über alle Teildatenmengen der Trainings-, Validierungs- und Testdaten mit dem primären Rang (1 / 1) 1, was eine zuverlässige Approximation auf der Kalibrationsstichprobe sowie die genaue Generalisierung auf den Testdaten ausdrückt. Allerdings sind die Ergebnisse nicht robust, so dass eine Bewertung nach *mdMLLC* trotz geringer Kosten auf der Training- und Validierungsmenge nur zu einem zehnten Rang auf den Testdaten führt, mit (2 / 1) 10. Auch eine Analyse des summierten Rangs *mwRang MLLC* zeigt für *mv.MLP.ACF* niedrige Dispositionskosten und einen ersten oder zweiten Rang auf der Kalibrationsstichprobe im Gegensatz zu einer nur durchschnittlichen Generalisierung und einem mittleren Rang auf den Testdaten. Dies erklärt die mangelnde statistische Signifikanz der Wilcoxon-Tests ($p > 0,05$), welche die Unterschiede der Verfahren nach ihrer ordinalen Rangfolge je Zeitreihe und nicht anhand der Mittelwerte bestimmen. Dennoch erlernt das *mv.MLP.ACF* eine kostenminimale Bestellmenge unmittelbar aus den Absatzdaten der Vergangenheit, so dass die Funktionalität des neuen Verfahrensansatzes auch auf empirischen Zeitreihen bestätigt wird. Zudem kann die Nullhypothese H_c verworfen werden, da die Genauigkeit von NN zur Warendisposition durch die simultane Prognose und Bestellmengenrechnung mit asymmetrischen Kostenfunktionen nicht signifikant geringer ist als die etablierter statistischer Benchmark-Verfahren.

Die Naiven Prognoseverfahren NF1 und NF2 mit konventioneller Berechnung von Sicherheitsbeständen zeigen konstant die niedrigste Dispositionsgüte aller Verfahren, mit den niedrigsten Rängen 11 und 12 über alle Datenmengen und Kostenmaße.

¹ Eine Analyse der summierten mittleren Kosten *sumMLLC* zeigt identische Rangreihung und relative Dispositionsgüte der Verfahren wie *mwMLLC*. Ferner ist der Rang nach *sumMLLC* und *mwMLLC* identisch mit dem Rang nach *sumSLLC* und *mwSLLC*, was die Validität und Konsistenz der Fehlerberechnung bestätigt. Somit wird nachfolgend auf ihre Abbildung und Diskussion verzichtet.

Tabelle 6.25: *MLLC* Kosten der Warenbestände und Rang nach *MLLC* über alle Absatzstellen

Verfahren	<i>sumMLLC</i>	<i>mwMLLC</i>	<i>mdMLLC</i>
NF1Bestand	(3,286 / 2,689) 2,338	(0,075 / 0,061) 0,053	(0,031 / 0,034) 0,029
NF2 Bestand	(2,818 / 2,558) 2,245	(0,064 / 0,058) 0,051	(0,033 / 0,031) 0,030
MA Bestand	(2,609 / 2,187) 1,940	(0,059 / 0,050) 0,044	(0,026 / 0,030) 0,025
S.ES Bestand	(2,532 / 2,076) 1,874	(0,058 / 0,047) 0,043	(0,026 / 0,027) 0,024
DT.ES Bestand	(2,552 / 1,982) 2,117	(0,058 / 0,045) 0,048	(0,025 / 0,024) 0,024
EXP.ES Bestand	(2,501 / 1,980) 1,867	(0,057 / 0,045) 0,042	(0,024 / 0,024) 0,024
EXP.ARIMA Bestand	(2,502 / 2,025) 1,886	(0,057 / 0,046) 0,043	(0,025 / 0,026) 0,023
EXP.ES&ARIMA Bestand	(2,593 / 2,052) 2,106	(0,059 / 0,047) 0,048	(0,027 / 0,024) 0,024
mv.EXP.ES Bestand	(2,090 / 2,001) 2,063	(0,047 / 0,045) 0,047	(0,021 / 0,026) 0,022
MLP Bestand	(2,611 / 2,059) 2,081	(0,059 / 0,047) 0,047	(0,024 / 0,028) 0,024
mv.MLP Bestand	(2,421 / 1,980) 1,945	(0,055 / 0,045) 0,044	(0,024 / 0,029) 0,024
mv.MLP.ACF Bestand	(1,709 / 1,587) 1,754	(0,039 / 0,036) 0,040	(0,023 / 0,022) 0,026
Verfahren	<i>Rang sumMLLC</i>	<i>Rang mwMLLC</i>	<i>Rang mdMLLC</i>
NF1Bestand	(12 / 12) 12	(12 / 12) 12	(11 / 12) 11
NF2 Bestand	(11 / 11) 11	(11 / 11) 11	(12 / 11) 12
MA Bestand	(9 / 10) 5	(9 / 10) 5	(8 / 10) 9
S.ES Bestand	(6 / 9) 3	(6 / 9) 3	(9 / 7) 3
DT.ES Bestand	(7 / 4) 10	(7 / 4) 10	(6 / 3) 8
EXP.ES Bestand	(4 / 3) 2	(4 / 3) 2	(5 / 4) 7
EXP.ARIMA Bestand	(5 / 6) 4	(5 / 6) 4	(7 / 5) 2
EXP.ES&ARIMA Bestand	(8 / 7) 9	(8 / 7) 9	(10 / 2) 5
mv.EXP.ES Bestand	(2 / 5) 7	(2 / 5) 7	(1 / 6) 1
MLP Bestand	(10 / 8) 8	(10 / 8) 8	(3 / 8) 4
mv.MLP Bestand	(3 / 2) 6	(3 / 2) 6	(4 / 9) 6
mv.MLP.ACF Bestand	(1 / 1) 1	(1 / 1) 1	(2 / 1) 10
Verfahren	<i>sumRang MLLC</i>	<i>mwRang MLLC</i>	<i>mdRang MLLC</i>
NF1Bestand	(472 / 471) 431	(10,7 / 10,7) 9,8	(11,0 / 11,0) 11,0
NF2 Bestand	(445 / 414) 441	(10,1 / 9,4) 10,0	(11,0 / 11,0) 11,0
MA Bestand	(344 / 372) 315	(7,8 / 8,5) 7,2	(8,0 / 9,0) 8,0
S.ES Bestand	(326 / 294) 251	(7,4 / 6,7) 5,7	(8,0 / 7,0) 5,5
DT.ES Bestand	(273 / 235) 268	(6,2 / 5,3) 6,1	(6,5 / 5,0) 6,0
EXP.ES Bestand	(281 / 230) 216	(6,4 / 5,2) 4,9	(6,0 / 5,0) 5,0
EXP.ARIMA Bestand	(284 / 259) 233	(6,5 / 5,9) 5,3	(6,0 / 5,0) 4,5
EXP.ES&ARIMA Bestand	(323 / 236) 234	(7,3 / 5,4) 5,3	(8,0 / 5,0) 5,0
mv.EXP.ES Bestand	(145 / 212) 194	(3,3 / 4,8) 4,4	(3,0 / 5,0) 4,0
MLP Bestand	(205 / 266) 215	(4,7 / 6,0) 4,9	(3,0 / 6,5) 3,0
mv.MLP Bestand	(182 / 260) 197	(4,1 / 5,9) 4,5	(3,0 / 6,0) 3,0
mv.MLP.ACF Bestand	(148 / 112) 241	(3,4 / 2,5) 5,5	(2,0 / 1,0) 4,5
Verfahren	<i>Rang sumRang MLLC</i>	<i>Rang mwRangMLLC</i>	<i>Rang mdRang MLLC</i>
NF1Bestand	(12 / 12) 11	(12 / 12) 11	(11 / 11) 11
NF2 Bestand	(11 / 11) 12	(11 / 11) 12	(11 / 11) 11
MA Bestand	(10 / 10) 10	(10 / 10) 10	(8 / 10) 10
S.ES Bestand	(9 / 9) 8	(9 / 9) 8	(8 / 9) 8
DT.ES Bestand	(5 / 4) 9	(5 / 4) 9	(7 / 2) 9
EXP.ES Bestand	(6 / 3) 4	(6 / 3) 4	(5 / 2) 6
EXP.ARIMA Bestand	(7 / 6) 5	(7 / 6) 5	(5 / 2) 4
EXP.ES&ARIMA Bestand	(8 / 5) 6	(8 / 5) 6	(8 / 2) 6
mv.EXP.ES Bestand	(1 / 2) 1	(1 / 2) 1	(2 / 2) 3
MLP Bestand	(4 / 8) 3	(4 / 8) 3	(2 / 8) 1
mv.MLP Bestand	(3 / 7) 2	(3 / 7) 2	(2 / 7) 1
mv.MLP.ACF Bestand	(2 / 1) 7	(2 / 1) 7	(1 / 1) 4

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

Darüber hinaus kann keine signifikante Überlegenheit von komplexen Ansätzen der multivariaten Prognoseverfahren oder der Expertensysteme gegenüber einfachen, univariaten Prognoseverfahren mit aggregierter Modellselektion festgestellt werden. Dies bestätigt die Hypothese H_1 auch für Anwendungen in der Warendisposition, da komplexe Prognoseverfahren hierbei nicht genauer sind als einfache. Ferner wird die Nullhypothese H_5 bestätigt, nach der Expertensysteme nur bei längeren Prognosehorizonten genauere Ergebnisse liefern als Ansätze mit aggregierter Modellselektion. Dagegen kann die Nullhypothese H_6 angesichts der vorliegenden Daten verworfen werden, da Verfahren mit aggregierter Modellselektion und Parameteroptimierung wie EXP.ES nicht genauer sind als vergleichbare Verfahren mit konstanten Parametern wie S.ES oder DT.ES. Die unterschiedlichen Varianten der NN mit einer konventionellen Berechnung von Bestellmengen zeigen ebenfalls keine signifikant niedrigeren Dispositionskosten als etablierte Benchmark-Verfahren, jedoch auch keine signifikant höheren Kosten in der Warendisposition. Somit wird auch für sie die Nullhypothese H_6 verworfen. Aufgrund ihrer statistisch zumindest gleichwertigen Leistung wie etablierte statistische Verfahren und Expertensysteme können die unterschiedlichen Varianten der NN daher als zusätzliche Benchmark-Verfahren sowohl zur einfachen Absatzprognose als auch in den Anwendungen zur Warendisposition etabliert werden. Ihre gezeigte Prognose- und Dispositionsgüte erfordert insbesondere ihre Berücksichtigung in zukünftigen Verfahrensvergleichen.

Die Güte der Dispositionsverfahren kann zudem durch eine Analyse der erzielten α - und β -Servicegrade aus den ermittelten Bestellmengen beurteilt und interpretiert werden. In Tabelle 6.26 sind die mittleren Servicegrade sowie die mittleren Abweichungen der Servicegrade von dem Ziel-Servicegrad von 96,5% je Dispositionsverfahren abgebildet.¹

¹ Die Abbildung der Abweichungen von α - und β -Servicegrad erlaubt grundsätzlich die Überprüfung der Zielerreichung des Newsboy-Modells auf den verrauschten Datenmengen. Aufgrund der erheblichen Zufallsschwankungen der Zeitreihen kann jedoch keine Zielerreichung einer Minimierung des α - oder eines β -Servicegrades wie auf den synthetischen Zeitreihen in Abschnitt 5.4 festgestellt werden.

Tabelle 6.26: Mittlere Servicegrade, Abweichungen vom 96,5%-Soll-Servicegrad und Rang über alle Zeitreihen

Verfahren	Servicegrad	
	<i>mwBETA Servicegrad</i>	<i>mwALPHA Servicegrad</i>
NF1Bestand	(97,7% / 97,6%) 99,2%	(97,6% / 97,6%) 97,6%
NF2 Bestand	(98,1% / 98,3%) 99,1%	(97,7% / 97,7%) 97,7%
MA Bestand	(98,0% / 97,8%) 99,1%	(97,6% / 97,4%) 97,4%
S.ES Bestand	(97,9% / 98,0%) 99,0%	(97,4% / 97,7%) 97,7%
DT.ES Bestand	(97,5% / 97,8%) 98,5%	(96,8% / 97,0%) 97,0%
EXP.ES Bestand	(97,5% / 97,8%) 98,8%	(96,9% / 97,1%) 97,1%
EXP.ARIMA Bestand	(97,3% / 97,7%) 98,8%	(96,7% / 97,0%) 97,0%
EXP.ES&ARIMA Bestand	(97,4% / 97,8%) 98,6%	(96,7% / 97,0%) 97,0%
mv.EXP.ES Bestand	(98,5% / 97,5%) 98,4%	(97,7% / 96,1%) 96,1%
MLP Bestand	(97,4% / 96,5%) 97,2%	(96,5% / 95,2%) 95,2%
mv.MLP Bestand	(97,2% / 96,3%) 97,4%	(95,8% / 94,9%) 94,9%
mv.MLP.ACF Bestand	(99,6% / 99,1%) 97,9%	(99,0% / 98,6%) 98,6%
Verfahren	Abweichung vom Soll- Servicegrad	
	<i>mw Abweichung vom Soll-BETA</i>	<i>mw Abweichung vom Soll-ALPHA</i>
NF1Bestand	(+1,2% / +1,0%) +2,7%	(+1,1% / +1,1%) +2,3%
NF2 Bestand	(+1,6% / +1,8%) +2,6%	(+1,2% / +1,2%) +2,4%
MA Bestand	(+1,5% / +1,3%) +2,6%	(+1,1% / +0,9%) +2,0%
S.ES Bestand	(+1,4% / +1,4%) +2,5%	(+0,9% / +1,1%) +1,9%
DT.ES Bestand	(+1,0% / +1,3%) +1,9%	(+0,3% / +0,5%) +1,3%
EXP.ES Bestand	(+1,0% / +1,3%) +2,3%	(+0,4% / +0,6%) +1,5%
EXP.ARIMA Bestand	(+0,8% / +1,2%) +2,2%	(+0,2% / +0,4%) +1,5%
EXP.ES&ARIMA Bestand	(+0,9% / +1,3%) +2,1%	(+0,2% / +0,5%) +1,3%
mv.EXP.ES Bestand	(+2,0% / +1,0%) +1,9%	(+1,2% / - 0,4%) +0,7%
MLP Bestand	(+0,9% / - 0,1%) +0,6%	(- 0,0% / - 1,3%) +0,2%
mv.MLP Bestand	(+0,6% / - 0,2%) +0,9%	(- 0,7% / - 1,6%) - 0,0%
mv.MLP.ACF Bestand	(+3,1% / +2,6%) +1,4%	(+2,5% / +2,1%) +0,7%
Verfahren	Rang der Abweichungen vom Servicegrad	
	<i>Rang mwAbweichung Soll-BETA</i>	<i>Rang mwAbweichung Soll-ALPHA</i>
NF1Bestand	(7 / 4) 12	(9 / 7) 11
NF2 Bestand	(10 / 11) 11	(10 / 9) 12
MA Bestand	(9 / 9) 10	(8 / 6) 10
S.ES Bestand	(8 / 10) 9	(7 / 8) 9
DT.ES Bestand	(5 / 6) 5	(4 / 3) 5
EXP.ES Bestand	(6 / 7) 8	(5 / 5) 7
EXP.ARIMA Bestand	(2 / 5) 7	(2 / 2) 8
EXP.ES&ARIMA Bestand	(3 / 8) 6	(3 / 4) 5
mv.EXP.ES Bestand	(11 / 3) 4	(11 / 1) 3
MLP Bestand	(4 / 1) 1	(1 / 10) 2
mv.MLP Bestand	(1 / 2) 2	(6 / 11) 1
mv.MLP.ACF Bestand	(12 / 12) 3	(12 / 12) 4

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

Die Servicegrade in Tabelle 6.26 zeigen für alle Verfahrensansätze der Disposition eine zuverlässige Erreichung der Soll-Servicegrade mit einer geringen maximalen Abweichung aller Verfahren von 2,1% der β -Servicegrade und von 2,4% der α -Servicegrade.¹ Somit sind selbst die Naiven Prognoseverfahren NF1 und NF2 trotz signifikant höherer Prognosefehler (vgl. Abschnitt 6.3.1) in der Lage, durch die Bildung einfacher Sicherheitsbestände in der Bestellmengenrechnung nur unterproportionale, geringfügige Abweichungen vom Soll-

¹ Diese entstehen aus den β -Servicegraden von NF1 und mv.MLP Bestand sowie den α -Servicegraden von NF2 und mv.MLP Bestand.

Servicegrad zu erzielen. Die Interaktion von Prognosefehlern und Sicherheitsbeständen wirkt somit selbstregulierend und führt zu einer Verringerung der Sensitivität der Verfahren in der Bestellmengenrechnung.

Zudem verdeutlichen die Abweichungen der Servicegrade die Tendenz der statistischen Verfahren zur systematischen Realisation höherer Servicegrade als durch den Soll-Servicegrad von 96,5% vorgegeben.¹ Dies entspricht einer Vermeidung von Präsenzlücken mit höheren Fehlerkosten, was unter Berücksichtigung der starken Zufallsschwankungen der Zeitreihen und der daraus resultierenden Probleme in der Vorhersage einer kostenminimalen Bestellmenge zu verringerten Dispositionskosten führt und somit sinnvoll erscheint.

Die Varianten der NN durch MLP, mv.MLP und mv.MLP.ACF zeigen im Vergleich geringere Abweichungen vom Soll-Servicegrad als andere Dispositionsverfahren. Die valide Erzielung der Soll-Servicegrade begründet die geringeren mittleren Dispositionskosten der Verfahren der NN auf vielen Zeitreihen, was theoretisch zu den niedrigsten Dispositionskosten führen sollte. Somit können die geringen mittleren Dispositionskosten der mv.MLP.ACF, trotz nur insignifikanter Unterschiede zu den statistischen Ansätzen der Bestellmengenrechnung, durch die wenigen Abweichungen von den kostenoptimalen Servicegraden begründet werden. Die genauere Berechnung der Sicherheitsbestände kann aber auch zu einer höheren Sensitivität gegenüber Präsenzlücken führen. Lediglich die Verfahren der NN mit einer konventionellen Berechnung von Sicherheitsbeständen und Bestellmengen unterschreiten den Soll-Servicegrad auf den Trainings- und / oder den Validierungsdaten. Somit führt die gesteigerte Prognosegenauigkeit der MLP und der mv.MLP bei konventioneller Berechnung der Sicherheitsbestände zu verringerten Bestellmengen und einem höheren Risiko des Auftretens von kostenintensiven Präsenzlücken. Dies kann insbesondere bei mittelwert- und varianzinstationären Zeitreihen als Folge von Zufallsschwankungen und Ausreißern auf der Testmenge auftreten, die nicht auf der Kalibrationsstichprobe beobachtet werden konnten. Aufgrund der asymmetrischen Kosten verursachen die entstehenden Präsenzlücken höhere Fehlerkosten. Somit führt eine erhöhte Prognosegenauigkeit bei schwankenden Absatzmustern zwischen den Teildatenmengen zu niedrigeren Bestellmengen und häufigeren Präsenzlücken im Vergleich zu statistischen Verfahren mit höheren Prognosefehlern und einer systematischen Überhöhung der Sicherheitsbestände auf Trainings-, Validierungs- und Testmenge.²

6.3.2.2 Ergebnisse nach Absatzstelle

Nachfolgend werden die Ergebnisse der Dispositionsgüte je Absatzstelle der Innen- und Außenautomaten diskutiert, um analog zu den Prognosefehlern mögliche Regelmäßigkeiten in den Dispositionskosten der einzelnen Verfahren zu identifizieren. In Tabelle 6.27 sind zunächst die Summen und Mittelwerte der mittleren linearen Kosten *sumMLLC* und *mwMLLC* sowie die entsprechenden Rangfolgen der Verfahren an den vier Absatzstellen IN-1, IN-2, AU-1 und AU-2 dargestellt.

¹ Dieser wurde in Abschnitt 6.2.4.1 aus den asymmetrischen Dispositionskosten berechnet.

² Die Probleme einer unterschiedlichen Variabilität sowie von Instationaritäten und Strukturbrüchen zwischen Kalibrations- und Teststichprobe wurden bislang nicht untersucht und entsprechend nicht in der Berechnung von Sicherheitsbeständen berücksichtigt. Hierzu sind weitere systematische Vergleichsstudien erforderlich, um einen validen Risikozuschlag auf die Sicherheitsbestände aus der Variabilität der Prognosefehler und somit Sicherheitsbestände über mehrere Zeitreihenursprünge zu bestimmen.

Tabelle 6.27: Summen und Mittelwerte der *MLLC* der Warenbestände je Absatzstelle¹

<i>sumMLLC</i>				
Verfahren	AU-1	AU-2	IN-1	IN-2
NF1 Bestand	(0,359 / 0,370) 0,321	(0,355 / 0,391) 0,324	(1,591 / 1,268) 1,063	(0,980 / 0,659) 0,631
NF2 Bestand	(0,351 / 0,306) 0,322	(0,374 / 0,352) 0,344	(1,239 / 1,041) 0,962	(0,854 / 0,859) 0,617
MA Bestand	(0,284 / 0,283) 0,267	(0,312 / 0,313) 0,271	(1,315 / 1,032) 0,894	(0,698 / 0,560) 0,508
S.ES Bestand	(0,285 / 0,282) 0,261	(0,299 / 0,285) 0,276	(1,304 / 0,966) 0,850	(0,644 / 0,542) 0,487
DT.ES Bestand	(0,253 / 0,259) 0,265	(0,277 / 0,284) 0,277	(1,340 / 0,926) 0,957	(0,682 / 0,513) 0,619
EXP.ES Bestand	(0,268 / 0,261) 0,267	(0,276 / 0,289) 0,264	(1,278 / 0,916) 0,729	(0,680 / 0,515) 0,607
EXP.ARIMA Bestand	(0,279 / 0,254) 0,268	(0,281 / 0,287) 0,254	(1,224 / 0,939) 0,839	(0,717 / 0,546) 0,526
EXP.ES&ARIMA B.	(0,271 / 0,248) 0,265	(0,285 / 0,286) 0,261	(1,331 / 1,002) 0,933	(0,705 / 0,516) 0,648
mv.EXP.ES Bestand	(0,236 / 0,261) 0,237	(0,250 / 0,272) 0,258	(1,134 / 0,957) 0,918	(0,471 / 0,511) 0,649
MLP Bestand	(0,227 / 0,259) 0,255	(0,260 / 0,313) 0,277	(1,379 / 0,950) 0,976	(0,746 / 0,537) 0,573
mv.MLP Bestand	(0,219 / 0,252) 0,258	(0,259 / 0,327) 0,265	(1,298 / 0,834) 0,879	(0,645 / 0,566) 0,544
mv.MLP.ACF Best.	(0,233 / 0,222) 0,263	(0,262 / 0,254) 0,294	(0,730 / 0,642) 0,726	(0,484 / 0,469) 0,470
<i>mwMLLC</i>				
Verfahren	AU-1	AU-2	IN-1	IN-2
NF1 Bestand	(0,030 / 0,031) 0,027	(0,027 / 0,030) 0,025	(0,159 / 0,127) 0,106	(0,109 / 0,073) 0,070
NF2 Bestand	(0,029 / 0,025) 0,027	(0,029 / 0,027) 0,026	(0,124 / 0,104) 0,096	(0,095 / 0,095) 0,069
MA Bestand	(0,024 / 0,024) 0,022	(0,024 / 0,024) 0,021	(0,132 / 0,103) 0,089	(0,078 / 0,062) 0,056
S.ES Bestand	(0,024 / 0,024) 0,022	(0,023 / 0,022) 0,021	(0,130 / 0,097) 0,085	(0,072 / 0,060) 0,054
DT.ES Bestand	(0,021 / 0,022) 0,022	(0,021 / 0,022) 0,021	(0,134 / 0,093) 0,096	(0,076 / 0,057) 0,069
EXP.ES Bestand	(0,022 / 0,022) 0,022	(0,021 / 0,022) 0,020	(0,128 / 0,092) 0,073	(0,076 / 0,057) 0,067
EXP.ARIMA Bestand	(0,023 / 0,021) 0,022	(0,022 / 0,022) 0,020	(0,122 / 0,094) 0,084	(0,080 / 0,061) 0,058
EXP.ES&ARIMA B.	(0,023 / 0,021) 0,022	(0,022 / 0,022) 0,020	(0,133 / 0,100) 0,093	(0,078 / 0,057) 0,072
mv.EXP.ES Bestand	(0,020 / 0,022) 0,020	(0,019 / 0,021) 0,020	(0,113 / 0,096) 0,092	(0,052 / 0,057) 0,072
MLP Bestand	(0,019 / 0,022) 0,021	(0,020 / 0,024) 0,021	(0,138 / 0,095) 0,098	(0,083 / 0,060) 0,064
mv.MLP Bestand	(0,018 / 0,021) 0,021	(0,020 / 0,025) 0,020	(0,130 / 0,083) 0,088	(0,072 / 0,063) 0,060
mv.MLP.ACF Best.	(0,019 / 0,019) 0,022	(0,020 / 0,020) 0,023	(0,073 / 0,064) 0,073	(0,054 / 0,052) 0,052
<i>Rang sumMLLC</i>				
Verfahren	AU-1	AU-2	IN-1	IN-2
NF1 Bestand	(12 / 12) 11	(11 / 12) 11	(12 / 12) 12	(12 / 11) 10
NF2 Bestand	(11 / 11) 12	(12 / 11) 12	(4 / 11) 10	(11 / 12) 8
MA Bestand	(9 / 10) 8	(10 / 8) 6	(8 / 10) 6	(7 / 9) 3
S.ES Bestand	(10 / 9) 4	(9 / 4) 7	(7 / 8) 4	(3 / 7) 2
DT.ES Bestand	(5 / 5) 6	(6 / 3) 8	(10 / 4) 9	(6 / 3) 9
EXP.ES Bestand	(6 / 8) 9	(5 / 7) 4	(5 / 3) 2	(5 / 4) 7
EXP.ARIMA Bestand	(8 / 4) 10	(7 / 6) 1	(3 / 5) 3	(9 / 8) 4
EXP.ES&ARIMA B.	(7 / 2) 7	(8 / 5) 3	(9 / 9) 8	(8 / 5) 11
mv.EXP.ES Bestand	(4 / 7) 1	(1 / 2) 2	(2 / 7) 7	(1 / 2) 12
MLP Bestand	(2 / 6) 2	(3 / 9) 9	(11 / 6) 11	(10 / 6) 6
mv.MLP Bestand	(1 / 3) 3	(2 / 10) 5	(6 / 2) 5	(4 / 10) 5
mv.MLP.ACF Best.	(3 / 1) 5	(4 / 1) 10	(1 / 1) 1	(2 / 1) 1

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

Die Dispositionskosten in Tabelle 6.27 bestätigen die hohe Dispositionsgüte der NN mit asymmetrischen Kostenfunktionen des Typs mv.MLP.ACF lediglich für die Absatzstellen der Innenautomaten, an denen sie mit 0,073 für IN-1 und 0,052 für IN-2 pro Dispositionsvorgang konstant die niedrigsten Dispositionskosten nach *mwMLLC* auf den Testdaten aufweisen. Darüber hinaus zeigen sie auch die niedrigsten Kosten auf der Trainings- und Validierungsmenge, was durch den Rang nach *sumMLLC* von (1 / 1) 1 für IN-1 und (2 / 1) 1

¹ Die Gütekriterien *Rang sumMLLC* und *Rang mwMLLC* sind identisch. Somit wird auf eine separate Darstellung verzichtet.

für IN-2 bestätigt wird. Somit sind die mv.MLP.ACF in der Lage, die Zeitreihen auf der Kalibrationsstichprobe valide und genau zu approximieren und auf den Testdaten mit hoher Genauigkeit zu generalisieren. Andere Ansätze der NN mit einer konventionellen Berechnung der Sicherheitsbestände zeigen dagegen nur mittlere Dispositionskosten und belegen Ränge im Mittelfeld, so dass die niedrigen Kosten des Verfahrens der mv.MLP.ACF nicht auf die Nichtlinearität der NN zurückzuführen sind, sondern auf die Integration der Entscheidungskosten in den Trainingsprozess mittels asymmetrischer Kostenfunktionen.

Im Vergleich dazu zeigen die mv.MLP.ACF nur eine mittelmäßige Dispositionsgüte an den Absatzstellen der Außenautomaten, die sich im Vergleich zu den Innengeräten durch einen geringeren Artikelumsatz auszeichnen. Das neuartige Verfahren der NN belegt bei Absatzstelle AU-1 mit (3 / 1) 5 nur einen fünften Rang auf den Testdaten, bei AU-2 mit (4 / 1) 10 sogar nur einen zehnten Rang. Für andere Varianten der NN, der statistischen Verfahren sowie der Expertensysteme ergibt sich erneut keine konsistente Rangfolge von Verfahren und Verfahrensklassen, wie bereits bei den insignifikanten Unterschieden der Verfahrenskosten über alle Zeitreihen. Dies wird auch durch eine Analyse der robusteren Fehlermaße der Mediane *mdMLLC* und ihre Rangfolge in Tabelle 6.28 sowie der Rangsummen nach *sumMLLC* in Tabelle 6.31 bestätigt.

Tabelle 6.28: Mediane der *MLLC* der Warenbestände je Absatzstelle

Verfahren	<i>mdMLLC</i>			
	AU-1	AU-2	IN-1	IN-2
NF1 Bestand	(0,028 / 0,026) 0,022	(0,028 / 0,025) 0,025	(0,087 / 0,061) 0,058	(0,043 / 0,032) 0,073
NF2 Bestand	(0,026 / 0,021) 0,022	(0,028 / 0,025) 0,026	(0,073 / 0,058) 0,078	(0,044 / 0,031) 0,053
MA Bestand	(0,021 / 0,021) 0,022	(0,025 / 0,023) 0,022	(0,070 / 0,057) 0,073	(0,041 / 0,030) 0,046
S.ES Bestand	(0,021 / 0,020) 0,021	(0,023 / 0,021) 0,021	(0,070 / 0,056) 0,065	(0,039 / 0,029) 0,051
DT.ES Bestand	(0,020 / 0,019) 0,020	(0,020 / 0,021) 0,021	(0,068 / 0,066) 0,060	(0,035 / 0,023) 0,040
EXP.ES Bestand	(0,020 / 0,018) 0,020	(0,021 / 0,021) 0,021	(0,069 / 0,059) 0,058	(0,035 / 0,023) 0,039
EXP.ARIMA Bestand	(0,019 / 0,018) 0,020	(0,022 / 0,021) 0,020	(0,060 / 0,072) 0,061	(0,038 / 0,026) 0,039
EXP.ES&ARIMA B.	(0,020 / 0,018) 0,020	(0,020 / 0,021) 0,020	(0,074 / 0,068) 0,063	(0,035 / 0,023) 0,039
mv.EXP.ES Bestand	(0,018 / 0,018) 0,019	(0,020 / 0,020) 0,020	(0,057 / 0,063) 0,064	(0,027 / 0,029) 0,038
MLP Bestand	(0,018 / 0,019) 0,016	(0,022 / 0,023) 0,019	(0,068 / 0,070) 0,063	(0,026 / 0,046) 0,059
mv.MLP Bestand	(0,018 / 0,017) 0,016	(0,020 / 0,025) 0,018	(0,064 / 0,063) 0,063	(0,030 / 0,031) 0,047
mv.MLP.ACF Best.	(0,021 / 0,016) 0,019	(0,021 / 0,020) 0,022	(0,039 / 0,043) 0,072	(0,024 / 0,023) 0,058
Verfahren	<i>Rang mdMLLC</i>			
	AU-1	AU-2	IN-1	IN-2
NF1 Bestand	(12 / 12) 11	(11 / 11) 11	(12 / 6) 2	(11 / 11) 12
NF2 Bestand	(11 / 11) 12	(12 / 10) 12	(10 / 4) 12	(12 / 9) 9
MA Bestand	(10 / 10) 10	(10 / 8) 10	(8 / 3) 11	(10 / 8) 6
S.ES Bestand	(8 / 9) 9	(9 / 7) 7	(9 / 2) 9	(9 / 6) 8
DT.ES Bestand	(7 / 8) 6	(3 / 3) 8	(5 / 9) 3	(5 / 2) 5
EXP.ES Bestand	(5 / 5) 7	(6 / 6) 6	(7 / 5) 1	(7 / 3) 3
EXP.ARIMA Bestand	(4 / 6) 7	(7 / 4) 3	(3 / 12) 4	(8 / 5) 4
EXP.ES&ARIMA B.	(6 / 3) 5	(4 / 5) 4	(11 / 10) 7	(6 / 1) 2
mv.EXP.ES Bestand	(1 / 4) 4	(1 / 2) 4	(2 / 8) 8	(3 / 7) 1
MLP Bestand	(2 / 7) 1	(8 / 9) 2	(6 / 11) 6	(2 / 12) 11
mv.MLP Bestand	(3 / 2) 2	(2 / 12) 1	(4 / 7) 5	(4 / 10) 7
mv.MLP.ACF Best.	(9 / 1) 3	(5 / 1) 9	(1 / 1) 10	(1 / 4) 10

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

Die Analyse der Mediane der mittleren Kosten bestätigt die Inkonsistenz der Dispositionsgüte aller Verfahren an den vier Absatzstellen. Während die Verfahren der Expertensysteme an der Absatzstelle IN-2 die vier niedrigsten Kosten aufweisen, zeigen sie an den anderen drei Absatzstellen lediglich eine mittlere Performanz. Die Verfahren des MLP und mv.MLP

erreichen bei den Absatzstellen AU-1 und AU-2 die geringsten Dispositionskosten, bei den Absatzstellen der Innengeräte hingegen nur einen mittleren bzw. unteren Rang. Allein die Naiven Verfahren NF1 und NF2 sowie die Mittelwerte zeigen eine konstante Leistung als die schlechtesten Verfahren mit den konsistent höchsten Dispositionskosten. Die geringe Dispositionsgüte der Naiven Verfahren sowie die inkonsistente Leistung anderer Verfahren wird auch durch die Summen der Ränge nach *sumMLLC* in Tabelle 6.29 bestätigt.

Tabelle 6.29: Summe der Ränge nach *sumMLLC* der Warenbestände je Absatzstelle¹

<i>Summe Rang sumMLLC</i>				
Verfahren	AU-1	AU-2	IN-1	IN-2
NF1 Bestand	(134 / 132) 124	(140 / 134) 134	(107 / 106) 84	(91 / 99) 89
NF2 Bestand	(128 / 118) 123	(147 / 119) 144	(74 / 75) 87	(96 / 102) 87
MA Bestand	(95 / 101) 100	(113 / 110) 90	(79 / 84) 73	(57 / 77) 52
S.ES Bestand	(96 / 101) 74	(100 / 63) 72	(73 / 63) 59	(57 / 67) 46
DT.ES Bestand	(64 / 70) 72	(72 / 61) 79	(80 / 66) 70	(57 / 38) 47
EXP.ES Bestand	(74 / 63) 67	(70 / 72) 66	(74 / 58) 36	(63 / 37) 47
EXP.ARIMA Bestand	(82 / 59) 71	(76 / 79) 56	(59 / 63) 53	(67 / 58) 53
EXP.ES&ARIMA Bestand	(78 / 51) 59	(96 / 67) 53	(83 / 76) 67	(66 / 42) 55
mv.EXP.ES Bestand	(47 / 67) 40	(42 / 45) 49	(36 / 60) 60	(20 / 40) 45
MLP Bestand	(46 / 62) 37	(54 / 90) 57	(53 / 67) 62	(52 / 47) 59
mv.MLP Bestand	(35 / 58) 46	(52 / 102) 45	(42 / 48) 54	(53 / 52) 52
mv.MLP.ACF Bestand	(56 / 30) 56	(51 / 42) 89	(19 / 12) 59	(22 / 28) 37
<i>Rang Rang sumMLLC</i>				
Verfahren	AU-1	AU-2	IN-1	IN-2
NF1 Bestand	(12 / 12) 12	(11 / 12) 11	(12 / 12) 11	(11 / 11) 12
NF2 Bestand	(11 / 11) 11	(12 / 11) 12	(7 / 9) 12	(12 / 12) 11
MA Bestand	(9 / 9) 10	(10 / 10) 10	(9 / 11) 10	(5 / 10) 6
S.ES Bestand	(10 / 9) 9	(9 / 4) 7	(6 / 5) 4	(5 / 9) 3
DT.ES Bestand	(5 / 8) 8	(6 / 3) 8	(10 / 7) 9	(5 / 3) 4
EXP.ES Bestand	(6 / 6) 6	(5 / 6) 6	(7 / 3) 1	(8 / 2) 4
EXP.ARIMA Bestand	(8 / 4) 7	(7 / 7) 4	(5 / 5) 2	(10 / 8) 8
EXP.ES&ARIMA Bestand	(7 / 2) 5	(8 / 5) 3	(11 / 10) 8	(9 / 5) 9
mv.EXP.ES Bestand	(3 / 7) 2	(1 / 2) 2	(2 / 4) 6	(1 / 4) 2
MLP Bestand	(2 / 5) 1	(4 / 8) 5	(4 / 8) 7	(3 / 6) 10
mv.MLP Bestand	(1 / 3) 3	(3 / 9) 1	(3 / 2) 3	(4 / 7) 6
mv.MLP.ACF Bestand	(4 / 1) 4	(2 / 1) 9	(1 / 1) 4	(2 / 1) 1

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

Nach den rangbasierten Gütemaßen der Summen der relativen Ränge zeigt das mv.MLP.ACF die niedrigsten Dispositionskosten an IN-2 mit (2 / 1) 1 und die viertniedrigsten Dispositionskosten an IN-1 mit (1 / 1) 4 und AU-1 mit (4 / 1) 4. Zudem erzielt mv.MLP.ACF niedrige Dispositionskosten auf den Trainings- und Validierungsmengen und belegt somit auch seine Fähigkeit zum validen Erlernen einer kostenminimalen Bestellmenge auf der Kalibrierungstichprobe. Auch die anderen multivariaten Verfahren mv.MLP und mv.EXP.ES mit konventioneller Berechnung der Bestellmengen zeigen eine konsistente und hohe Dispositionsgüte als eines der drei besten Verfahren. Für andere Verfahren kann erneut keine konsistente Rangreihung gefunden werden; einzig die Naiven Verfahren NF1 und NF2 und der MA bestätigen ihre signifikant schlechte Dispositionsgüte durch jeweils einen der unteren drei Ränge bei allen Absatzstellen.

¹ Der Rang nach der Summe der Ränge (*Rang Rang sumMLLC*) ist identisch mit dem *Rang Rang mwMLLC* der mittleren *MLLC* sowie nach der *sumSLLC* und *mwSLLC*. Auf ihre Abbildung wird daher verzichtet.

Eine Untersuchung der erzielten Servicegrade der Bestellmengen je Absatzstelle in Tabelle 6.32 zeigt, dass erneut alle Verfahren bis auf die NN mit konventioneller Bestellmengenrechnung homogen zur Bildung von kostengünstigeren Warenüberbeständen und somit überhöhten Servicegraden neigen.¹

Tabelle 6.30: Mittlere Servicegrade, Abweichungen vom 96,5%-Soll-Servicegrad und entsprechender Rang

<i>mwBETA</i> Servicegrad je Absatzstelle				
Verfahren	AU-1	AU-2	IN-1	IN-2
NF1 Bestand	(98,7%/97,0%) 99,0%	(99,0%/97,2%) 99,7%	(95,6%/97,2%) 99,7%	(96,9%/99,2%) 98,1%
NF2 Bestand	(98,8%/98,9%) 98,9%	(99,0%/98,2%) 99,4%	(96,5%/97,5%) 98,7%	(97,6%/98,7%) 99,5%
MA Bestand	(99,4%/98,1%) 99,1%	(99,1%/98,0%) 99,7%	(95,0%/96,1%) 98,2%	(98,1%/99,2%) 99,2%
S.ES Bestand	(99,1%/97,9%) 99,0%	(98,8%/98,2%) 99,6%	(95,1%/96,5%) 98,4%	(98,0%/99,4%) 99,0%
DT.ES Bestand	(99,1%/98,2%) 99,2%	(98,8%/97,8%) 99,3%	(93,9%/95,6%) 97,4%	(97,3%/99,6%) 97,5%
EXP.ES Bestand	(99,1%/98,0%) 99,3%	(98,8%/97,5%) 99,3%	(94,1%/96,2%) 98,8%	(97,3%/99,6%) 97,6%
EXP.ARIMA B.	(98,5%/98,3%) 99,4%	(98,6%/97,7%) 99,3%	(94,5%/95,5%) 97,7%	(96,9%/99,4%) 98,3%
EXP.ES&ARIMA	(99,1%/98,3%) 99,3%	(98,7%/97,6%) 99,3%	(94,0%/95,7%) 98,0%	(96,9%/99,6%) 97,3%
mv.EXP.ES B.	(99,7%/97,6%) 99,4%	(99,1%/97,6%) 99,2%	(96,2%/96,2%) 97,9%	(98,6%/98,7%) 96,5%
MLP Bestand	(99,5%/97,6%) 98,9%	(98,4%/96,0%) 97,3%	(93,7%/94,1%) 95,2%	(97,4%/98,4%) 96,9%
mv.MLP Bestand	(99,1%/97,4%) 98,3%	(98,1%/95,3%) 97,7%	(94,3%/94,7%) 96,2%	(96,4%/98,1%) 97,1%
mv.MLP.ACF B.	(100,0%/99,2%) 98,7%	(99,9%/99,4%) 98,1%	(98,9%/98,4%) 97,0%	(99,4%/99,6%) 97,6%
Abweichung vom Soll-Servicegrad je Absatzstelle				
Verfahren	AU-1	AU-2	IN-1	IN-2
NF1 Bestand	(+2,2%/+0,5%) +2,5%	(+2,5%/+0,7%) +3,2%	(- 0,9%/+0,6%) +3,1%	(+0,4%/+2,7%) +1,6%
NF2 Bestand	(+2,3%/+2,3%) +2,4%	(+2,4%/+1,6%) +2,9%	(- 0,0%/+1,0%) +2,1%	(+1,0%/+2,2%) +2,9%
MA Bestand	(+2,9%/+1,6%) +2,6%	(+2,5%/+1,4%) +3,2%	(- 1,5%/- 0,4%) +1,7%	(+1,6%/+2,7%) +2,6%
S.ES Bestand	(+2,5%/+1,4%) +2,5%	(+2,3%/+1,7%) +3,1%	(- 1,4%/- 0,1%) +1,9%	(+1,5%/+2,9%) +2,5%
DT.ES Bestand	(+2,6%/+1,6%) +2,6%	(+2,3%/+1,3%) +2,8%	(- 2,6%/- 0,9%) +0,9%	(+0,8%/+3,1%) +1,0%
EXP.ES Bestand	(+2,6%/+1,5%) +2,8%	(+2,3%/+1,0%) +2,8%	(- 2,4%/- 0,4%) +2,2%	(+0,8%/+3,1%) +1,1%
EXP.ARIMA B.	(+1,9%/+1,7%) +2,9%	(+2,1%/+1,2%) +2,8%	(- 2,0%/- 1,0%) +1,2%	(+0,4%/+2,9%) +1,7%
EXP.ES&ARIMA	(+2,6%/+1,8%) +2,8%	(+2,2%/+1,1%) +2,8%	(- 2,5%/- 0,8%) +1,5%	(+0,3%/+3,1%) +0,8%
mv.EXP.ES B.	(+3,2%/+1,1%) +2,9%	(+2,6%/+1,1%) +2,7%	(- 0,3%/- 0,3%) +1,4%	(+2,1%/+2,1%) - 0,0%
MLP Bestand	(+3,0%/+1,1%) +2,4%	(+1,8%/- 0,5%) +0,8%	(- 2,8%/- 2,5%) -1,3%	(+0,9%/+1,8%) +0,3%
mv.MLP Bestand	(+2,6%/+0,9%) +1,8%	(+1,5%/- 1,2%) +1,2%	(- 2,2%/- 1,8%) -0,4%	(- 0,2%/+1,6%) +0,6%
mv.MLP.ACF B.	(+3,4%/+2,7%) +2,2%	(+3,4%/+2,9%) +1,6%	(+2,4%/+1,8%) +0,5%	(+2,9%/+3,1%) +1,1%
Rang Abweichung vom Soll-Servicegrad				
Verfahren	AU-1	AU-2	IN-1	IN-2
NF1 Bestand	(2 / 1) 6	(9 / 2) 12	(3 / 5) 12	(3 / 5) 8
NF2 Bestand	(3 / 11) 3	(8 / 10) 9	(1 / 8) 10	(8 / 4) 12
MA Bestand	(9 / 7) 7	(10 / 9) 11	(5 / 4) 8	(10 / 6) 11
S.ES Bestand	(4 / 5) 5	(7 / 11) 10	(4 / 1) 9	(9 / 7) 10
DT.ES Bestand	(5 / 8) 8	(6 / 8) 6	(11 / 7) 3	(6 / 11) 5
EXP.ES Bestand	(6 / 6) 9	(5 / 3) 6	(9 / 3) 11	(5 / 11) 6
EXP.ARIMAB.	(1 / 9) 11	(3 / 7) 5	(6 / 9) 4	(4 / 8) 9
EXP.ES&ARIMA	(8 / 10) 9	(4 / 5) 6	(10 / 6) 7	(2 / 10) 4
mv.EXP.ES B.	(11 / 3) 11	(11 / 4) 4	(2 / 2) 6	(11 / 3) 1
MLP Bestand	(10 / 4) 3	(2 / 1) 1	(12 / 12) 5	(7 / 2) 2
mv.MLP Bestand	(7 / 2) 1	(1 / 6) 2	(7 / 10) 1	(1 / 1) 3
mv.MLP.ACF B.	(12 / 12) 2	(12 / 12) 3	(8 / 11) 2	(12 / 9) 7

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

¹ Lediglich auf der Kalibrationsstichprobe der Absatzstelle IN-1 zeigen alle Verfahren bis auf mv.MLP.ACF zu niedrige Warenbestände. Dies bestätigt sich jedoch nicht auf den Testdaten, so dass keine einheitliche, absatzstellenspezifische Verzerrung identifiziert werden kann.

Auch bei den erreichten Servicegraden zeigen das mv.MLP.ACF sowie die Varianten der NN aus MLP und mv.MLP als die drei Verfahren mit den geringsten Abweichungen vom Soll-Servicegrad eine hohe Dispositionsgüte auf der Testmenge für AU-1 und AU-2 sowie auf IN-1 und IN-2 als jeweils zwei der besten drei Verfahrensansätze.

Die Untersuchung der Kosten der Warendisposition zeigt auch bei der Analyse der Verfahren an unterschiedlichen Absatzstellen keine Ableitung von Regelmäßigkeiten oder signifikanten Unterschieden, in denen ein statistisches oder neuronales Verfahren andere Verfahren systematisch übertrifft. Für eine weitergehende Analyse werden daher nachfolgend zunächst die einzelnen Zeitreihen betrachtet und im Anschluss nach ihren Charakteristika gruppiert.

6.3.2.3 Ergebnisse nach Zeitreihen

Nachfolgend wird die relative Dispositionsgüte der Verfahren je Zeitreihe nach den mittleren Dispositionskosten $mwMLLC$ analysiert. Hierzu werden die Dispositionskosten der univariaten und der multivariaten Prognoseverfahren je Zeitreihe in jeweils zwei Kivitagraphen der Innen- und Außenautomaten abgebildet.

Abbildung 6.23 vergleicht zunächst die multivariaten Dispositionsverfahren des mv.MLP.ACF mit asymmetrischen Kostenfunktionen mit dem mv.MLP mit konventioneller Bestellmengenrechnung und mit dem Expertensystem des mv.EXP.ES.¹ Der Kivitagraph in Abbildung 6.23 (a.) zeigt die Dispositionskosten der $mwMLLC$ für alle Zeitreihen der Außenautomaten AU-1 und AU-2, Kivitagraph (b.) die $MLLC$ für alle Zeitreihen der Innenautomaten IN-1 und IN-2.

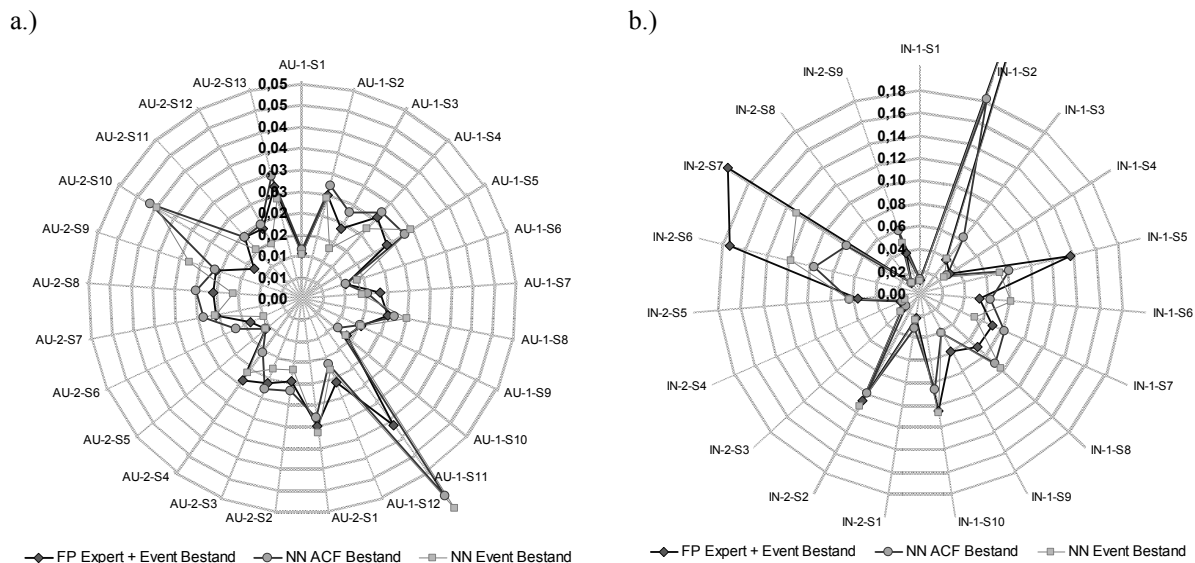


Abbildung 6.23: Dispositionsgüte der multivariaten Verfahren mv.MLP.ACF, mv.MLP und mv.EXP.ES nach $MLLC$ je Zeitreihe auf den Testmengen der Außenautomaten (a.) und der Innenautomaten (b.)

Im Gegensatz zur Analyse der Prognosefehler je Zeitreihe in Abbildung 6.19 verdeutlichen Abbildung 6.23 (a.) und (b.) die geringen Unterschiede zwischen den multivariaten Verfahren durch die geringen Abstände der Linien im Kivitagraph und den schwankenden Rang des besten Verfahrens. Dies erklärt die zum Teil insignifikanten Unterschiede der mittleren Kosten und die Inkonsistenz in der Analyse der Dispositionsverfahren nach ihrer Rangordnung. Als Konsequenz kann das mv.MLP.ACF in der Disposition nicht eindeutig als

¹ Wie bereits in Abbildung 6.19 und Abbildung 6.20 zeigt jeder Kivitagraph die Fehlerprofile der einzelnen Dispositionsverfahren als Linien über die Achsen der einzelnen Zeitreihen an, wobei näher am Ursprung liegende Linien geringere Fehlerkosten der $MLLC$ auf der Testmenge anzeigen.

überlegenes Verfahren identifiziert werden. Auch das mv.MLP unterscheidet sich hinsichtlich der mittleren Kosten $MLLC$ nicht signifikant von der Dispositionsgüte des Expertensystems der Exponentiellen Glättung mv.EXP.ES, obwohl hinsichtlich ihrer Prognosegenauigkeit signifikante Unterschiede nachgewiesen wurden (vgl. Abschnitt 6.3.1.3).

Abbildung 6.24 bestätigt diese Ergebnisse im Vergleich der univariaten Dispositionsverfahren des MLP und ausgewählten statistischen Prognoseverfahren unter konventioneller Berechnung von Sicherheitsbeständen in zwei Kivitagraphen der Innen- und Außenautomaten.

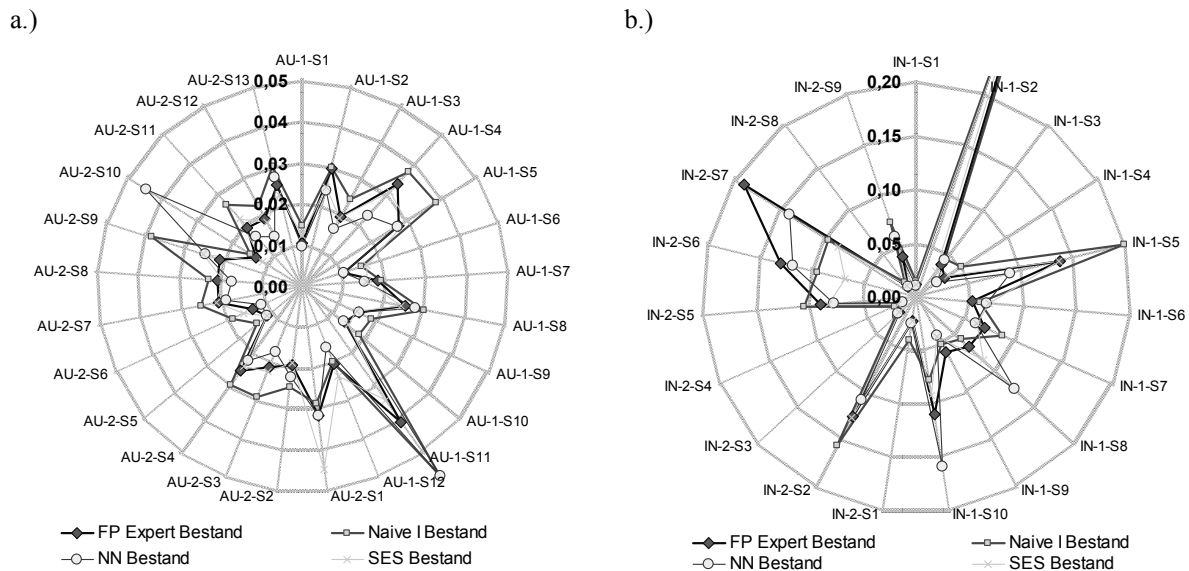


Abbildung 6.24: Dispositionsgüte der univariaten Verfahren MLP, NF1, S.ES, und EXP.ES&ARIMA nach $MLLC$ je Zeitreihe an Außenautomaten (a.) und Innenautomaten (b.)

Die Abbildung zeigt die niedrigsten Dispositionskosten des univariaten MLP auf der Mehrzahl der Zeitreihen im Vergleich zu anderen Prognoseverfahren. Die Abstände der Fehlerwerte sind jedoch gering, und das Verfahren wird wechselweise durch unterschiedliche andere Verfahren übertroffen, was die nur insignifikanten Unterschiede nach der Rangreihung der Verfahren begründet. Die Ergebnisse können erneut nicht die vorhergegangenen Erkenntnisse im Vergleich der Verfahren hinsichtlich ihrer Prognosegenauigkeit bestätigen. Vielmehr kann erneut keine signifikante Überlegenheit eines Verfahrens oder einer Verfahrensklasse der Disposition ermittelt werden.

Dennoch zeigen die unterschiedlichen Varianten der NN in Form der univariaten MLP, der multivariaten mv.MLP sowie der mv.MLP.ACF zur direkten Ermittlung kostenminimaler Bestellmengen eine wettbewerbsfähige Dispositionsgüte, gemessen an den mittleren Dispositionskosten.

6.3.2.4 Ergebnisse nach Zeitreihencharakteristika

Anschließend wird die Dispositionsgüte der Verfahren, geordnet nach unterschiedlichen warenlogistischen Kriterien der Anwendungsdomäne, analysiert. Hierzu werden die Zeitreihen nach der Struktur des Zeitreihenverlaufs, der Geschwindigkeit des Warenumschlages und der Bedeutung der Ware im Sortiment kategorisiert und analysiert. In Tabelle 6.31 sind zunächst die Ergebnisse aller Zeitreihen, kategorisiert nach der Struktur des Absatzverlaufs (siehe Abschnitt 6.2.2) für Zeitreihen mit regelmäßigen, unregelmäßigen und sporadischen Verlaufsformen, dargestellt.

Tabelle 6.31: *mwMLLC* und *mdMLLC* über alle Zeitreihen nach Absatzverlauf

Verfahren	<i>mwMLLC</i>		
	Sporadisch	Unregelmäßig	Regelmäßig
NF1 Bestand	(0,034 / 0,031) 0,025	(0,065 / 0,051) 0,045	(0,094 / 0,078) 0,068
NF2 Bestand	(0,036 / 0,027) 0,035	(0,058 / 0,048) 0,042	(0,077 / 0,075) 0,064
MA Bestand	(0,032 / 0,027) 0,030	(0,052 / 0,045) 0,037	(0,073 / 0,059) 0,055
S.ES Bestand	(0,034 / 0,027) 0,030	(0,051 / 0,042) 0,035	(0,070 / 0,057) 0,053
DT.ES Bestand	(0,032 / 0,026) 0,027	(0,053 / 0,041) 0,038	(0,069 / 0,054) 0,064
EXP.ES Bestand	(0,033 / 0,026) 0,026	(0,051 / 0,040) 0,037	(0,068 / 0,054) 0,052
EXP.ARIMA Bestand	(0,032 / 0,025) 0,028	(0,049 / 0,042) 0,034	(0,070 / 0,055) 0,056
EXP.ES&ARIMA Bestand	(0,036 / 0,024) 0,026	(0,052 / 0,043) 0,040	(0,071 / 0,056) 0,061
mv.EXP.ES Bestand	(0,026 / 0,024) 0,026	(0,042 / 0,041) 0,040	(0,058 / 0,055) 0,059
MLP Bestand	(0,028 / 0,029) 0,048	(0,041 / 0,044) 0,035	(0,087 / 0,053) 0,061
mv.MLP Bestand	(0,025 / 0,026) 0,039	(0,041 / 0,041) 0,032	(0,077 / 0,053) 0,058
mv.MLP.ACF Bestand	(0,018 / 0,021) 0,039	(0,032 / 0,031) 0,035	(0,051 / 0,045) 0,046

Verfahren	Rang <i>mwMLLC</i>		
	Sporadisch	Unregelmäßig	Regelmäßig
NF1 Bestand	(10 / 12) 1	(12 / 12) 12	(12 / 12) 12
NF2 Bestand	(12 / 9) 9	(11 / 11) 11	(9 / 11) 11
MA Bestand	(5 / 10) 8	(8 / 10) 7	(8 / 10) 4
S.ES Bestand	(9 / 8) 7	(6 / 7) 5	(5 / 9) 3
DT.ES Bestand	(7 / 7) 5	(10 / 3) 8	(4 / 4) 10
EXP.ES Bestand	(8 / 5) 2	(7 / 2) 6	(3 / 5) 2
EXP.ARIMA Bestand	(6 / 4) 6	(5 / 6) 2	(6 / 6) 5
EXP.ES&ARIMA Bestand	(11 / 3) 2	(9 / 8) 10	(7 / 8) 8
mv.EXP.ES Bestand	(3 / 2) 4	(4 / 4) 9	(2 / 7) 7
MLP Bestand	(4 / 11) 12	(2 / 9) 4	(11 / 2) 9
mv.MLP Bestand	(2 / 6) 11	(3 / 5) 1	(10 / 3) 6
mv.MLP.ACF Bestand	(1 / 1) 10	(1 / 1) 3	(1 / 1) 1

Verfahren	<i>mdMLLC</i>		
	Sporadisch	Unregelmäßig	Regelmäßig
NF1 Bestand	(0,018 / 0,024) 0,015	(0,028 / 0,031) 0,025	(0,033 / 0,038) 0,030
NF2 Bestand	(0,019 / 0,020) 0,015	(0,032 / 0,027) 0,026	(0,034 / 0,037) 0,032
MA Bestand	(0,016 / 0,023) 0,013	(0,025 / 0,029) 0,021	(0,028 / 0,030) 0,026
S.ES Bestand	(0,015 / 0,022) 0,012	(0,024 / 0,025) 0,019	(0,027 / 0,030) 0,025
DT.ES Bestand	(0,014 / 0,020) 0,012	(0,023 / 0,023) 0,021	(0,026 / 0,027) 0,029
EXP.ES Bestand	(0,014 / 0,020) 0,012	(0,024 / 0,023) 0,021	(0,027 / 0,030) 0,026
EXP.ARIMA Bestand	(0,015 / 0,019) 0,012	(0,021 / 0,026) 0,020	(0,025 / 0,026) 0,028
EXP.ES&ARIMA Bestand	(0,016 / 0,019) 0,012	(0,028 / 0,023) 0,021	(0,028 / 0,025) 0,026
mv.EXP.ES Bestand	(0,014 / 0,019) 0,012	(0,020 / 0,025) 0,021	(0,022 / 0,028) 0,023
MLP Bestand	(0,011 / 0,020) 0,027	(0,021 / 0,026) 0,022	(0,026 / 0,031) 0,027
mv.MLP Bestand	(0,010 / 0,020) 0,025	(0,025 / 0,027) 0,017	(0,024 / 0,032) 0,025
mv.MLP.ACF Bestand	(0,012 / 0,017) 0,027	(0,023 / 0,022) 0,025	(0,024 / 0,022) 0,027

Verfahren	Rang <i>mdMLLC</i>		
	Sporadisch	Unregelmäßig	Regelmäßig
NF1 Bestand	(11 / 12) 8	(11 / 12) 11	(11 / 12) 11
NF2 Bestand	(12 / 9) 9	(12 / 9) 12	(12 / 11) 12
MA Bestand	(10 / 11) 7	(9 / 11) 4	(10 / 8) 6
S.ES Bestand	(8 / 10) 1	(7 / 5) 2	(8 / 7) 2
DT.ES Bestand	(6 / 8) 2	(5 / 3) 4	(5 / 4) 10
EXP.ES Bestand	(4 / 7) 6	(6 / 4) 8	(7 / 6) 4
EXP.ARIMA Bestand	(7 / 2) 2	(2 / 7) 3	(4 / 3) 9
EXP.ES&ARIMA Bestand	(9 / 3) 2	(10 / 2) 4	(9 / 2) 4
mv.EXP.ES Bestand	(5 / 4) 2	(1 / 6) 4	(1 / 5) 1
MLP Bestand	(2 / 6) 12	(3 / 8) 9	(6 / 9) 7
mv.MLP Bestand	(1 / 5) 10	(8 / 10) 1	(3 / 10) 2
mv.MLP.ACF Bestand	(3 / 1) 11	(4 / 1) 10	(2 / 1) 8

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

Die Ergebnisse geben Hinweise auf mögliche Ursachen der inkonsistenten Dispositionsgüte der einzelnen Verfahren. Das Verfahren des mv.MLP.ACF zeigt die geringsten Kosten der *mwMLLC* für die Zeitreihen mit regelmäßigem Verlauf auf Trainings-, Validierungs- und Testmenge (1 / 1) 1 sowie die drittniedrigsten Kosten auf den Zeitreihen mit unregelmäßigem Verlauf (1 / 1) 3. Dagegen ist das mv.MLP.ACF zwar in der Lage, die Struktur der sporadischen Zeitreihen auf der Kalibrierungsstichprobe besser als jedes andere Verfahren zu erlernen, nicht aber auf den Testdaten zu generalisieren, mit einem zehnten Rang (1 / 1) 10. Auch die anderen Verfahren der NN zeigen auf den Testdaten der sporadischen Zeitreihen die schlechteste Dispositionsgüte aller Verfahren, das MLP mit (4 / 11) 12 und das mv.MLP mit (2 / 6) 11.

Im Gegensatz zu den NN zeigen alle Varianten der Expertensysteme eine robuste Dispositionsgüte auf den Zeitreihen mit sporadischem Verlauf, was auf die speziellen Routinen zur Erkennung und Vorhersage von Zeitreihen mit Nullperioden zurückzuführen ist.¹ Die geringe Dispositionsgüte der NN im Gegensatz zu der robusten Dispositionsleistung der Expertensysteme wird ebenfalls durch die Analyse der medianbasierten Kosten der *mdMLLC* sowie der Rangordnung in Tabelle 6.31 bestätigt. Somit scheinen Verfahren der NN weniger zur Bestellmengenrechnung auf sporadischen Zeitreihen geeignet zu sein als statistische Verfahren. Die geringe Dispositionsgüte auf sporadischen Zeitreihen im Gegensatz zur hohen Dispositionsgüte auf regelmäßigen Zeitreihen führt zu einer inkonsistenten Rangreihung der Verfahren der NN über alle Zeitreihen, hat aber wegen des geringen Umsatzes von wenigen Verkaufseinheiten auf den sporadischen Zeitreihen nur unterproportionale Auswirkungen auf die mittelwertbasierten Gütekriterien. Dies trägt zur Erklärung der inkonsistenten und statistisch insignifikanten Rangordnungen der Dispositionsverfahren bei und dient als Indiz für die Notwendigkeit weiterführender Untersuchungen der NN.²

In Tabelle 6.32 sind die Dispositionskosten der Verfahren, kategorisiert nach der Umschlagsgeschwindigkeit der einzelnen Waren in Schnelldreher und Langsamdreher, aufgeführt. Auf den Zeitreihen mit schnellem Warenumschlag wird erneut das mv.MLP.ACF als Verfahren mit den niedrigsten Dispositionskosten nach dem Mittelwert *mwMLLC* sowie dem Median der Kosten *mdMLLC* identifiziert, mit dem jeweils höchsten Rang (1 / 1) 1. Auf den Zeitreihen mit langsamem Warenumschlag zeigt mv.MLP.ACF dagegen nur eine durchschnittliche Dispositionsgüte von (1 / 1) 4 nach *mwMLLC* und (4 / 1) 9 nach *mdMLLC*. Dies ist zu erwarten, da diese Kategorie von Zeitreihen auch alle Artikel mit sporadischem Verlauf umfasst, auf denen die Ansätze der NN keine hohe Dispositionsgüte erreichen. Wie zuvor zeigen darüber hinaus lediglich die Naiven Verfahren NF1 und NF2 eine konsistente Leistung als schlechteste Verfahren, während andere Verfahrensansätze und Verfahrensklassen keine klaren Unterschiede zwischen der Disposition von Waren mit schnellem bzw. langsamem Artikelumschlag erkennen lassen.

¹ Die höchste Dispositionsgüte des Naiven Verfahrens NF1 auf den Testdaten bei einer geringen Güte auf der Kalibrationsstichprobe mit (10 / 12) 1 erscheint hingegen zufällig.

² Das Erlernen sporadischer Zeitreihenmuster durch NN erfordert im Gegensatz zu kontinuierlichen Zeitreihen ohne Nullperioden eine spezifische Modellierung der Eingabedaten, beispielsweise unter Kodierung der Distanzen zwischen Perioden mit positivem Bedarf durch eine intervallskalierte Erklärungsvariable ähnlich wie bei der Exponentiellen Glättung nach CROSTON. Erste Experimente zeigen vielversprechende Ergebnisse, widersprechen jedoch dem hier gewählten Verfahrensaufbau von nur einer konstanten Modellstruktur der NN und erfordern somit eine systematische Überprüfung und Analyse in einem separaten Verfahrenvergleich.

Tabelle 6.32: *mwMLLC* und *mdMLLC* über alle Absatzstellen nach Umschlagsgeschwindigkeit

Verfahren	<i>mwMLLC</i>		<i>mdMLLC</i>	
	Langsamdreher	Schnelldreher	Langsamdreher	Schnelldreher
NF1 Bestand	(0,039 / 0,036) 0,033	(0,235 / 0,175) 0,145	(0,029 / 0,032) 0,028	(0,163 / 0,141) 0,100
NF2 Bestand	(0,039 / 0,034) 0,034	(0,178 / 0,168) 0,130	(0,031 / 0,029) 0,027	(0,148 / 0,145) 0,099
MA Bestand	(0,033 / 0,030) 0,028	(0,178 / 0,140) 0,115	(0,025 / 0,026) 0,023	(0,114 / 0,126) 0,090
S.ES Bestand	(0,032 / 0,028) 0,028	(0,173 / 0,131) 0,109	(0,024 / 0,023) 0,023	(0,107 / 0,122) 0,087
DT.ES Bestand	(0,031 / 0,028) 0,030	(0,180 / 0,121) 0,128	(0,023 / 0,022) 0,022	(0,129 / 0,109) 0,109
EXP.ES Bestand	(0,031 / 0,028) 0,029	(0,172 / 0,122) 0,102	(0,023 / 0,022) 0,021	(0,127 / 0,109) 0,087
EXP.ARIMA Best.	(0,032 / 0,029) 0,027	(0,171 / 0,124) 0,115	(0,023 / 0,023) 0,021	(0,143 / 0,117) 0,114
EXP.ES&ARIMA	(0,032 / 0,028) 0,030	(0,179 / 0,130) 0,130	(0,023 / 0,022) 0,021	(0,134 / 0,123) 0,118
mv.EXP.ES Bestand	(0,026 / 0,028) 0,029	(0,144 / 0,126) 0,126	(0,020 / 0,021) 0,021	(0,090 / 0,105) 0,106
MLP Bestand	(0,029 / 0,030) 0,031	(0,197 / 0,123) 0,123	(0,022 / 0,025) 0,022	(0,140 / 0,112) 0,099
mv.MLP Bestand	(0,028 / 0,029) 0,029	(0,176 / 0,117) 0,112	(0,021 / 0,026) 0,022	(0,114 / 0,107) 0,090
mv.MLP.ACF Best.	(0,023 / 0,023) 0,029	(0,109 / 0,095) 0,088	(0,022 / 0,020) 0,023	(0,084 / 0,085) 0,084
Verfahren	Rang <i>mwMLLC</i>		Rang <i>mdMLLC</i>	
	Langsamdreher	Schnelldreher	Langsamdreher	Schnelldreher
NF1 Bestand	(12 / 12) 11	(12 / 12) 12	(11 / 12) 12	(12 / 11) 8
NF2 Bestand	(11 / 11) 12	(7 / 11) 11	(12 / 11) 11	(11 / 12) 6
MA Bestand	(10 / 9) 3	(8 / 10) 6	(10 / 9) 10	(4 / 10) 5
S.ES Bestand	(8 / 6) 2	(5 / 9) 3	(9 / 7) 8	(3 / 8) 3
DT.ES Bestand	(5 / 5) 9	(10 / 3) 9	(7 / 3) 6	(7 / 5) 10
EXP.ES Bestand	(6 / 3) 7	(4 / 4) 2	(5 / 5) 4	(6 / 4) 2
EXP.ARIMA Best.	(7 / 7) 1	(3 / 6) 5	(6 / 6) 2	(10 / 7) 11
EXP.ES&ARIMA	(9 / 4) 8	(9 / 8) 10	(8 / 4) 3	(8 / 9) 12
mv.EXP.ES Bestand	(2 / 2) 6	(2 / 7) 8	(1 / 2) 1	(2 / 2) 9
MLP Bestand	(4 / 10) 10	(11 / 5) 7	(3 / 8) 7	(9 / 6) 7
mv.MLP Bestand	(3 / 8) 5	(6 / 2) 4	(2 / 10) 5	(5 / 3) 4
mv.MLP.ACF Best.	(1 / 1) 4	(1 / 1) 1	(4 / 1) 9	(1 / 1) 1

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

Dies wird auch in der Analyse der Zeitreihen, kategorisiert nach ihrer Umsatzbedeutung als A-, B- oder C-Artikel, bestätigt, deren Dispositionskosten in Tabelle 6.33 aufgeführt sind.¹ Erneut ist das mv.MLP.ACF das beste Verfahren auf Trainings-, Validierungs- und Testmenge mit den niedrigsten Dispositionskosten nach *mwMLLC* und *mdMLLC* und den höchsten Rängen mit (1 / 1) 1 auf den A-Artikeln. Auch auf den B-Artikeln der zweithöchsten Umsatzklasse zeigt das mv.MLP.ACF eine hohe Dispositionsgüte nach *mwMLLC* von (1 / 1) 1, was jedoch durch den mittleren Rang auf den Testdaten von (4 / 1) 6 nach dem *mdMLLC* nur bedingt bestätigt wird. Bei den C-Artikeln zeigt das Verfahren dagegen nur die geringe Leistung eines elften Ranges nach *mwMLLC* (1 / 1) 11 sowie (3 / 2) 6 nach *mdMLLC*, was erneut auf den Einfluss der sporadischen Zeitreihen zurückzuführen ist. Damit bestätigt sich, dass multivariate MLP mit asymmetrischer Kostenfunktion im Vergleich zu Expertensystemen besser zur Disposition von Waren mit regelmäßigem Absatzverlauf und höherem Warenumsatz geeignet sind als zur Disposition an unregelmäßigen Zeitreihen mit geringem Warenumsatz. Die Naiven Verfahren zeigen eine konstant niedrige Dispositionsgüte mit hohen Kosten nach *mwMLLC* und *mdMLLC*. Alle anderen Verfahren zeigen auch nach der ABC-Klassifizierung erneut keine einheitliche Überlegenheit auf einer der Kategorien von Zeitreihen.

¹ Die Artikel der Schnelldreher entsprechen den A-Artikeln, während die Langsamdreher nach Gesamtumsatz in B- und C-Artikel unterteilt werden.

Tabelle 6.33: *mwMLLC* und *mdMLLC* über alle Zeitreihen nach Artikelklassifikation

<i>mwMLLC</i>			
Verfahren	A-Artikel	B-Artikel	C-Artikel
NF1 Bestand	(0,235 / 0,175) 0,145	(0,052 / 0,048) 0,041	(0,029 / 0,026) 0,026
NF2 Bestand	(0,178 / 0,168) 0,130	(0,051 / 0,047) 0,047	(0,029 / 0,023) 0,023
MA Bestand	(0,178 / 0,140) 0,115	(0,043 / 0,038) 0,038	(0,025 / 0,023) 0,021
S.ES Bestand	(0,173 / 0,131) 0,109	(0,041 / 0,037) 0,038	(0,025 / 0,021) 0,020
DT.ES Bestand	(0,180 / 0,121) 0,128	(0,041 / 0,038) 0,044	(0,023 / 0,020) 0,020
EXP.ES Bestand	(0,172 / 0,122) 0,102	(0,041 / 0,038) 0,042	(0,023 / 0,020) 0,019
EXP.ARIMA Bestand	(0,171 / 0,124) 0,115	(0,042 / 0,039) 0,037	(0,023 / 0,020) 0,019
EXP.ES&ARIMA Bestand	(0,179 / 0,130) 0,130	(0,042 / 0,037) 0,043	(0,024 / 0,021) 0,019
mv.EXP.ES Bestand	(0,144 / 0,126) 0,126	(0,032 / 0,036) 0,042	(0,021 / 0,021) 0,019
MLP Bestand	(0,197 / 0,123) 0,123	(0,038 / 0,041) 0,044	(0,022 / 0,021) 0,020
mv.MLP Bestand	(0,176 / 0,117) 0,112	(0,036 / 0,040) 0,042	(0,022 / 0,020) 0,019
mv.MLP.ACF Bestand	(0,109 / 0,095) 0,088	(0,029 / 0,029) 0,037	(0,018 / 0,018) 0,023
<i>Rang mwMLLC</i>			
Verfahren	Rang A-Artikel	Rang B-Artikel	Rang C-Artikel
NF1 Bestand	(12 / 12) 12	(12 / 12) 5	(11 / 12) 12
NF2 Bestand	(7 / 11) 11	(11 / 11) 12	(12 / 11) 10
MA Bestand	(8 / 10) 6	(10 / 7) 4	(9 / 10) 9
S.ES Bestand	(5 / 9) 3	(6 / 4) 3	(10 / 9) 7
DT.ES Bestand	(10 / 3) 9	(5 / 6) 11	(5 / 5) 6
EXP.ES Bestand	(4 / 4) 2	(7 / 5) 7	(6 / 2) 5
EXP.ARIMA Bestand	(3 / 6) 5	(8 / 8) 2	(7 / 3) 1
EXP.ES&ARIMA Bestand	(9 / 8) 10	(9 / 3) 9	(8 / 6) 4
mv.EXP.ES Bestand	(2 / 7) 8	(2 / 2) 8	(2 / 7) 2
MLP Bestand	(11 / 5) 7	(4 / 10) 10	(3 / 8) 8
mv.MLP Bestand	(6 / 2) 4	(3 / 9) 6	(4 / 4) 3
mv.MLP.ACF Bestand	(1 / 1) 1	(1 / 1) 1	(1 / 1) 11
<i>mdMLLC</i>			
Verfahren	A-Artikel	B-Artikel	C-Artikel
NF1 Bestand	(0,163 / 0,141) 0,100	(0,035 / 0,048) 0,034	(0,025 / 0,024) 0,019
NF2 Bestand	(0,148 / 0,145) 0,099	(0,036 / 0,038) 0,035	(0,027 / 0,021) 0,019
MA Bestand	(0,114 / 0,126) 0,090	(0,029 / 0,033) 0,026	(0,021 / 0,021) 0,018
S.ES Bestand	(0,107 / 0,122) 0,087	(0,027 / 0,034) 0,027	(0,021 / 0,018) 0,017
DT.ES Bestand	(0,129 / 0,109) 0,109	(0,026 / 0,031) 0,030	(0,019 / 0,019) 0,017
EXP.ES Bestand	(0,127 / 0,109) 0,087	(0,027 / 0,033) 0,030	(0,020 / 0,018) 0,019
EXP.ARIMA Bestand	(0,143 / 0,117) 0,114	(0,026 / 0,030) 0,030	(0,019 / 0,018) 0,016
EXP.ES&ARIMA Bestand	(0,134 / 0,123) 0,118	(0,028 / 0,028) 0,028	(0,020 / 0,019) 0,019
mv.EXP.ES Bestand	(0,090 / 0,105) 0,106	(0,024 / 0,031) 0,025	(0,018 / 0,018) 0,018
MLP Bestand	(0,140 / 0,112) 0,099	(0,025 / 0,034) 0,027	(0,016 / 0,018) 0,015
mv.MLP Bestand	(0,114 / 0,107) 0,090	(0,025 / 0,033) 0,026	(0,016 / 0,016) 0,015
mv.MLP.ACF Bestand	(0,084 / 0,085) 0,084	(0,025 / 0,024) 0,028	(0,016 / 0,018) 0,018
<i>Rang mdMLLC</i>			
Verfahren	Rang A-Artikel	Rang B-Artikel	Rang C-Artikel
NF1 Bestand	(12 / 11) 8	(11 / 12) 11	(11 / 12) 12
NF2 Bestand	(11 / 12) 6	(12 / 11) 12	(12 / 11) 11
MA Bestand	(4 / 10) 5	(10 / 7) 2	(9 / 10) 8
S.ES Bestand	(3 / 8) 3	(8 / 10) 5	(10 / 7) 4
DT.ES Bestand	(7 / 5) 10	(6 / 4) 8	(6 / 8) 5
EXP.ES Bestand	(6 / 4) 2	(7 / 6) 10	(8 / 4) 9
EXP.ARIMA Bestand	(10 / 7) 11	(5 / 3) 9	(5 / 5) 3
EXP.ES&ARIMA Bestand	(8 / 9) 12	(9 / 2) 7	(7 / 9) 9
mv.EXP.ES Bestand	(2 / 2) 9	(1 / 5) 1	(4 / 3) 7
MLP Bestand	(9 / 6) 7	(3 / 9) 4	(1 / 6) 2
mv.MLP Bestand	(5 / 3) 4	(2 / 8) 3	(2 / 1) 1
mv.MLP.ACF Bestand	(1 / 1) 1	(4 / 1) 6	(3 / 2) 6

Ergebnisse werden in der Form „(Fehler Trainingsmenge / Fehler Validierungsmenge) Fehler Testmenge“ präsentiert.

6.3.3 Zusammenhang zwischen Prognosegenauigkeit und Dispositionskosten

Prognosegenauigkeit und Dispositionsgüte wurden in den vorangegangenen Abschnitten separat analysiert und die konsistent hohe Güte unterschiedlicher Varianten der NN nachgewiesen. Unter ausschließlicher Betrachtung als Prognoseverfahren zeigten NN eine statistisch signifikant bessere Prognosegenauigkeit nach $mwMAE$ auf den Absatzzeitreihen der Warenautomaten als alle anderen statistischen Benchmark-Verfahren. Dagegen wurde in der Analyse der Dispositionsgüte nach den mittleren Kosten des $mwMLCC$ aus Überbeständen und Fehlmengen deutlich, dass dieselben Prognoseverfahren einschließlich der vorher überragenden NN keine statistisch signifikant unterschiedlichen Dispositionskosten verursachten.¹ Die Diskrepanz zwischen den statistisch signifikanten Unterschieden der Prognosegüte und den nur insignifikanten Unterschieden in der Dispositionsgüte – trotz einheitlicher Berechnung der Sicherheitsbestände und Bestellmengen – erfordert eine separate Analyse.

Hierzu wird nachfolgend die Übereinstimmung von Prognosegenauigkeit und Dispositionskosten je Zeitreihe und Verfahren analysiert. In Abbildung 6.25 wird die Prognosegenauigkeit der Verfahren nach $mwMAE$ ihren Dispositionskosten nach $mwMLCC$ als Punkte in einem Streudiagramm über alle Zeitreihen gegenübergestellt.

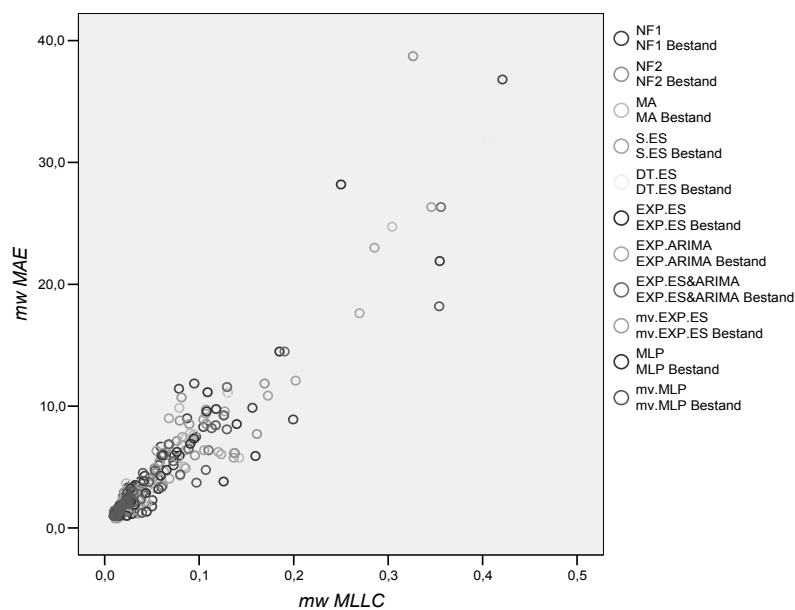


Abbildung 6.25: Streudiagramm der Prognosefehler nach $mwMAE$ und der Dispositionskosten nach $mwMLCC$ je Prognoseverfahren

Die Abbildung verdeutlicht eine positive Korrelation zwischen dem mittleren Prognosefehler und den mittleren Dispositionskosten aller Verfahren über alle Absatzstellen und Zeitreihen.² Somit führen Prognoseverfahren mit einem geringen $mwMAE$ auf den Testdaten ebenfalls zu niedrigen Dispositionskosten nach dem $mwMLCC$, auch wenn bei den Dispositionskosten keine statistisch signifikanten Unterschiede festgestellt werden konnten. Die Prognosegenau-

¹ Das neuartige Verfahren der NN, trainiert mit asymmetrischen Kostenfunktionen mv.MLP.ACF ohne Pendant in den Prognoseverfahren, erzielte zwar die niedrigsten Dispositionskosten, diese unterschieden sich jedoch aufgrund der inkonsistenten Rangreihung über alle Zeitreihen mit sporadischen, unregelmäßigen und regelmäßigen Verläufen ebenfalls nicht statistisch signifikant von den anderen Verfahren.

² Die Darstellung wird durch überproportional hohe Fehler- und Kostenwerte der drei Zeitreihen IN-1-S2, IN-1-S0 und IN-2-S6 dominiert, die im oberen rechten Quadranten dargestellt sind und zu einer verzerrten Darstellung führen. Da sie jedoch Rückschlüsse auf die Konsistenz der Ergebnisse nach den Fehlermaßen der Mittelwerte im Gegensatz zu den Medianen erlauben, werden sie dennoch abgebildet.

igkeit kann daher bedingt als Indikator für die Güte in der Entscheidungsunterstützung zur Disposition verwendet werden, wodurch die Erkenntnisse aus vorangehenden Verfahrenvergleichen zur Prognosegenauigkeit auch Rückschlüsse auf die Güte der Verfahren in der Warendisposition und der Bestellmengenrechnung zulassen.

In Abbildung 6.26 sind die vergleichsweise geringen Prognosefehler und Dispositionskosten aller Verfahren der NN im unteren linken Quadranten des Streudiagramms verdeutlicht. Dies lässt jedoch noch keine Rückschlüsse auf eine Überlegenheit gegenüber anderen Verfahren zu, da die relative Güte je Zeitreihe jeweils im Vergleich zu anderen Verfahren und somit anhand der relativen Rangordnung je Zeitreihe zu bestimmen ist. Hierzu werden nachfolgend Streudiagramme des Rangs jedes Verfahrens nach den Prognosefehlern *Rang mwMAE* im Vergleich zum Rang nach den Dispositionskosten *Rang mwMLLC* gebildet und in Abbildung 6.26 dargestellt.¹

Die Naiven Verfahren NF1 und NF2 erreichen konsistent einen niedrigen Rang nach den Prognosefehlern und Dispositionskosten, dargestellt durch die Konzentration der Punktwolken im oberen rechten Quadranten der jeweiligen Darstellung. Das Verfahren der MA zeigt dagegen eine leicht bessere Leistung um den achten und neunten Rang nach Prognosefehlern und Dispositionskosten, jedoch mit höherer Streuung als die Naiven Verfahren. Dazu erzielen die gleitenden Mittelwerte vereinzelt einen höheren, nämlich zweiten bis fünften Rang nach den Prognosefehlern trotz abschließend hoher Dispositionskosten. Die Prognoseverfahren mit aggregierter Modellselektion S.ES und DT.ES weisen eine gesteigerte Prognose- und Dispositionsgüte gegenüber den Naiven Verfahren auf, jedoch mit einer zunehmenden Streuung über alle Ränge. Dabei zeigt das DT.ES eine positive Korrelation zwischen einigen oberen Rängen nach Prognosefehlern und mittleren Rängen nach der Dispositionsgüte. Diese positive Korrelation durch eine stärkere Übereinstimmung zwischen der Rangordnung nach den Prognosefehlern und der Rangordnung nach Dispositionskosten wird insbesondere durch die Verfahren der Expertensysteme mit individueller Modellselektion bestätigt, die gleichmäßig verteilt sowohl obere, mittlere als auch untere Ränge erzielen. Dabei zeigen die Expertensysteme der univariaten und der multivariaten Exponentiellen Glättung EXP.ES und mv.EXP.ES konsistent obere Ränge mit niedrigen Prognosefehlern und niedrigen Dispositionskosten, während Expertensysteme mit ARIMA-Verfahren eine stärkere Verteilung über mittlere und niedrigere Ränge aufweisen.

Die Verfahren der NN in Form des MLP und des mv.MLP zeigen dagegen eine starke Konzentration der höchsten Ränge nach Prognosefehlern und Dispositionskosten im unteren linken Quadranten der Darstellungen, trotz der nur insignifikanten Unterschiede der Dispositionsgüte im Vergleich zu anderen Verfahren. Sie bilden so einen klaren Gegensatz zu den Streudiagrammen der Naiven Verfahren sowie der Expertensysteme mit deutlich stärker verteilten Punktwolken und somit stärkeren Diskrepanzen zwischen Prognosegenauigkeit und Dispositionskosten.

¹ Die Streudiagramme der Ränge zeigen ein mehrfaches Auftreten gleicher Rang-Paare auf mehreren Zeitreihen durch überlagerte Punkte mit zunehmendem Durchmesser an, so dass größere Kreise mehrfache Ausprägungen übereinstimmender Rangordnungen anzeigen. Punkte auf der Diagonalen zeigen eine Übereinstimmung der Prognosegüte mit der Dispositionsgüte an, so dass eine hohe Prognosegenauigkeit auch zu einer hohen Dispositionsgenauigkeit führt, und umgekehrt eine niedrige Prognosegüte zu einer niedrigen Dispositionsgüte. Verfahren mit einer hohen Prognosegenauigkeit und niedrigen Dispositionskosten zeigen größere Punktwolken im unteren linken Quadranten eines Diagramms, Verfahren mit einer konstant schlechten relativen Leistung entsprechend im oberen rechten Quadranten. Das Streudiagramm zeigt so beispielsweise in der obersten linken Abbildung den Rang des Verfahrens NF1 nach *mwMAE* im Vergleich zu seinem Rang nach *mwMLLC* an. Für das Verfahren NF1 nehmen drei Zeitreihen nach *mwMAE* den ersten Rang als bestes Prognoseverfahren ein. Sie erzielen damit jedoch lediglich eine geringe Dispositionsgüte nach *mwMLLC* und einen neunten, elften und zwölften Rang. Dagegen nimmt eine große Anzahl von Zeitreihen für den NF1-Prognosefehler einen zehnten oder elften Rang ein, welcher ebenfalls durch einen zehnten oder elften Rang nach den Dispositionskosten *mwMLLC* bestätigt wird.

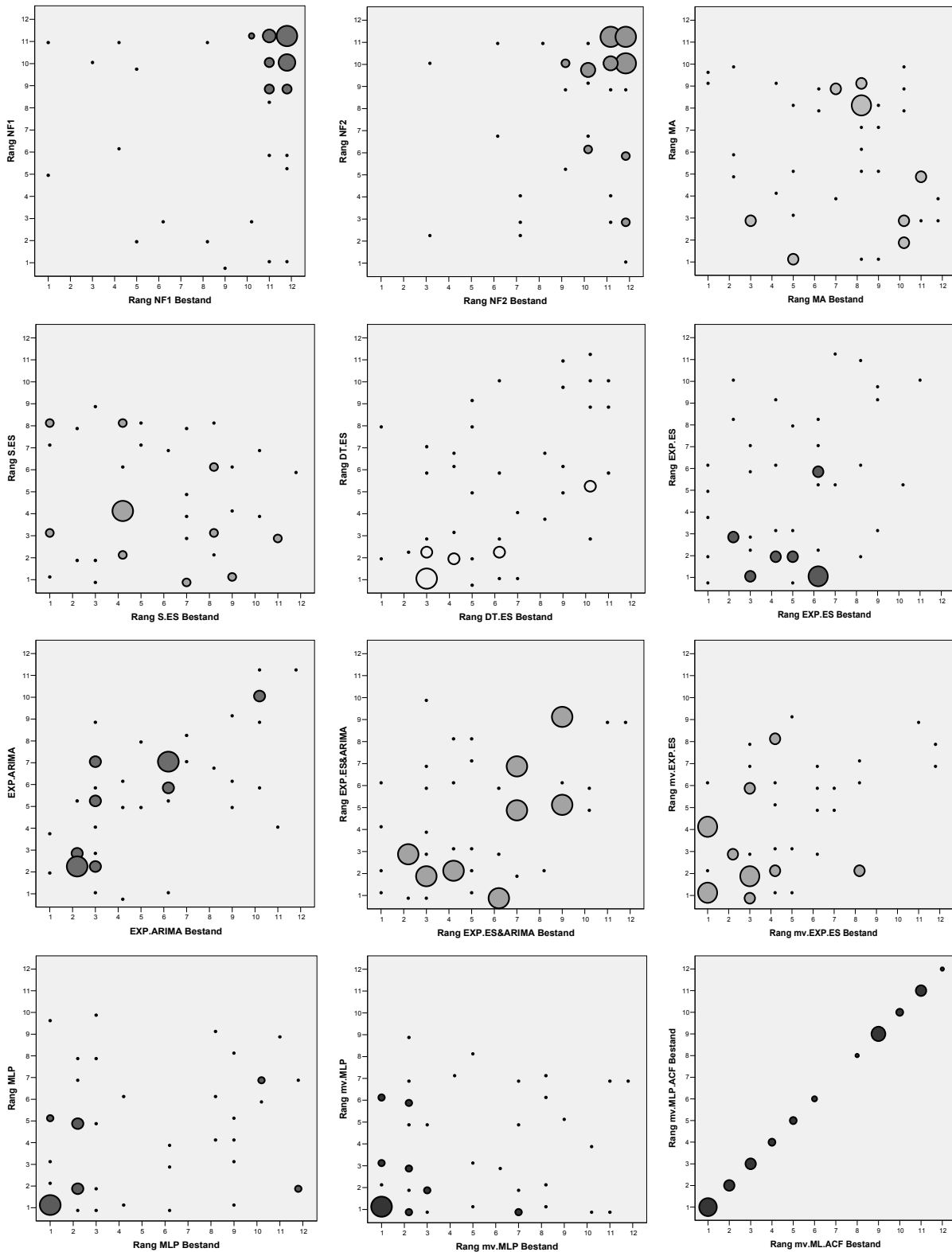


Abbildung 6.26: Streudiagramme der Prognosefehler nach *Rang mwMAE* und der Dispositionskosten nach *Rang mwMLLC* je Prognoseverfahren

Das Verfahren der NN mit asymmetrischen Kostenfunktionen *mv.MLP.ACF* besitzt kein korrespondierendes Prognoseverfahren, so dass kein Streudiagramm abgebildet werden kann. Zur konsistenten Visualisierung der Streudiagramme wird die Rangordnung nach *mwMLLC* mit sich selbst korreliert und so eine Verteilung der Ränge abgebildet.¹ Die Darstellung

¹ Die Abbildung verdeutlicht somit die Verteilung der Ränge nach *mwMLLC*, nicht jedoch ihre Korrelation.

verdeutlicht, dass mv.MLP.ACF aufgrund seiner geringen Dispositionskosten überwiegend hohe Ränge und nur mit abnehmender Häufigkeit niedrigere Ränge erzielt. Allerdings belegt das neuartige Verfahren auch mehrfach hintere Ränge, welche den Zeitreihen mit sporadischem Absatz zugeordnet werden können.

Darüber hinaus verdeutlichen die Diagramme aller Verfahren die erhebliche Streuung der rangbasierten Analyse durch die stark verteilten Punktwolken mit nur vereinzelt Konzentrationen oder deutlichen Korrelationen. Selbst die Naiven Verfahren und die Verfahren der NN zeigen vereinzelt niedrige Ränge der Prognosefehler mit hohen Dispositionskosten und umgekehrt, was die geringe Korrelation von Prognosegenauigkeit und Dispositionskosten auf den Testdaten einiger Zeitreihen begründet. Diese Ergebnisse bestätigen daher Erkenntnisse vorangegangener Studien – insbesondere hinsichtlich der Notwendigkeit einer Beurteilung von Verfahren über mehrere repräsentative Zeitreihen und unter Verwendung multipler Gütekriterien gemäß der Nullhypothesen H_2 , H_3 und H_4 – und erlauben so eine Bestätigung der Validität und Reliabilität des vorliegenden Experimentdesigns.

Die positive, aber aufgrund der Streuung der Ränge nur insignifikante Korrelation zwischen Prognosefehlern und Dispositionskosten verdeutlicht, dass die Prognosegenauigkeit nicht unmittelbar für die Beurteilung eines Verfahrens in der Entscheidungsfindung verwendet werden kann. Somit kann die Nullhypothese H_4 verworfen werden, da die Genauigkeit eines Verfahrens zur Absatzprognose nicht immer als Indikator für die Genauigkeit des Verfahrens in der Warendisposition dient. Vielmehr ist diese im Einzelnen zu untersuchen und die Prognosegenauigkeit den Dispositionskosten gegenüberzustellen.

7 Schlussbetrachtung

7.1 Zusammenfassung

Die vorliegende Arbeit untersucht die Verfahrensklasse der künstlichen Neuronen Netze zur Prognose und erweitert sie durch die Verwendung asymmetrischer Kostenfunktionen zur simultanen Bedarfsprognose und Bestellmengenrechnung im Sinne einer integrierten Warendisposition. Die Fähigkeit der Neuronen Netze wird in einer Vergleichsstudie auf synthetischen Zeitreihen ermittelt und anschließend in einem objektiven empirischen Verfahrensvergleich mit etablierten Verfahren der Bedarfsprognose bestätigt.

Die Einordnung der Warendisposition in die Prozesse des Handels verdeutlicht ihre zentrale Rolle in der Koordination der Warenbestände durch die Bestimmung von Bestellmengen und Zeitpunkten. Dabei sind auf Grund der Bedarfsstrukturen im Handel stochastische instationäre Modelle der Bedarfsprognose und Bestellmengenrechnung erforderlich, die eine genaue und unverzerrte Bedarfsprognose erfordern. Die Genauigkeit der Bedarfsprognose bestimmt dabei nachhaltig die Kosten der Lagerhaltung, weshalb ihr eine besondere Bedeutung zukommt. Dazu können unterschiedliche Prognoseverfahren zur Abbildung von Erklärungs- und Prognosemodellen unterschiedlicher Komplexität zum Einsatz kommen. Die Verfahrensklasse der Neuronen Netze erweitert das Arsenal statistischer Prognoseverfahren für eine datengetriebene, nichtlineare und semi-parametrische Bedarfsprognose.

Die Grundlagen der Verfahrensklasse der Neuronen Netze wurden intensiv aufgearbeitet. Dabei wurde gezeigt, dass Neuronen Netze weder aufgrund der Signalverarbeitung in Neuronen noch der Informationsverarbeitung im Netzwerk von Neuronen oder der Algorithmen zur Parametrisierung auf die Verwendung quadratischer Zielfunktionen beschränkt sind. Das Architekturkonzept Neuronen Netze erlaubt somit die Erweiterung auf die Verwendung alternativer Zielfunktionen, was in der wissenschaftlichen Literatur bisher vernachlässigt wurde. Die Analyse asymmetrischer Entscheidungskosten der Warendisposition erlaubt die Entwicklung asymmetrischer Kostenfunktionen nach GRANGER [1969], die eine Veränderung der Lernziele und des Verhaltens der Verfahrensklasse bewirkt, unter Erweiterung des Lernalgorithmus der Backpropagation auf die Minimierung generalisierter Kostenfunktionen.

Die Fähigkeit von Neuronen Netzen zur Minimierung asymmetrischer Kostenfunktion wird anhand einer Studie auf synthetischen Zeitreihen sowie einer Stichprobe empirischer Zeitreihen des Absatzes an Warenautomaten überprüft. Das Neuronen Netz prognostiziert unter Verwendung der asymmetrischen Kostenfunktion einen kostenminimalen Warenbestand anstelle eines fehlerminimalen Bedarfswertes. Dies entspricht der Bestimmung beliebiger Quantile einer Verteilungsfunktion durch die Minimierung einer asymmetrischen Kostenfunktion gemäß der Theorie generalisierter Kostenfunktionen nach GRANGER [1969]¹ und bestätigt den formalen Beweis der Fähigkeit von NN zur Approximation beliebiger Quantile einer Verteilungsfunktion nach WHITE [1992] erstmals empirisch. Die jeweiligen Quantile können unmittelbar anhand der asymmetrischen Kosten der Überbestände und Unterbestände spezifiziert und anhand der erzielten Servicegrade gemessen werden, ein Verfahren, welches konsistent mit den Erkenntnissen der Forschungsbereiche der Warendisposition und der Absatzprognose ist. Es wurde gezeigt, dass MLPs in der Lage sind, die Quantile auf unterschiedlichen Zeitreihenmustern stationärer sowie saisonaler Zeitreihen der Normalverteilung, der Negativen Binomialverteilung oder der Poissonverteilung mit unterschiedlicher Streuung

¹ Auf eine Gültigkeit zu der generalisierten Quantil-Regression nach KOENKER [2005] kann nicht geschlossen werden, da diese auch die Veränderung der Quantile im Zeitablauf im Sinne von ARCH- oder GARCH-Modellen zulässt. Dies ist in weiterführenden Studien mit längeren Prognosehorizonten zu untersuchen.

und ohne A-priori-Kenntnis oder separate Bestimmung der Verteilungsfunktion zu ermitteln. Somit sind weder eine vorangehende Untersuchung der Beobachtungswerte noch explizite Verteilungsannahmen wie für statistische Verfahren erforderlich, z. B. durch statistische Tests zur Überprüfung der Prämisse normalverteilter Beobachtungswerte. Das MLP erfordert lediglich die Spezifikation der Kosten einer Über- und Unterprognose in einer asymmetrischen Kostenfunktion. Die Entscheidungskosten aus den so erzielten Servicegraden der MLPs sind signifikant geringer als die herkömmlicher, statistischer Prognoseverfahren sowie führender Expertensysteme zur Prognose. Die Verbesserungen in der Effizienz der Warenbestände werden konsistent erzielt, sowohl auf den historischen Daten zur Parametrisierung als auch auf unbekanntem Testdaten aller Zeitreihenmuster.

Die dargestellte Funktionalität und die Güte der Ergebnisse rechtfertigen den Einsatz von MLPs mit asymmetrischen Kostenfunktionen zur Disposition auf Zeitreihen bei instationärem Bedarf unterschiedlicher Streuung und beliebiger Verteilungsformen des Bedarfs. In der empirischen Anwendung kann somit nach der Erhebung repräsentativer Dispositionskosten eine einzige Konfiguration eines MLP mit asymmetrischen Kostenfunktionen produktübergreifend zur effizienten Disposition eingesetzt werden. Neben einer Kostensenkung bieten MLPs aber noch weitere Vorteile bei der Warendisposition, indem sie die Teilschritte zur Ermittlung kostenminimaler Warenbestände vereinfachen und in einen datengetriebenen Trainingsprozess integrieren. Insbesondere erfordern sie weder Ex-ante-Annahmen oder statistische Tests über die Verteilung der Bedarfe bzw. der Prognoseresiduen noch eine iterative Modellbildung durch Experten oder Expertensysteme. Die Modellierung erfolgt zudem näher an der betrieblichen Realität, da lediglich die betriebswirtschaftlichen Kosten der Entscheidungsfindung und die Daten des vergangenen Bedarfs verwendet werden.

Die Überlegenheit der NN in der Prognosegenauigkeit kann jedoch bei der Untersuchung der Bestandsrechnung zur Warendisposition nur bedingt bestätigt werden. Es wird deutlich, dass die Berechnung von Sicherheitsbeständen diverse Einflüsse und Ineffizienzen selbst von signifikant unterlegenen Prognoseverfahren kompensieren und nur unterproportional und insignifikant verschlechterte Warenbestände bewirken. Dies kann eine mögliche Erklärung für die weite Verbreitung und nahezu ausschließliche Verwendung von einfachen Verfahren in der Praxis liefern, welche die Weiterentwicklungen von komplexen Verfahrensansätzen zur Prognose und ihre Bestätigung in theoretischen Verfahrensvergleichen zu vernachlässigen scheint. Wenngleich komplexere Prognoseverfahren signifikante Verbesserungen in der Prognosegenauigkeit erlauben, führen diese in der vorliegenden Studie nur zu unterproportionalen Einsparungen in den Kosten der Entscheidungsfindung zur Warendisposition.¹

7.2 Ausblick

Die Erkenntnisse der vorliegenden Arbeit tragen in mehrfacher Hinsicht zur gegenwärtigen wissenschaftlichen Diskussion von NN zur Prognose bei. Die Möglichkeiten des Trainings von NN mit asymmetrischen Zielfunktionen und betrieblichen Kostenfunktionen wird hier erstmalig entwickelt und anhand einer konkreten Anwendung diskutiert. Hierzu werden die Einsatzmöglichkeiten bestehender statistischer Zielfunktionen zum Training von NN in dieser Untersuchung erstmals umfassend aufgearbeitet. Dies beinhaltet die Auswertung einer Vielzahl von Literaturquellen der interdisziplinären Forschung und Anwendung von NN in der Statistik, der Ökonometrie, der Informatik und in den Ingenieurwissenschaften.

¹ Es erscheint zudem plausibel, dass der Grad der Einsparungen durch den Grad der Asymmetrie der Kosten bestimmt wird, welche durch die Höhe der Sicherheitsbestände eine selbstregulierende Wirkung erzielen. Dies ist nachfolgend in weiterführenden Verfahrensvergleichen zu untersuchen, die über die Entwicklung des neuartigen Ansatzes zum Training von NN mit asymmetrischen Zielfunktionen hinaus gehen.

Die Erweiterung der Theorie der Neuronalen Netze auf das Erlernen betriebswirtschaftlicher Zielfunktionen ermöglicht neuartige Anwendungen über die Disziplinen der Betriebswirtschaftslehre und der Prognostik hinaus, sofern eine empirische Kostenfunktion bestimmt werden kann. Asymmetrischen Zielfunktionen ermöglichen der Verfahrensklasse der NN die Abbildung eines neuartigen Verhalten in der Interaktion mit Daten, welche ihre Anwendung erstmalig auf die unmittelbare und realitätsnahe Entscheidungsunterstützung allein aus vergangenen Beobachtungen ermöglicht. Aufgrund der neuartigen Entwicklung und Anwendung nichtsymmetrischer Zielfunktionen für die Verfahren der Neuronalen Netze ist auch ein Erkenntniswert für Informatik, Ingenieurwissenschaften und Statistik anzunehmen. Ferner lässt die vorgestellte Anwendung eine unmittelbare Erprobung in der Handelspraxis zu und kann aufgrund der gestiegenen Dispositionsgüte zu erheblichen Kostensenkungen und einer Vereinfachung der Prognose- und Dispositionsprozesse führen.

Auch wenn die Entwicklung und die systematische Beurteilung einer allgemeinen Methodologie zur Modellierung von NN für alle Formen der Prognose den Rahmen dieser Untersuchung übersteigt, wird dennoch ein prototypisches Vorgehenskonzept für die valide und reliable Modellierung von NN auf beliebigen Zeitreihenmustern vorgestellt. Diese Voraussetzung für die Anwendung von NN zu Zwecken der Prognose sowie der Disposition bietet Potenzial für weiterführende Forschung zur Modellierung und Automatisierung von NN in vielen Anwendungsbereichen. Die Arbeit zeigt somit Ansätze zur Überwindung einer Hauptkritik gegen den Einsatz von NN in der betriebswirtschaftlichen Forschung und Anwendung auf.

Des Weiteren belegt diese Untersuchung die Einsatzfähigkeit von NN bei der Absatzprognose im Handel erstmalig in einem umfassenden, empirischen Verfahrenvergleich gegenüber etablierten Prognoseverfahren auf der Basis eines validen und reliablen Versuchsaufbaus der Statistik. Der Themenkomplex „NN zu Zwecken der Absatzprognose“ ist ein Forschungsfeld, das in der Vergangenheit weniger durch empirische Untersuchungen als durch Erweiterungen der Methoden und Verfahren wissenschaftlich bearbeitet wurde. Jedoch erst die systematische, empirische Beurteilung erlaubt zuverlässige Schlüsse hinsichtlich der Dispositions- und Prognosegüte von NN in der betrieblichen Praxis - im Gegensatz zu einseitigen Veröffentlichungen mit suboptimalen Versuchsaufbauten der Vergleichsstudien. Somit trägt diese Untersuchung auch dazu bei, Qualität und Aussagekraft zukünftiger Verfahrensvergleiche zu verbessern.

Der konsistente Vergleich der Verfahrensgüte sowohl mithilfe statistischer Prognosefehler als auch mittels betrieblicher Dispositionskosten erlaubt darüber hinaus empirische und theoretische Erkenntnisse zur Interdependenz und Sensitivität der Teilfunktionen der Bedarfsrechnung und Bestellmengenrechnung auf die Güte der Warendisposition. Diese Interdependenz wird bei der üblichen disjunkten Untersuchung der Bedarfsprognose unter Vernachlässigung des nachfolgenden Entscheidungsmodells oder bei der Bestellmengenrechnung unter Annahme einer bekannten, konstanten Bedarfsverteilung nicht erkannt. Die Untersuchung trägt so zu den Erkenntnissen des Bestandsmanagements und der Lagerlogistik hinsichtlich des Einflusses der Prognosegenauigkeit bei.

Das Untersuchungsgebiet dieser Arbeit beschränkt sich auf Betriebe des Handels. Die Anwendung von Neuronalen Netzen mit asymmetrischen Kostenfunktionen auf unmittelbar verwandte Problemstellungen, wie die Disposition an Geldautomaten, oder auch auf andere Formen des stationären Einzelhandels liegt nahe. Aufgrund der branchenübergreifenden Problemstellung von Prognose und Disposition sowie der unmittelbaren Vernetzung von Industrie und Handel entlang der Supply Chain sind die Erkenntnisse aber auch auf andere

Betriebsformen und Branchen übertragbar und insbesondere für Betriebe der Konsumgüterindustrie von Interesse.²

Zukünftige Erweiterungen der Netzwerkarchitektur auf mehrere Ausgabeneuronen unter Verwendung jeweils unterschiedlicher asymmetrischer Kostenfunktionen können die Verfahrensklasse der Neuronalen Netze weiter zu einer nichtlinearen, semi-parametrischen Intervallprognose und bis hin zur Verteilungsprognose erweitern.

In der Zukunft sind jedoch zunächst die Erkenntnisse dieser Arbeit durch Vergleichsstudien auf repräsentativen Stichproben von empirischen Zeitreihen kritisch zu überprüfen. Bis auf weiteres sind die getroffenen Erkenntnisse nur für die vorliegenden Datensätze der synthetischen Zeitreihen sowie die Auswahl empirischer Zeitreihen des Handels gültig, und nicht unbegrenzt generalisierbar. Die so gewonnenen theoretischen Erkenntnisse aus der Analyse weiterer Zeitreihen müssen dabei im Zuge spezifischer Problemstellungen und Anwendungen, unter Verwendung der tatsächlichen Kosten der Entscheidung, empirisch durch einen wissenschaftlich fundierten Verfahrensvergleich für jeden Einzelfall auf ihre Güte überprüft werden. Erst dies erlaubt den Rückschluss auf eine allgemeine Einsatzfähigkeit von künstlichen Neuronalen Netzen mit asymmetrischen Kostenfunktionen in der betrieblichen Entscheidungsunterstützung.

² Dies wird ebenfalls durch die zitierten Studien und Aktivitäten der Industrieverbände hinsichtlich Fehlmen- gen und Überbeständen belegt.

Literaturverzeichnis

- Aberdeen Group (2004) The Demand Management Benchmark Report, Technical Report, Nr. 2004-3, Aberdeen Group: Boston, 42 Seiten.
- Aberdeen Group (2005) Best Practices in Sales & Operations Planning, Technical Report, Nr. 2005-6, Aberdeen Group: Boston, 43 Seiten.
- Abramowitz, M. und Stegun, I. A. (1972) Handbook of mathematical functions with formulas, graphs, and mathematical tables, 10. print., Dec. 1972, with corr, United States Department of Commerce: Washington, DC.
- Adya, M. und Collopy, F. (1998) How effective are neural networks at forecasting and prediction? A review and evaluation, in: Journal of Forecasting, Vol. 17, S. 481-495.
- Agarwal, A. (1993) Neural networks and their extensions for business decision making
- Agrawal, N. und Smith, S. A. (1996) Estimating negative binomial demand for retail inventory management with unobservable lost sales, in: Naval Research Logistics, Vol. 43, S. 839-861.
- Ahlburg, D. A. (1992) Error measures and the choice of a forecast method, in: International Journal of Forecasting, Vol. 8, S. 99-100.
- Ahlert, D. (2002) Distributionspolitik: das Management des Absatzkanals, 4. Aufl, G. Fischer: Stuttgart [u.a.].
- Al-Shayji, K. A. und Liu, Y. A. (2002) Predictive modeling of large-scale commercial water desalination plants: Data-based neural network and model-based process simulation, in: Industrial & Engineering Chemistry Research, Vol. 41, S. 6460-6474.
- Alam, P., Booth, D., Lee, K. und Thordarson, T. (2000) The use of fuzzy clustering algorithm and self-organizing neural networks for identifying potentially failing banks: an experimental study, in: Expert Systems With Applications, Vol. 18, S. 185-199.
- Alekseev, K. P. G. und Seixas, J. M. (2002) Forecasting the air transport demand for passengers with neural modelling, in: Proceedings of the VII Brazilian Symposium on Neural Networks, SBRN 2002, Vol., S. 86-91.
- Alex, B. (1998) Künstliche neuronale Netze in Management-Informationssystemen: Grundlagen und Einsatzmöglichkeiten, Gabler: Wiesbaden.
- Alex, B. und Stahlbock, R. (1996) Unterstützung von Managementaufgaben mit NeuralWare Predict, in: NewsServer, Vol. 96, S. 28-30.
- Allen, G. und Fildes, R. (2001) Econometric Forecasting, in: Principles of forecasting: a handbook for researchers and practitioners (Hrsg. Armstrong, J. S.), Kluwer Academic: Boston; London.
- Allende, H., Moraga, C. und Salas, R. (2002) Robust Estimator for the Learning Process in Neural Networks applied in Time Series, in: Proceedings of the ICANN 2002, Vol., S. 1080-1086.
- Alon, I., Qi, M. und Sadowski, R. J. (2001) Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods, in: Journal of Retailing and Consumer Services, Vol. 8, S. 147-156.
- Alotto, P., Magele, C., Brandstatter, B. R., Luschin, M. und Ebner, T. (1999) Approximation of the objective function: multiquadrics versus neural networks, in: Compel-The International Journal For Computation And Mathematics In Electrical And Electronic Engineering, Vol. 18, S. 250-265.
- Alstrom, P. (2001) Numerical computation of inventory policies, based on the EOQ/sigma(x) value for order-point systems, in: International Journal of Production Economics, Vol. 71, S. 235-245.

- Amari, S., Barron, A. R., Bienenstock, E., Geman, S., Breiman, L., McClelland, J. L., Ripley, B. D., Tibshirani, R., Cheng, B. und Titterington, D. M. (1994) Neural Networks - A Review From A Statistical Perspective - Comments And Rejoinders, in: *Statistical Science*, Vol. 9, S. 31-54.
- Amoozegar, F. (1998) Neural-network-based target tracking state-of-the-art survey, in: *Optical Engineering*, Vol. 37, S. 836-846.
- Anderson, J. A. (1995) *An introduction to neural networks*, MIT Press: Cambridge, Mass.
- Anderson, J. A. und Rosenfeld, E. (1988) *Neurocomputing: foundations of research*, MIT Press: Cambridge, Mass.
- Anderson, J. A., Rosenfeld, E. und Pellionisz, A. (1988) *Neurocomputing*, MIT Press: Cambridge, Mass.
- Andler, K. (1929) *Rationalisierung der Fabrikation und optimale Losgröße*, Oldenbourg: München.
- Andras, P. (2002) The equivalence of support vector machine and regularization neural networks, in: *Neural Processing Letters*, Vol. 15, S. 97-104.
- Ansuji, A. P., Camargo, M. E., Radharamanan, R. und Petry, D. G. (1996) Sales forecasting using time series and neural networks, in: *Computers & Industrial Engineering*, Vol. 31, S. 421-424.
- APICS (2005) Doing away with Forecast Error, in: *APICS e-News*, APICS, S. 2-3.
- Arentzen, U. (1993) *Gabler Wirtschafts-Lexikon*, [Taschenbuchausg., 13., vollst. überarb. Aufl. ff.], Gabler: Wiesbaden.
- Arminger, G. und Götz, N. (1999) Asymmetric loss functions for evaluating the quality of forecasts in time series for goods management systems, Technical Report, Nr. 22/1999, Sonderforschungsbereich SFB 475 - Komplexitätsreduktion in Multivariaten Datenstrukturen, Universität Dortmund: Dortmund.
- Armstrong J. S. und Collopy F. (1993) Error measures for generalizing about forecasting methods: Empirical comparisons: *International Journal of Forecasting*, 8 (1), 69-80 (June 1992), in: *Long Range Planning*, Vol. 26, S. 150.
- Armstrong, J. S. (1978) *Long-range forecasting: from crystal ball to computer*, Wiley: New York.
- Armstrong, J. S. (1984) Forecasting By Extrapolation - Conclusions From 25 Years Of Research, in: *Interfaces*, Vol. 14, S. 52-61.
- Armstrong, J. S. (1985) *Long-range forecasting: from crystal ball to computer*, 2, Wiley: New York.
- Armstrong, J. S. (2001a) Evaluating Forecasting Methods, in: *Principles of forecasting: a handbook for researchers and practitioners* (Hrsg. Armstrong, J. S.), Kluwer Academic: Boston; London, S. 443-472.
- Armstrong, J. S. (2001b) Extrapolation for Time-Series and Cross-Sectional Data, in: *Principles of forecasting: a handbook for researchers and practitioners* (Hrsg. Armstrong, J. S.), Kluwer Academic: Boston; London, S. 217-243.
- Armstrong, J. S. (2001c) Introduction to Forecasting Principles, in: *Principles of forecasting: a handbook for researchers and practitioners* (Hrsg. Armstrong, J. S.), Kluwer Academic: Boston; London, S. 1-13.
- Armstrong, J. S. (Hrsg.) (2001d) *Principles of forecasting: a handbook for researchers and practitioners*, Kluwer Academic: Boston; London.
- Armstrong, J. S. (2001e) Standards and Practices for Forecasting, in: *Principles of forecasting: a handbook for researchers and practitioners* (Hrsg. Armstrong, J. S.), Kluwer Academic: Boston; London, S. 679-732.

- Armstrong, J. S. (2005) File-drawer problem of Neural Networks, A comment on Chatfield [2005], (persönliche Kommunikation mit Crone, S. F.), ISF'05, San Antonio, Texas, 15.06.2005
- Armstrong, J. S. und Collopy, F. (1992) Error measures for generalizing about forecasting methods: Empirical comparisons, in: *International Journal of Forecasting*, Vol. 8, S. 69-80.
- Armstrong, J. S. und Fildes, R. (1995) On The Selection Of Error Measures For Comparisons Among Forecasting Methods, in: *Journal Of Forecasting*, Vol. 14, S. 67-71.
- Armstrong, J. S. und Grohman, M. C. (1972) A Comparative Study Of Methods For Long-Range Market Forecasting, in: *Management Science*, Vol. 19, S. 211-221.
- Armstrong, J. S., Koehler, A. B., Fildes, R., Hibon, M., Makridakis, S. und Meade, N. (1998) Commentaries on "Generalizing about univariate forecasting methods: further empirical evidence", in: *International Journal Of Forecasting*, Vol. 14, S. 359-366.
- Arnold, D., Isermann, H., Kuhn, A. und Tempelmeier, H. (Hrsg.) (2004) *Handbuch Logistik*, Springer: Berlin [u.a.].
- Arnolds, H., Heege, F. und Tussing, W. (2001) *Materialwirtschaft und Einkauf: praxisorientiertes Lehrbuch*, 10., durchges. Aufl., Nachdr, Gabler: Wiesbaden.
- Arrow, K. J., Harris, T. und Marschak, J. (1951) Optimal Inventory Policy, in: *Econometrica*, Vol. 19, S. 250-272.
- Arrow, K. J., Karlin, S. und Scarf, H. (1958) *Studies in the mathematical theory of inventory and production*, Stanford University Press: Stanford, CA.
- Artto, K. A. und Pylkkanen, E. (1999) An effective procedure for the distribution of magazines, in: *International Transactions in Operational Research*, Vol. 6, S. 289.
- Arul, M. (1996) Neural net classifiers for communication systems: A review, in: *Iete Technical Review*, Vol. 13, S. 3-10.
- Ashenayi, K. (1991) Application of Neural Networks In Optimization Problems - A Review, in: *Applications of Optical Engineering: Proceedings of OE/Midwest 90*, Vol. 1396 *Spie - Int Soc Optical Engineering*: Bellingham, S. 285-296.
- Askin, R. G. (1981) A Procedure for Production Lot Sizing with Probabilistic Dynamic Demand, in: *AIIE Transactions*, Vol. 13, S. 132-137.
- Assfalg, H. (1976) *Lagerhaltungsmodelle für mehrere Produkte*, Hain: Meisenheim.
- Atiya, A. F. (2001) Bankruptcy prediction for credit risk using neural networks: A survey and new results, in: *Ieee Transactions On Neural Networks*, Vol. 12, S. 929-935.
- Atzberger, M. (Hrsg.) (1998a) *Automatische Disposition - Bestandsaufnahme und Perspektiven*, DHI - Deutsches Handelsinstitut: Köln.
- Atzberger, M. (1998b) *Automatische Disposition - Von den Lehrbüchern in die Filialen*, in: *Automatische Disposition - Bestandsaufnahme und Perspektiven* (Hrsg. Atzberger, M.), DHI - Deutsches. Handelsinstitut: Köln, S. 6-12.
- Ausschuss für Begriffsdefinitionen aus der Handels- und Absatzwirtschaft (Hrsg.) (1995) *Katalog E. Begriffsdefinitionen aus der Handels- und Absatzwirtschaft*: Köln.
- Axsäter, S. (2000) Exact analysis of continuous review (R, Q) policies in two-echelon inventory systems with compound Poisson demand, in: *Operations Research*, Vol. 48, S. 686-696.
- Axsäter, S. (1993) Continuous Review Policies for Multi-Level Inventory Systems with Stochastic Demand, in: *Logistics of production and inventory* (Hrsg. Graves, S. C., Rinnooy Kan, A. H. G. und Zipkin, P. H.), North-Holland: Amsterdam [u.a.], S. 175-197.
- Baba, N. und Suto, H. (2000) Utilization of artificial neural networks & GAs for constructing an intelligent sales prediction system, in: *Ijcn 2000: Proceedings Of The Ieee-Inns*

- Enns International Joint Conference On Neural Networks, Vol Vi, Ieee Computer Soc: Los Alamitos, S. 565-570.
- Babel, W. (1997) Einsatzmöglichkeiten neuronaler Netze in der Industrie: Mustererkennung anhand überwachter Lernverfahren - mit Beispielen aus der Verkehrs- und Medizintechnik, Expert: Renningen.
- Babri, H. A., Chen, Y. Q. und Ahsan, K. (2003) Back propagation with randomized cost function for training neural networks, in: Rough Sets, Fuzzy Sets, Data Mining, And Granular Computing, Vol. 2639, Springer-Verlag Berlin: Berlin, S. 471-475.
- Backhaus, K., Erichson, B., Plinke, W. und Weiber, R. (1994) Multivariate Analysemethoden: eine anwendungsorientierte Einführung, 7., vollst. überarb. und erw. Aufl, Springer: Berlin [u.a.].
- Backhaus, K., Erichson, B., Plinke, W. und Weiber, R. (2003) Multivariate Analysemethoden, 10. Auflage, Springer: Berlin.
- Bakircioglu, H. und Kocak, T. (2000) Survey of random neural network applications, in: European Journal Of Operational Research, Vol. 126, S. 319-330.
- Balakrishnan, K. und Honavar, V. (1992) Improving Convergence of Back-Propagation by Handling Flat-Spots in the Output Layer, in: Proceedings of the ICANN-92 - International Conference on Artificial Neural Networks, Brighton, UK, Vol. IEEE.
- Balakrishnan, P. V. (1992a) A Cluster analysis of brand choice data using neural networks, College of Business The Ohio State University: Columbus.
- Balakrishnan, P. V. (1992b) A Study of the classification capabilities of neural networks for market segmentation: a comparison with k-means clustering, College of Business The Ohio State University: Columbus.
- Balkin, S. D. und Ord, J. K. (2000) Automatic neural network modeling for univariate time series, in: International Journal of Forecasting, Vol. 16, S. 509-515.
- Bansal, K., Vadhavkar, S. und Gupta, A. (1998) Brief Application Description. Neural Networks Based Forecasting Techniques for Inventory Control Applications, in: Data Mining and Knowledge Discovery, Vol. V2, S. 97.
- Bartezzaghi, E., Verganti, R. und Zotteri, G. (1999) Measuring the impact of asymmetric demand distributions on inventories, in: International Journal of Production Economics, Vol. 60-1, S. 395-404.
- Bartolomei, S. M. und Sweet, A. L. (1989) A note on a comparison of exponential smoothing methods for forecasting seasonal series, in: International Journal of Forecasting, Vol. 5, S. 111-116.
- Bass, F. M. (1969) New Product Growth For Model Consumer Durables, in: Management Science Series A-Theory, Vol. 15, S. 215-227.
- BDTA (2005) Die Branche, Vol. 2006 (Hrsg. BDTA) Bundesverband der Tabakwaren-Großhändler und Automatenaufsteller: Köln, <http://www.bdta.de>.
- Becker, J. und Schütte, R. (1996) Handelsinformationssysteme, MI Verlag: Landsberg/Lech.
- Berger, R. (2000) Analyse zum Kaufverhalten an Zigarettensautomaten: München.
- Bichler, K. und Lorsch, W. (1985) Optimale Bestandsplanung im Handel, Kohlhammer: Stuttgart [u.a.].
- Bierhals, E. (1993) Organisation der Materialwirtschaft, 1. Aufl, Gabler: Wiesbaden.
- Binner, J. M., Elger, T., Nilsson, B. und Tepper, J. A. (2004) Tools for non-linear time series forecasting in economics - An empirical comparison of regime switching vector autoregressive models and recurrent neural networks, in: Applications Of Artificial Intelligence In Finance And Economics, Vol. 19, Jai-Elsevier Sci Bv: Amsterdam, S. 71-91.

- Bishop, C. M. (1995) Neural networks for pattern recognition, Oxford University Press: Oxford.
- Bode, J. (1998) Decision support with neural networks in the management of research and development: Concepts and application to cost estimation, in: Information & Management, Vol. 34, S. 33-40.
- Bohr, H., Bohr, J., Brunak, S., Cotterill, R. M. J., Fredholm, H., Lautrup, B. und Petersen, S. B. (1990) A Novel-Approach To Prediction Of The 3-Dimensional Structures Of Protein Backbones By Neural Networks, in: Febs Letters, Vol. 261, S. 43-46.
- Bougaardt, G. (2002) An investigation into the application of artificial neural networks and cluster analysis in long-term load forecasting
- Bourbia, W. A. (1989) Application of artificial neural networks to business
- Box, G. E. P. und Jenkins, G. M. (1970) Time series analysis: forecasting and control, Holden-Day: San Francisco.
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C. und Box Jenkins, R. (1994) Time series analysis: forecasting and control, 3. ed, Prentice Hall: Englewood Cliffs, NJ [u.a.].
- Box, G. E. P. und Tiao, G. C. (1975) Intervention Analysis With Applications To Economic And Environmental Problems, in: Journal Of The American Statistical Association, Vol. 70, S. 70-79.
- Brink, A. (1988) Operative Lager- und Bestellmengenplanung unter besonderer Berücksichtigung von Lagerkapazitätsrestriktionen, Eul: Bergisch Gladbach.
- Brockhoff, K. (1977) Prognoseverfahren für die Unternehmensplanung, Gabler: Wiesbaden.
- Brockhoff, K. (1979) Delphi-Prognosen im Computer-Dialog: experimentelle Erprobung und Auswertung kurzfristiger Prognosen, Mohr: Tübingen.
- Brown, C. E., Coakley, J. und Phillips, M. E. (1995) Neural networks enter the world of management accounting, in: Management Accounting, Vol. 76, S. 51.
- Brown, R. G. (1956) Exponential Smoothing for predicting demand, in: Tenth national meeting of the Operations Research Society of America, San Francisco.
- Brown, R. G. (1959) Statistical forecasting for inventory control, McGraw-Hill: New York.
- Brown, R. G. (1963) Smoothing, forecasting and prediction of discrete time series, Prentice-Hall: Englewood Cliffs, NJ.
- Buckler, F. (2001) NEUSREL - neuer Kausalanalyseansatz auf Basis neuronaler Netze als Instrument der Marketingforschung, Cuvillier: Göttingen
- Burge, P. und Shawe-Taylor, J. (2001) An unsupervised neural network approach to profiling the behavior of mobile phone users for use in fraud detection, in: Journal Of Parallel And Distributed Computing, Vol. 61, S. 915-925.
- Burgin, T. A. (1972a) Inventory Control with Normal Demand and Gamma Lead Times, in: Operational Research Quarterly, Vol. 23, S. 73-&.
- Burgin, T. A. (1972b) Inventory Control with Normal Demand and Gamma Lead Times, in: Operational Research Quarterly, Vol. 23, S. 73-87.
- Burgin, T. A. (1975) Gamma Distribution and Inventory Control, in: Operational Research Quarterly, Vol. 26, S. 507-525.
- Burgin, T. A. und Norman, J. M. (1976) Table for Determining Probability of a Stock out and Potential Lost Sales for a Gamma-Distributed Demand, in: Operational Research Quarterly, Vol. 27, S. 621-631.
- Burrascano, P. (1991) A norm selection criterion for the generalized delta rule, in: IEEE Transactions on Neural Networks, Vol. 2, S. 125.

- Burrascano, P., Fiori, S. und Mongiardo, M. (1999) A review of artificial neural networks applications in microwave computer-aided design (invited article), in: *International Journal of RF and Microwave Computer-Aided Engineering*, Vol. 9, S. 158-174.
- Cain, M. (1989) Median Predictive Cost Of Error With An Asymmetric Cost Function, in: *Journal of the Operational Research Society*, Vol. 40, S. 735-740.
- Cao, Q., Leggio, K. B. und Schniederjans, M. J. (2005a) A comparison between Fama and French's model and artificial neural networks in predicting the Chinese stock market, in: *Computers & Operations Research*, Vol. 32, S. 2499-2512.
- Cao, Y., Li, Y. und Liao, X. (2005b) Applying Modified Fuzzy Neural Network to Customer Classification of E-Business.
- Capistrán-Carmona, C. (2002) A Review of Forecast Theory using Generalised Cost Functions, Technical Report, Nr. X1, UCSD: San Diego.
- Capistrán-Carmona, C. (2005) Essays on forecast evaluation under general loss functions, unpublished PhD thesis, University of California San Diego (UCSD): San Diego
- Capobianco, E. (2000) Neural networks and statistical inference: seeking robust and efficient learning, in: *Computational Statistics & Data Analysis*, Vol. 32, S. 443.
- Caudil, M. (1992) A view from now, in: *AI Expert*, Vol. 4, S. 24-31.
- Cerny, V. (1985) Thermodynamical Approach to the Traveling Salesman Problem - an Efficient Simulation Algorithm, in: *Journal of Optimization Theory and Applications*, Vol. 45, S. 41-51.
- Chakrabarty, T., Giri, B. C. und Chaudhuri, K. S. (1998) An EOQ model for items with Weibull distribution deterioration, shortages and trended demand: An extension of Philip's model, in: *Computers & Operations Research*, Vol. 25, S. 649-657.
- Chambers, J. C., Mullick, S. K. und Smith, D. D. (1971) How to choose the right forecasting technique, in: *Harvard Business Review*, Vol. 29 Wiley: New York, S. 45-50.
- Chan, C. K., Kingsman, B. G. und Wong, H. (2004) Determining when to update the weights in combined forecasts for product demand - an application of the CUSUM technique, in: *European Journal of Operational Research*, Vol. 153, S. 757-768.
- Chang, F. J., Chen, Y. C. und Liang, J. M. (2002) Fuzzy clustering neural network as flood forecasting model, in: *Nordic Hydrology*, Vol. 33, S. 275-290.
- Chang, P. C., Wang, Y. W. und Tsai, C. Y. (2005) Evolving neural network for printed circuit board sales forecasting, in: *Expert Systems With Applications*, Vol. 29, S. 83-92.
- Charalambous, C., Charitou, A. und Kaourou, F. (2000) Comparative analysis of artificial neural network models: Application in bankruptcy prediction, in: *Annals of Operations Research*, Vol. 99, S. 403-425.
- Chatfield, C. (1988) Apples, oranges and the mean squared error, in: *Journal of Forecasting*, S. 515-518.
- Chatfield, C. (1993a) Calculating interval forecasts, in: *Journal of business and economic statistics*, S. 121-144.
- Chatfield, C. (1993b) Neural Networks - Forecasting Breakthrough or Passing Fad, in: *International Journal of Forecasting*, Vol. 9, S. 1-3.
- Chatfield, C. (1998) Time series forecasting with neural networks, in: *Neural Networks For Signal Processing Viii*, IEEE: New York, S. 419-427.
- Chatfield, C. (2005) File-drawer problem of Neural Networks, (persönliche Kommunikation mit Crone, S. F.), Santander, Spain, 14.06.2006
- Chatfield, C. und Yar, M. (1991) Prediction intervals for multiplicative Holt-Winters, in: *International Journal of Forecasting*, Vol. 7, S. 31.

- Chattopadhyay, S., Chaturvedi, A. und Sengupta, R. N. (2000) Sequential estimation of a linear function of normal means under asymmetric loss function, in: *Metrika*, Vol. 52, S. 225-235.
- Chen, L. und Chen, G. R. (2000) Fuzzy modeling, prediction, and control of uncertain chaotic systems based on time series, in: *IEEE Transactions On Circuits And Systems I-Fundamental Theory And Applications*, Vol. 47, S. 1527-1531.
- Chen, M. S. und Chuang, C. C. (2000) An extended newsboy problem with shortage-level constraints, in: *International Journal of Production Economics*, Vol. 67, S. 269-277.
- Chen, S. T., Yu, D. C., Moghaddamjo, A. R., Lu, C. N. und Vemuri, S. (1992) Weather Sensitive Short-Term Load Forecasting Using Nonfully Connected Artificial Neural Network, in: *Ieee Transactions On Power Systems*, Vol. 7, S. 1098-1105.
- Chen, Z. (2000) Trimmed and Winsorized M- and Z-estimators, with applications to robust estimation in neural network models, DAI
- Cheng, B. und Titterington, D. M. (1994) Neural Networks - A Review From A Statistical Perspective, in: *Statistical Science*, Vol. 9, S. 2-30.
- Chiu, C.-C. (1994) On adaptive radial basis function neural networks for the forecasting of time series processes, unpublished PhD thesis, Texas A&M University
- Chow, G. C. (1960) Tests Of Equality Between Sets Of Coefficients In 2 Linear Regressions, in: *Econometrica*, Vol. 28, S. 591-605.
- Chow, W. M. (1965) Adaptive Control of the Exponential Smoothing Constant, in: *Journal of Industrial Engineering*, S. 314-332.
- Chowdhury, F. N., Wahi, P., Raina, R. und Kaminedi, S. (2001) A survey of neural networks applications in automatic control, in: *Proceedings Of The 33rd Southeastern Symposium On System Theory*, IEEE: New York, S. 349-353.
- Christoffersen, P. F. und Diebold, F. X. (1994) Optimal prediction under asymmetric loss, Technical Report, Nr. 167, (Hrsg. National Bureau of Economic Reserach), NBER Technical Working Paper, NBER: Cambridge, Massachusetts, 38 Seiten.
- Christoffersen, P. F. und Diebold, F. X. (1996) Further results on forecasting and model selection under asymmetric loss, in: *Journal of Applied Econometrics*, Vol. 11, S. 561-571.
- Christoffersen, P. F. und Diebold, F. X. (1997a) Optimal prediction under asymmetric loss, Technical Report, Nr. 97-11, International Monetary Fund Working Papers, IMF: Philadelphia, 808-817 Seiten.
- Christoffersen, P. F. und Diebold, F. X. (1997b) Optimal prediction under asymmetric loss, in: *Econometric Theory*, Vol. 13, S. 808-817.
- Chu, C.-W. und Zhang, G. P. (2003) A comparative study of linear and nonlinear models for aggregate retail sales forecasting, in: *International Journal of Production Economics*, Vol. In Press, Corrected Proof.
- Churchland, P. M. (1989) *A neurocomputational perspective: the nature of mind and the structure of science*, MIT Press: Cambridge, MA.
- Churchland, P. S. (1986) *Neurophilosophy: toward a unified science of the mind-brain*, MIT Press: Cambridge, Mass.
- Churchland, P. S., Koch, C. und Sejnowski, T. (1990) *What is Computational Neuroscience?* in: *Computational Neuroscience* (Hrsg. Schwartz, E. L.), MIT Press: Cambridge, Mass.
- Clark, A. J. und Scarf, H. (1960) Optimal policies for a multiechelon inventory problem, in: *Management science*, Vol. 6, S. 475-490.
- Coakley, J. R. und Brown, C. E. (2000) Artificial neural networks in accounting and finance: modeling issues, in: *International Journal of Intelligent Systems in Accounting, Finance and Management*, Vol. 9, S. 119.

- Connor, J. T. (1996) A robust neural network filter for electricity demand prediction, in: *Journal of Forecasting*, Vol. 15, S. 437-458.
- Connor, J. T., Martin, R. D. und Atlas, L. E. (1994) Recurrent Neural Networks and Robust Time-Series Prediction, in: *Ieee Transactions on Neural Networks*, Vol. 5, S. 240-254.
- Conrad, O. (1996) Die Einsatzmöglichkeiten von neuronalen Netzen im Rahmen der Absatzprognose und Bezugsregulierung im Vertrieb von Publikumszeitschriften, P. Lang: Frankfurt am Main [u.a.].
- Corradi, V. und White, H. (1995) Regularized Neural Networks - Some Convergence Rate Results, in: *Neural Computation*, Vol. 7, S. 1225-1244.
- Cramer, C. (1998) Neural networks for image and video compression: A review, in: *European Journal Of Operational Research*, Vol. 108, S. 266-282.
- Crone, S. F. (2002) Training artificial neural networks for time series prediction using asymmetric cost functions, in: *Computational Intelligence For The E-Age, Proceedings of the ICONIP'02, The 9th International Conference On Neural Information Processing*, Nanyang Technological University: Singapore, S. 2374-2380.
- Crone, S. F. (2003a) Künstliche neuronale Netze zur betrieblichen Entscheidungsunterstützung, in: *WISU*, Vol. 32, S. 452-458.
- Crone, S. F. (2003b) Prediction of white noise time series using artificial neural networks and asymmetric cost functions, in: *Proceedings of the International Joint Conference on Neural Networks 2003*, Vol. 4, IEEE: New York, S. 2460-2465.
- Crone, S. F. (2004) A business forecasting competition approach to modeling artificial neural networks for time series prediction, in: *Proceedings of the ICAI'04 & MLMTA'04, Las Vegas*, Vol. 1 (Hrsg. Arabnia, H.) CSREA Press, S. 207-213.
- Crone, S. F. (2005) Stepwise Selection of Artificial Neural Network Models for Time Series Prediction, in: *Journal of Intelligent Systems*, Vol. 15.
- Crone, S. F. und Graffelle, P. C. (2004) An evaluation framework for publications on artificial neural networks in sales forecasting, in: *Proceedings of the International Conference on Artificial Intelligence, IC-AI'04, Las Vegas*, Vol. 1 (Hrsg. Arabnia, H.) CSREA Press, S. 221-227.
- Crone, S. F., Guajardo, J. und Weber, R. (2006a) The impact of Data Preprocessing on Support Vector Regression and Artificial Neural Networks in Time Series Forecasting, in: *Proceedings of the World Congress in Computational Intelligence, WCCI'06, Vancouver, Canada*, Vol. 1 (Hrsg. Yen, G. G.) IEEE: New York.
- Crone, S. F., Guajardo, J. und Weber, R. (2006b) A study on the ability of Support Vector Regression and Neural Networks to Forecast Basic Time Series Patterns, in: *Proceedings of the IFIP World Congress in Computation, WCC'06, Santiago, Chile*, Vol. 1 (Hrsg. IFIP) IFIP.
- Crone, S. F., Lessmann, S. und Pietsch, S. (2006c) An empirical Evaluation of Support Vector Regression versus Artificial Neural Networks to Forecast basic Time Series Patterns, in: *Proceedings of the World Congress in Computational Intelligence, WCCI'06, Vancouver, Canada*, Vol. 1 (Hrsg. Yen, G. G.) IEEE: New York.
- Crone, S. F., Lessmann, S. und Pietsch, S. (2006d) Parameter Sensitivity of Support Vector Regression and Neural Networks for Forecasting, in: *Proceedings of the International Conference on Data Mining, DMIN'06, Las Vegas, USA*, Vol. 1 (Hrsg. Crone, S. F., Lessmann, S. und Stahlbock, R.) CSREA Press.
- Crone, S. F., Lessmann, S. und Stahlbock, R. (2004) Empirical comparison and evaluation of classifier performance for data mining in customer relationship management, in: *2004 Ieee International Joint Conference On Neural Networks*, Vols 1-4, Proceedings, Ieee: New York, S. 443-448.
- Crone, S. F., Lessmann, S. und Stahlbock, R. (2005) Utility based data mining for time series analysis: cost-sensitive learning for neural network predictors, in: *Proceedings of the 1st*

- International Workshop on Utility-based Data Mining, 2005 Conference on Knowledge Discovery in Databases (KDD), Chicago, Illinois, Vol. ACM Press, S. 59-68.
- Crone, S. F., Lessmann, S. und Stahlbock, R. (2006e) The impact of preprocessing on data mining: An evaluation of classifier sensitivity in direct marketing, in: *European Journal Of Operational Research*, Vol. 173, S. 781-800.
- Crone, S. F. und Preßmar, D. B. (2006) An extended evaluation framework for publications on artificial neural networks in sales forecasting, in: *Proceedings of the AIA'06, Innsbruck*, Vol. 1 (Hrsg. IASTED) IASTED Press.
- Crone, S. F. und Soopramanien, D. (2005) Predicting customer online shopping adoption an evaluation of data mining and market modelling approaches, in: *Dmin '05: Proceedings Of The 2005 International Conference On Data Mining*, C S R E A Press: Athens, S. 215-221.
- Croston, J. D. (1972) Forecasting and Stock Control for Intermittent Demands, in: *Operational Research Quarterly*, Vol. 23, S. 289-&.
- Cybenko, G. (1989) Approximation by superposition of a sigmoid function, in: *Mathematics of Control Signal and Systems*, S. 303-314.
- Dagli, C. H. (1994) *Artificial neural networks for intelligent manufacturing*, 1st, Chapman & Hall: London; New York.
- Dagli, C. H. (2002) Smart engineering system design: neural networks, fuzzy logic, evolutionary programming, complex systems and artificial life: proceedings of the Artificial Neural Networks in Engineering Conference (ANNIE 2002), held November 10-13, 2002, in St. Louis, Missouri, U.S.A, ASME Press: New York, New York.
- Dagum, E. B. (1988) X-11-ARIMA/88 seasonal adjustment method: foundation and user manuals, Statistics Canada.
- Dalrymple, D. J. (1987) Sales forecasting practices: Results from a United States survey, in: *International Journal of Forecasting*, Vol. 3, S. 379-391.
- DARPA (1989) DARPA neural network study final report, The Laboratory: Lexington, Mass.
- Dassen, W. R. M., Egmont-Petersen, M. und Mulleneers, R. G. A. (1998) Artificial neural networks in cardiology; A review, in: *Cardiac Arrhythmias, Pacing & Electrophysiology - The Expert View*, Vol. 201, Kluwer Academic: Dordrecht, S. 205-211.
- De Gooijer, J. G. und Kumar, K. (1992) Some recent developments in non-linear time series modelling, testing, and forecasting, in: *International Journal of Forecasting*, Vol. 8, S. 135-156.
- de Menezes, L. M., W. Bunn, D. und Taylor, J. W. (2000) Review of guidelines for the use of combined forecasts, in: *European Journal of Operational Research*, Vol. 120, S. 190-204.
- Dekok, A. G., Tijms, H. C. und Schouten, F. A. V. (1984) Approximations for the Single-Product Production-Inventory Problem with Compound Poisson Demand and Service-Level Constraints, in: *Advances in Applied Probability*, Vol. 16, S. 378-401.
- del Aguila, M. D. R., Requena, I., Bernier, J. L., Ros, E. und Mota, S. (2004) Neural networks and statistics: A review of the literature, in: *Soft Methodology And Random Information Systems*, Springer-Verlag Berlin: Berlin, S. 597-604.
- Desai, V. S. und Bharati, R. (1998) The efficacy of neural networks in predicting returns on stock and bond indices, in: *Decision Sciences*, Vol. 29, S. 405-425.
- Desai, V. S., Crook, J. N. und Overstreet, G. A. (1996) A comparison of neural networks and linear scoring models in the credit union environment, in: *European Journal Of Operational Research*, Vol. 95, S. 24-37.
- Dhawan, R. und O'Connor, M. (2005) Time series forecasting: An approach to modeling seasonality using neural networks, in: *Proceedings of the ICAI'05: The 2005 Interna-*

- tional Conference On Artificial Intelligence, Athens, Vol. 1 (Hrsg. Arabnia, H.) CSREA Press, S. 41-45.
- Diamantaras, K. I. und Kung, S. Y. (1996) *Principal component neural networks: theory and applications*, Wiley: New York.
- Dickey, D. A. und Fuller, W. A. (1979) *Distribution Of The Estimators For Autoregressive Time-Series With A Unit Root*, in: *Journal Of The American Statistical Association*, Vol. 74, S. 427-431.
- Diebold, F. X. (2001) *Elements of forecasting*, 2nd, South-Western: Cincinnati, Ohio.
- Diebold, F. X. und Mariano, R. S. (1995) *Comparing Predictive Accuracy*, in: *Journal Of Business & Economic Statistics*, Vol. 13, S. 253-263.
- Diebold, F. X. und Mariano, R. S. (2002) *Comparing predictive accuracy (Reprinted)*, in: *Journal Of Business & Economic Statistics*, Vol. 20, S. 134-144.
- Dittrich, M. (2000) *Ein Beitrag zur Grobplanung der Lagerlogistik im Handel*, unpublished PhD thesis, Universität von St.Gallen: St.Gallen
- Domschke, W., Scholl, A. und Voß, S. (1997) *Produktionsplanung - Ablauforganisatorische Aspekte*, 2., überarb. und erw. Aufl, Springer: Berlin [u.a.].
- Donaldson, W. A. (1977) *Inventory Replenishment Policy for a Linear Trend in Demand – An Analytical Solution*, in: *Operational Research Quarterly*, Vol. 28, S. 663-670.
- Donaldson, W. A. (1984) *An Equation for the Optimal Value of T, the Inventory Replenishment Review Period When Demand Is Normal*, in: *Journal of the Operational Research Society*, Vol. 35, S. 137-139.
- Dong, X. N. und Wen, G. R. (2006) *An improved Neural Networks Prediction Model and its Application in Supply Chain*, in: *Nature and Science*, Vol. 3 Tongji Univ Press: Lansing, S. 24-27.
- Dong, X. N., Zhang, T. und Wen, G. R. (2005) *Application of recurrent neural networks prediction in supply chain*, in: *Proceedings Of The 2005 Conference Of System Dynamics And Management Science*, Vol 2 - Sustainable Development Of Asia Pacific, Tongji Univ Press: Shanghai, S. 718-722.
- Dorffner, G. (1996) *Neural Networks for Time Series Processing*, in: *Neural Network World*, Vol. 6, S. 447-468.
- Dorling, S. R., Foxall, R. J., Mandic, D. P. und Cawley, G. C. (2003) *Maximum likelihood cost functions for neural network models of air quality data*, in: *Atmospheric Environment*, Vol. 37, S. 3435-3443.
- Dougherty, M. (1995) *A Review Of Neural Networks Applied To Transport*, in: *Transportation Research Part C-Emerging Technologies*, Vol. 3, S. 247-260.
- Drago, G. P., Morando, M. und Ridella, S. (1995) *An Adaptive Momentum Back-Propagation (Ambp)*, in: *Neural Computing & Applications*, Vol. 3, S. 213-221.
- Dueck, G. und Scheuer, T. (1990) *Threshold Accepting - a General-Purpose Optimization Algorithm Appearing Superior to Simulated Annealing*, in: *Journal of Computational Physics*, Vol. 90, S. 161-175.
- Duncan, J. C. (1996) *A comparison of radial basis function and multilayer perceptron neural networks with linear multiple regression in cohort-survival based enrollment projection*, unpublished PhD thesis, Kent State University
- Durbin, J. und Watson, G. S. (1950) *Testing For Serial Correlation In Least Squares Regression.1*, in: *Biometrika*, Vol. 37, S. 409-428.
- Durbin, J. und Watson, G. S. (1951) *Testing For Serial Correlation In Least Squares Regression.2*, in: *Biometrika*, Vol. 38, S. 159-178.
- Durbin, J. und Watson, G. S. (1971) *Testing For Serial Correlation In Least Squares Regression.3*, in: *Biometrika*, Vol. 58, S. 1-&.

- ECR Europe und Roland Berger (2003) ECR - Optimal Shelf Availability, Technical Report, (Hrsg. ECR Europe), 65 Seiten.
- Eglese, R. W. (1990) Simulated Annealing - a Tool for Operational-Research, in: European Journal of Operational Research, Vol. 46, S. 271-281.
- Egmont-Petersen, M., de Ridder, D. und Handels, H. (2002) Image processing with neural networks - a review, in: Pattern Recognition, Vol. 35, S. 2279-2301.
- EHI (2006) EHI-Logistik-Studie 2006, Technical Report, (Hrsg. EHI Retail Institute), EHI Retail Institute, 37 Seiten.
- Ehrhardt, R. (1979) The Power Approximation for Computing (s,S) Inventory Policies, in: Management Science, Vol. 25, S. 777-786.
- El-Gindy, M. und Palkovics, L. (1993) Possible application of artificial neural networks to vehicle dynamics and control: a literature review, Technical Report, Nr. 208, Centre for Surface Transportation Technology (Canada), p. 592-614 Seiten.
- Elgindy, M. und Palkovics, L. (1993) Possible Application Of Artificial Neural Networks To Vehicle Dynamics And Control - A Literature-Review, in: International Journal Of Vehicle Design, Vol. 14, S. 592-614.
- Elrod, D. W., Maggiora, G. M. und Trenary, R. (1989) Application Of Neural Networks In Chemical-Reaction Prediction, in: Abstracts Of Papers Of The American Chemical Society, Vol. 197.
- Elrod, D. W., Maggiora, G. M. und Trenary, R. G. (1990) Applications Of Neural Networks In Chemistry.1. Prediction Of Electrophilic Aromatic-Substitution Reactions, in: Journal Of Chemical Information And Computer Sciences, Vol. 30, S. 477-484.
- Elsevier (2003) Top requested papers of the International Journal of Forecasting, Vol. 2003 Elsevier: online im Internet, <http://www.elsevier.com/homepage/sae/econworld/econbase/intfor/frame.htm>.
- Emde (1989) Prognosetechniken und -systeme, in: Handwörterbuch der Planung - Enzyklopädie der Betriebswirtschaftslehre, Vol. 9 (Hrsg. Szyperski, N.), Schäffer-Poeschel: Stuttgart.
- Encyclopaedia Britannica inc. (2005) Encyclopaedia Britannica 2005, Encyclopaedia Britannica, Inc.: Chicago, IL, S. 1 DVD-ROM.
- Engle, R. (2002) New frontiers for arch models, in: Journal Of Applied Econometrics, Vol. 17, S. 425-446.
- Engle, R. F. (1982) Autoregressive Conditional Heteroscedasticity With Estimates Of The Variance Of United-Kingdom Inflation, in: Econometrica, Vol. 50, S. 987-1007.
- Englich, H. und Hiemstra, Y. (1994) The Correlation As Cost Function In Neural Networks, in: 1994 Ieee International Conference On Neural Networks, Vol 1-7, I E E E: New York, S. 3170-3172.
- Eppen, G. D. und Martin, R. K. (1988) Determining safety stock in the presence of stochastic lead time and demand, in: Management science, Vol. 34, S. 1380-1390.
- Eurohandelsinstitut (1999) Der Tabakwarenfachgroßhandel, Ehi: Köln.
- European Vending Association (2004) The European Representation of the Vending Industry, Technical Report, EVA: Brussels, 29 Seiten.
- EVA (2004) The European Representation of the Vending Industry, Technical Report, (Hrsg. EVA), European Vending Association: Brussels, 29 Seiten.
- Fahlman, S. E. (1988) An empirical study of learning speed in back-propagation networks, in: Proceedings of the 1988 Connectionist Models Summer School, Carnegie Mellon University, Vol. (Hrsg. Touretzky, D. S., Hinton, G. und Sejnowski, T. J.) Morgan Kaufmann.

- Fang, Y. (2003) Forecasting combination and encompassing tests, in: *International Journal of Forecasting*, Vol. 19, S. 87-94.
- Faraway, J. und Chatfield, C. (1995) Time series forecasting with neural networks: A case study, Technical Report, Research report 95-06 of the statistics group: University of Bath, 1-21 Seiten.
- Faraway, J. und Chatfield, C. (1998) Time series forecasting with neural networks: A comparative study using the airline data, in: *Applied statistics*, Vol. 47, S. 20.
- FDI, FMI und GMA (2002) 2002 Unsaleables Benchmark Report, Technical Report, (Hrsg. Lightburn, A.), FDI, Food Distributors International; FMI, Food Marketing Institute; GMA, Grocery Manufacturers of America: Washington DC.
- Federgruen, A. (1993) Centralized Planning Models for Multi-Echelon Inventory Systems under Uncertainty, in: *Logistics of production and inventory* (Hrsg. Graves, S. C., Rinnooy Kan, A. H. G. und Zipkin, P. H.), North-Holland: Amsterdam [u.a.], S. 133-173.
- Feng, J. (2004) *Computational neuroscience: comprehensive approach*, Chapman & Hall/CRC: Boca Raton.
- Fernandes, E., Silva, P., Melo, M. und d'Assuncao, A. G. (2005) Modular configuration RProp/SF-ANN for microwave waveguide on EBG structures modeling, in: *2005 Ieee Mtt-S International Microwave Symposium*, Vols 1-4, Ieee: New York, S. 1135-1138.
- Ferrano, F. J. und Wong, K. V. (1990) Prediction Of Thermal Storage Loads Using A Neural Network, in: *Ashrae Transactions*, Vol 96, Pt 2, 1990 - Technical And Symposium Papers Presented At The 1990 Annual Meeting, Vol. 96, Amer Soc Heating, Refrigerating And Air-Conditioning Eng: Atlanta, S. 723-726.
- Fildes, R. (1989) Evaluation Of Aggregate And Individual Forecast Method Selection-Rules, in: *Management Science*, Vol. 35, S. 1056-1065.
- Fildes, R. (1992a) The Evaluation Of Extrapolative Forecasting Methods, in: *International Journal Of Forecasting*, Vol. 8, S. 81-98.
- Fildes, R. (1992b) On Error Measures - A Response To The Commentators - The Best Error Measure, in: *International Journal Of Forecasting*, Vol. 8, S. 109-111.
- Fildes, R., Hibon, M., Makridakis, S. und Meade, N. (1998) Generalising about univariate forecasting methods: further empirical evidence, in: *International Journal Of Forecasting*, Vol. 14, S. 339-358.
- Fildes, R. und Makridakis, S. (1988) Forecasting and loss functions, in: *International Journal of Forecasting*, Vol. 4, S. 545-550.
- Fildes, R. und Makridakis, S. (1995) The Impact Of Empirical Accuracy Studies On Time-Series Analysis And Forecasting, in: *International Statistical Review*, Vol. 63, S. 289-308.
- Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C. und Chen, B.-C. (1998) New capabilities and methods of the X-12-ARIMA seasonal-adjustment program, in: *Journal of business & economic statistics*, Vol. 16, S. 127-152.
- Flood, I. (2001) Neural networks in civil engineering: A review, in: *Civil And Structural Engineering Computing: 2001*, Saxe-Coburg Publications: Edinburgh, S. 185-209.
- Fogel, D. B. und Fogel, L. J. (1994) Evolutionary Computation, in: *Ieee Transactions on Neural Networks*, Vol. 5, S. 1-1.
- Fogelman-Soulié, F. und Gallinari, P. (1998) *Industrial applications of neural networks*, World Scientific: Singapore; River Edge, NJ.
- Forsberg, R. (1997) Exact evaluation of (R,Q)-policies for two-level inventory systems with Poisson demand, in: *European Journal of Operational Research*, Vol. 96, S. 130-138.
- Forti, M., Manetti, S. und Marini, M. (1992) Neural Networks For Optimization Of Nonquadratic Cost-Functions With Application To Adaptive Signal-Processing, in: *1992 Ieee*

- International Symposium On Circuits And Systems, Vols 1-6, I E E E: New York, S. 2909-2912.
- Foster, W. R., Collopy, F. und Ungar, L. H. (1992) Neural Network Forecasting of Short, Noisy Time-Series, in: *Computers & Chemical Engineering*, Vol. 16, S. 293-297.
- Fratzl, H. (1992) Ein- und mehrstufige Lagerhaltung, *Physica*: Heidelberg.
- Fricke, F. und Rube, K.-H. (2001) *Betriebswirtschaftslehre*, 4., aktualisierte Aufl., Bayerischer Schulbuch-Verl.: München.
- Fricker, R. D. und Goodhart, C. A. (2000) Applying a bootstrap approach for setting reorder points in military supply systems, in: *Naval Research Logistics*, Vol. 47, S. 459-478.
- Fulda, E., Härter, M. und Lenk, H. (1989) Prognoseprobleme, in: *Handwörterbuch der Planung - Enzyklopädie der Betriebswirtschaftslehre*, Vol. 9 (Hrsg. Szyperski, N.), Schäffer-Poeschel: Stuttgart.
- Funahashi, K.-I. (1989) On the Approximate Realization of Continuous Mappings by Neural Networks, in: *Neural Networks*, Vol. 2, S. 183-192.
- Fyfe, C. und MacDonald, D. (2002) epsilon-insensitive Hebbian learning, in: *Neurocomputing*, Vol. 47, S. 35-57.
- Gardner, E. S. und Dannenbring, D. G. (1980) Forecasting with Exponential Smoothing: some guidelines for model selection, in: *Decision sciences*, Vol. 11, S. 370-383.
- Gardner, E. S. und MacKenzie, E. (1989) Seasonal exponential smoothing with damped trends, in: *Management Science*, Vol. 35, S. 372-376.
- Gemello, R., Albesano, D., Mana, F. und Moisa, L. (2000) Multi-Source Neural Networks for speech recognition: a review of recent results, in: *Ijcn 2000: Proceedings Of The Ieee-Inns-Enns International Joint Conference On Neural Networks*, Vol V, Ieee Computer Soc: Los Alamitos, S. 265-270.
- Gerfin, H. (1964) *Wirtschaftsprognose*: Basel.
- Geunes, J. und Zeng, A. Z. (2001) Impacts of inventory shortage policies on transportation requirements in two-stage distribution systems, in: *European Journal of Operational Research*, Vol. 129, S. 299-310.
- Geunes, J. P., Ramasesh, R. V. und Hayya, J. C. (2001) Adapting the newsvendor model for infinite-horizon inventory systems, in: *International Journal of Production Economics*, Vol. 72, S. 237-250.
- Ghamasae, R., Amoozegar, F., Pang, H. Y., Chin, Y. und Blackman, S. (1997) Survey of neural networks as applied to target tracking, in: *Automatic Target Recognition Vii*, Vol. 3069, *Spie - Int Soc Optical Engineering*: Bellingham, S. 412-423.
- Ghysels, E. und Osborn, D. R. (2001) *The econometric analysis of seasonal time series*, Cambridge University Press: New York.
- Gilchrist, W. (1976) *Statistical forecasting*, Wiley: London [u.a.].
- Girosi, F., Jones, M. und Poggio, T. (1995) Regularization Theory and Neural Networks Architectures, in: *Neural Computation*, Vol. 7, S. 219-269.
- Girosi, F. und Poggio, T. (1994) Networks and the Best Approximation Property, in: *Neuro-control systems: theory and applications* (Hrsg. Gupta, M. M. und Rao, D. H.), IEEE Press: New York, S. 257-264.
- Gisholt, O. (1976) *Marketing-Prognosen: unter besonderer Berücksichtigung der Delphi-Methode*, Haupt: Bern [u.a.].
- Glover, F. (1990) Tabu Search - a Tutorial, in: *Interfaces*, Vol. 20, S. 74-94.
- GMA (2002) Full Shelf Satisfaction - Reducing Out-of-Stocks in the Grocery Channel, Technical Report, (Hrsg. Grocery Manufacturers' Association), Grocery Manufacturers' Association.

- GMA, CIES und FMI (2002) Retail Out-of-Stocks - A Worldwide Examination of Extent, Causes and Consumer Responses, Technical Report, (Hrsg. GMA), GMA Grocery Manufacturers of America, CIES Food Business Forum, FMI Food Marketing Institute.
- GMA und IBM (2005) The GMA 2005 Logistics Survey, Technical Report, (Hrsg. Grocery Manufacturers' Association), Grocery Manufacturers of America.
- Goita, K., Gonzalezrubio, R., Benie, G. B., Royer, A. und Michaud, F. (1994) Literature-Review Of Artificial Neural Networks And Knowledge-Based Systems For Image-Analysis And Interpretation Of Data In Remote-Sensing, in: Canadian Journal Of Electrical And Computer Engineering-Revue Canadienne De Genie Electrique Et Informatique, Vol. 19, S. 53-61.
- Goodrich, R. L. (2000) The Forecast Pro methodology, in: International Journal of Forecasting, Vol. 16, S. 533-535.
- Goonatilake, S. und Treleaven, P. C. (1995) Intelligent systems for finance and business, Wiley: Chichester; New York.
- Gorzalczany, M. B. (1998) Business data modelling and forecasting with the use of fuzzy neural networks, in: Ieee International Symposium On Industrial Electronics (Isie 98) - Proceedings, Vols 1 And 2, IEEE: New York, S. 396-401.
- Götz, N. D. (1999) Automatische Dispositionssysteme für den Handel, Eul: Lohmar [u.a].
- Graepel, T. und Schraudolph, N. N. (2002) Stable adaptive momentum for rapid online learning in nonlinear systems, in: Artificial Neural Networks - Icann 2002, Vol. 2415, Springer-Verlag Berlin: Berlin, S. 450-455.
- Graf, H. (1968) Die kurzfristige Absatzprognose als Bestandteil der Planung in der Konsumgüterindustrie, Keller: Zürich.
- Grambsch, P. und Stahel, W. A. (1990) Forecasting demand for special telephone services*1: A case study, in: International Journal of Forecasting, Vol. 6, S. 53-64.
- Granger, C. W. J. (1969) Prediction With A Generalized Cost Of Error Function, in: Operational Research Quarterly, Vol. 20, S. 199-208.
- Granger, C. W. J. (1980) Forecasting in business and economics, Academic Press: New York.
- Granger, C. W. J. (1993) On the limitations of comparing mean square forecast errors: comment, in: Journal of Forecasting, Vol. 12, S. 651-652.
- Granger, C. W. J. (1999) Outline of Forecast Theory using Generalised Cost Functions, in: Spanish Economic Review, Vol. 1, S. 161-173.
- Granger, C. W. J. (2001) Evaluation of Forecasts, in: Understanding economic forecasts (Hrsg. Hendry, D. F. und Ericsson, N. R.), MIT Press: Cambridge, Mass.
- Granger, C. W. J. (2004) Personal conversation between C.W. Granger and S.F. Crone at ISF04, (persönliche Kommunikation mit Crone, S. F.), Sydney, 22.07.2004
- Granger, C. W. J. und Andersen, A. P. (1978) An introduction to bilinear time series models, Vandenhoeck und Ruprecht: Göttingen.
- Granger, C. W. J. und Newbold, P. (1986) Forecasting economic time series, 2. edition, Acad. Press: San Diego [u.a].
- Granger, C. W. J. und Pesaran, M. H. (2000a) A Decision Theoretic Approach to Forecast Evaluation, in: Statistics and Finance: An Interface (Hrsg. Chon, W. S., Li, W. K. und Tong, H.), Imperial College Press: London, S. 261-278.
- Granger, C. W. J. und Pesaran, M. H. (2000b) Economic and statistical measures of forecast accuracy, in: Journal Of Forecasting, Vol. 19, S. 537-560.
- Granger, C. W. J. und Teräsvirta, T. (1993) Modelling nonlinear economic relationships, Oxford University Press: Oxford [England]; New York.

- Graves, S. C., Rinnooy Kan, A. H. G. und Zipkin, P. H. (Hrsg.) (1993) Logistics of production and inventory, North-Holland: Amsterdam [u.a.].
- Green, B. P. und Choi, J. H. (1997) Assessing the risk of management fraud through neural network technology, in: Auditing-A Journal Of Practice & Theory, Vol. 16, S. 14-28.
- Greenwood, D. (1990) Review Of Applications Of Neural Networks To Robotics, in: Life-Sciences Simulation: Then, Now And When, Soc Computer Simulation Int: San Diego, S. 61-63.
- Grochla, E. (1992) Grundlagen der Materialwirtschaft: das materialwirtschaftliche Optimum im Betrieb, 3., gründl. durchges. Aufl., unveränd. Nachdr, Gabler: Wiesbaden.
- Grossberg, S. (1976a) Adaptive Pattern-Classification And Universal Recoding.1. Parallel Development And Coding Of Neural Feature Detectors, in: Biological Cybernetics, Vol. 23, S. 121-134.
- Grossberg, S. (1976b) Adaptive Pattern-Classification And Universal Recoding.2. Feedback, Expectation, Olfaction, Illusions, in: Biological Cybernetics, Vol. 23, S. 187-202.
- Grosser, H., Britos, P. und Garcia-Martinez, R. (2005) Detecting fraud in mobile telephony using neural networks, in: Innovations In Applied Artificial Intelligence, Vol. 3533, Springer-Verlag Berlin: Berlin, S. 613-615.
- Gullu, R., Onol, E. und Erkip, N. (1997) Analysis of a deterministic demand production/inventory system under nonstationary supply uncertainty, in: Iie Transactions, Vol. 29, S. 703-709.
- Günther, H.-J. und Busch, R. (1965) Das Bilden von Rationalisierungsschwerpunkten in Industriebetrieben mit Hilfe der ABC-Analyse, Gesellschaft f. Betriebsberatung u. Absatzförderung mbH: Ludwigshafen-Rheingönheim.
- Günther, H.-O. (1991) Bestellmengenplanung aus logistischer Sicht, in: Zeitschrift für Betriebswirtschaft, Vol. 61, S. 641-666.
- Gupta, A. und Lam, S. M. (1998) Weight decay backpropagation for noisy data, in: Neural Networks, Vol. 11, S. 1127-1137.
- Gupta, J. N. D. und Smith, K. A. (2000) Neural networks in business, Pergamon: New York.
- Gupta, M. M. und Rao, D. H. (Hrsg.) (1994a) Neuro-control systems: theory and applications, IEEE Press: New York.
- Gupta, M. M. und Rao, D. H. (1994b) Theory of Neuronal Approximations, in: Neuro-control systems: theory and applications (Hrsg. Gupta, M. M. und Rao, D. H.), IEEE Press: New York, S. 245-246.
- Gutenberg, E. (1990) Einführung in die Betriebswirtschaftslehre, 1, Nachdr., Gabler: Wiesbaden.
- Gutteridge, A., Bartlett, G. J. und Thornton, J. M. (2003) Using a neural network and spatial clustering to predict the location of active sites in enzymes, in: Journal Of Molecular Biology, Vol. 330, S. 719-734.
- Haberbosch, B. F. und Volkmann, J. (1994) Schöne alte Automaten: Waren-, Spiel- und Unterhaltungsautomaten, Battenberg: Augsburg.
- Hadeler, T. (2000) Gabler Wirtschafts-Lexikon: die ganze Welt der Wirtschaft; Betriebswirtschaft; Volkswirtschaft; Recht; Steuern, 15., vollst. überarb. und aktualisierte Aufl., Gabler: Wiesbaden.
- Hadley, G. und Whitin, T. M. (1963) Analysis of inventory systems, Prentice-Hall: Englewood Cliffs, NJ.
- Hagiwara, M. und Sato, A. (1995) Analysis Of Momentum Term In Backpropagation, in: Ieice Transactions On Information And Systems, Vol. E78D, S. 1080-1086.
- Hallier, B. (1995) Direkte Produkt-Rentabilität, in: Handwörterbuch des Marketing (Hrsg. Tietz, B.), Schäffer-Poeschel: Stuttgart, S. 492-498.

- Ham, F. M. und Kostanic, I. (2000) Principles of neurocomputing for science and engineering, McGraw Hill: New York, NY.
- Hampel, F. R. (1986) Robust statistics: the approach based on influence functions, Wiley: New York [u.a.].
- Hansen, J. V. und Nelson, R. D. (1997) Neural networks and traditional time series methods: A synergistic combination in state economic forecasts, in: Ieee Transactions on Neural Networks, Vol. 8, S. 863-873.
- Hansmann, K.-W. (1979) Heuristische Prognoseverfahren, in: WISU, Vol. 8, S. 229-233.
- Hansmann, K.-W. (1980) Die Anwendung der Spektral- und Regressionsanalyse bei der Absatzprognose von Zeitschriften, in: Angewandte Prognoseverfahren (Hrsg. Schwarze, J. und Backert, K.), Verl. Neue Wirtschafts-Briefe: Herne [u.a.], S. 231-244.
- Hansmann, K.-W. (1983) Kurzlehrbuch Prognoseverfahren: mit Aufgaben u. Lösungen, Gabler: Wiesbaden.
- Hansmann, K.-W. (2001) Industrielles Management, 7., völlig überarb. und erw. Aufl, Oldenbourg: München [u.a.].
- Hanson, S. J. und Burr, D. J. (1988) Minkowski-r Back-Propagation: Learning in Connectionist Models with Non-Euclidian Error Signals, in: Proceedings of the NIPS, Whistler, Vol. (Hrsg. NIPS) NIPS.
- Harris, F. (1915) Operations and Costs, AW Shaw: Chicago.
- Harrison, P. J. (1967) Exponential Smoothing and Short-Term Sales Forecasting, in: Management Science, Vol. 13, S. 821-842.
- Hartmann, H. (2002) Materialwirtschaft: Organisation, Planung, Durchführung, Kontrolle, 8., überarb. und erw. Aufl., Dt. Betriebswirte-Verl.: Gernsbach.
- Hartung, J., Elpelt, B. und Klösener, K.-H. (2005) Statistik: Lehr- und Handbuch der angewandten Statistik; mit zahlreichen, vollständig durchgerechneten Beispielen, 14., unwesentlich veränd. Aufl, Oldenbourg: München [u.a.].
- Harvey, A. C. (1993) Time series models, 2nd, MIT Press: Cambridge, Mass.
- Hax, A. C. und Candea, D. (1984) Production and inventory management, Prentice-Hall: Englewood Cliffs, N.J.
- Hayashi, Y. und Setiono, R. (2002) Combining neural network predictions for medical diagnosis, in: Computers In Biology And Medicine, Vol. 32, S. 237-246.
- Haykin, S. S. (1999) Neural networks: a comprehensive foundation, 2nd, Prentice Hall: Upper Saddle River, NJ.
- He, H. X., Wang, J. C., Graco, W. und Hawkins, S. (1997) Application of neural networks to detection of medical fraud, in: Expert Systems With Applications, Vol. 13, S. 329-336.
- Hebb, D. O. (1949) The Organisation of Behavior. Introduction and Chapter 4, The first stage of perception: growth and the assembly (reprint), in: Neurocomputing: foundations of research (Hrsg. Anderson, J. A. und Rosenfeld, E.), MIT Press: Cambridge, Mass.
- Hecht-Nielsen, R. (1987) Kolmogorov's mapping neural network existence theorem, in: Proceedings of the IEEE First International Conference on Neural Networks, San Diego, Vol. III (Hrsg. IEEE) SOS Printing, S. 11-14.
- Hecht-Nielsen, R. (1988) Applications of counterpropagation networks, in: Neural Networks, Vol. 1, S. 131-139.
- Hecht-Nielsen, R. (1989) Theory of the back-propagation neural network, in: Proceedings of the International Joint Conference on Neural Networks, San Diego, Vol. I (Hrsg. IEEE) SOS Printing.
- Hecht-Nielsen, R. (1990) Neurocomputing, Addison-Wesley Pub. Co.: Reading, Mass.

- Hecht-Nielsen, R. und McKenna, T. (2003) Computational models for neuroscience: human cortical information processing, Springer: London; New York.
- Hero, o. A. und Baldi, B. (1589) Di Herone Alessandrino De gli avtomati, Appresso Girolamo Porro: In Venetia.
- Hill, T., Marquez, L., O'Connor, M. und Remus, W. (1994a) Artificial Neural-Network Models for Forecasting and Decision-Making, in: International Journal of Forecasting, Vol. 10, S. 5-15.
- Hill, T., Marquez, L., O'Connor, M. und Remus, W. (1994b) Artificial Neural-Network Models for Forecasting and Decision-Making, in: International Journal of Forecasting, Vol. 10, S. 5-15.
- Hill, T., O'Connor, M. und Remus, W. (1996) Neural network models for time series forecasts, in: Management Science, Vol. 42, S. 1082-1092.
- Hillier, F. S. und Lieberman, G. J. (1997) Operations Research: Einführung, 5. Aufl. (unveränd. Nachdr. der 4. Aufl.), Oldenbourg: München [u.a.].
- Hippert, H. S., Pedreira, C. E. und Souza, R. C. (2001) Neural networks for short-term load forecasting: A review and evaluation, in: Ieee Transactions On Power Systems, Vol. 16, S. 44-55.
- Ho, K. L., Hsu, Y. Y. und Yang, C. C. (1992) Short-Term Load Forecasting Using a Multi-layer Neural Network with an Adaptive Learning Algorithm, in: Ieee Transactions on Power Systems, Vol. 7, S. 141-149.
- Hobbs, B. F., Helman, U., Jitprapaikularn, S., Konda, S. und Maratukulam, D. (1998) Artificial neural networks for short-term energy forecasting: Accuracy and economic value, in: Neurocomputing, Vol. 23, S. 71-84.
- Hochstädter, D. (1969) Stochastische Lagerhaltungsmodelle, Springer: Berlin.
- Hochstädter, D. und Uebe, G. (1970) Ökonometrische Methoden, Springer: Berlin u.a.
- Holley, L. H. und Karplus, M. (1989) Protein Secondary Structure Prediction With A Neural Network, in: Proceedings Of The National Academy Of Sciences Of The United States Of America, Vol. 86, S. 152-156.
- Holt, C. C. (1957) Forecasting seasonals and trends by exponentially weighted moving averages, in: ONR Reserach Memorandum, Vol. 52.
- Holt, C. C. (2004a) Author's retrospective on 'Forecasting seasonals and trends by exponentially weighted moving averages', in: International Journal of Forecasting, Vol. 20, S. 11-13.
- Holt, C. C. (2004b) Forecasting seasonals and trends by exponentially weighted moving averages, in: International Journal of Forecasting, Vol. 20, S. 5-10.
- Holt, M. J. J. (1992) Comparison Of Generalization In Multilayer Perceptrons With The Log-Likelihood And Least-Squares Cost-Functions, in: 11th Iaprr International Conference On Pattern Recognition, Proceedings, Vol Ii - Conference B: Pattern Recognition Methodology And Systems, I E E E, Computer Soc Press: Los Alamitos, S. 17-20.
- Holt, M. J. J. und Semnani, S. (1990) Convergence Of Back-Propagation In Neural Networks Using A Log-Likelihood Cost Function, in: Electronics Letters, Vol. 26, S. 1964-1965.
- Hopfield, J. J. (1984) SC! Neurons with graded response have collective computational properties like those of two-state neurons (reprint), in: Neurocomputing: foundations of research (Hrsg. Anderson, J. A. und Rosenfeld, E.), MIT Press: Cambridge, Mass., S. 579-583.
- Hopfield, J. J. und Tank, D. W. (1985) Neural Computation of Decisions in Optimization Problems, in: Biological Cybernetics, Vol. 52, S. 141-152.
- Horne, B., Jamshidi, M. und Vadiiee, N. (1990) Neural Networks In Robotics - A Survey, in: Journal Of Intelligent & Robotic Systems, Vol. 3, S. 51-66.

- Hornik, K., Stinchcombe, M. und White, H. (1989) Multilayer Feedforward Networks Are Universal Approximators, in: *Neural Networks*, Vol. 2, S. 359-366.
- Hornik, K., Stinchcombe, M. und White, H. (1990) Universal Approximation Of An Unknown Mapping And Its Derivatives Using Multilayer Feedforward Networks, in: *Neural Networks*, Vol. 3, S. 551-560.
- Hou, C. (1989) Digital simulation of artificial neural networks for business applications
- Hruschka, H. (1980) Kurzfristprognose ökonomischer Zeitreihen: eine empirischer Leistungsvergleich alternativer Verfahren, Deutsch: Thun [u.a.].
- Hruschka, H. (1992) Using neural networks for clustering-based market segmentation, Institut für Höhere Studien: Wien.
- Hrycej, T. (1997) *Neurocontrol: towards an industrial control methodology*, J. Wiley & Sons: New York.
- Hu, M. J. C. (1964) Application of the adaline system to weather forecasting, Technical Report 6775-1, Stanford University, Stanford Electronic Laboratories: San Francisco
- Hu, Y. H. und Hwang, J.-N. (2002) *Handbook of neural network signal processing*, CRC Press: Boca Raton, Fla., S. 1 v. (various pagings).
- Huang, D. S. (1999) Neural pattern analyses based on a novel error cost function, in: *International Conference On Parallel And Distributed Processing Techniques And Applications*, Vol Vi, Proceedings, C S R E A Press: Athens, S. 2875-2880.
- Huang, S. H. und Zhang, H. C. (1993) Neural Networks In Manufacturing - A Survey, in: *Fifteenth Ieee/Chmt International Electronics Manufacturing Technology Symposium - Electronics Manufacturing For The Year 2000*, IEEE: New York, S. 177-190.
- Huber, P. J. (1981) *Robust statistics*, Wiley: New York, NY [u.a.].
- Hubert, C. (1993) Pattern Completion With The Random Neural-Network Using The Rprop Learning Algorithm, in: *1993, International Conference On Systems, Man And Cybernetics: Systems Engineering In The Service Of Humans*, Vol 2 - Conference Proceedings, I E E E: New York, S. 613-617.
- Hunt, K. J., Irwin, G. und Warwick, K. (1995) *Neural network engineering in dynamic control systems*, Springer: Berlin; New York.
- Hunt, K. J., Sbarbaro, D., Zbikowski, R. und Gawthrop, P. J. (1992) Neural Networks For Control-Systems - A Survey, in: *Automatica*, Vol. 28, S. 1083-1112.
- Husken, M. und Stagge, P. (2003) Recurrent neural networks for time series classification, in: *Neurocomputing*, Vol. 50, S. 223-235.
- Husmeier, D. (1999) *Neural networks for conditional probability estimation: forecasting beyond point predictions*, Springer: London [u.a.].
- Husmeier, D. (2000) Learning non-stationary conditional probability distributions, in: *Neural Networks*, Vol. 13, S. 287-290.
- Husmeier, D. und Taylor, J. G. (1997) Predicting conditional probability densities of stationary stochastic time series, in: *Neural Networks*, Vol. 10, S. 479-497.
- Hussain, M. A. (1999) Review of the applications of neural networks in chemical process control - simulation and online implementation, in: *Artificial Intelligence In Engineering*, Vol. 13, S. 55-68.
- Hüttner, M. (1986) *Prognoseverfahren und ihre Anwendung*, de Gruyter: Berlin [u.a.].
- Ibnkahla, M. (2000) Applications of neural networks to digital communications - a survey, in: *Signal Processing*, Vol. 80, S. 1185-1215.
- Igel, C. und Husken, M. (2003) Empirical evaluation of the improved Rprop learning algorithms, in: *Neurocomputing*, Vol. 50, S. 105-123.

- Informationsgemeinschaft Münzspiel (2001) Taschenbuch der Unterhaltungsautomaten-Wirtschaft 2000, VDAI: Braunschweig.
- Irie, B. und Miyake, S. (1988) Capabilities of three-layered perceptrons, in: Proceedings of the IEEE International Conference on Neural Networks, Vol. I (Hrsg. IEEE) IEEE Press, S. 641-648.
- Irwin, G. W., Warwick, K., Hunt, K. J. und Institution of Electrical Engineers. (1995) Neural network applications in control, Institution of Electrical Engineers: London, U.K.
- Istook, E. J., III (2001) Improved backpropagation learning in neural networks with windowed momentum
- Jacobs, R. A. (1988) Increased Rates Of Convergence Through Learning Rate Adaptation, in: Neural Networks, Vol. 1, S. 295-307.
- Jagielska, I. und Jacob, A. (1993) A neural network model for sales forecasting, in: Proceedings of the First New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems, Vol., S. 284-287.
- Jain, C. L. (2003) Benchmarking forecasting practices, Graceway Publishing: Flushing, N.Y.
- Jain, L. C. und Vemuri, V. (1999) Industrial applications of neural networks, CRC Press: Boca Raton.
- Jalan, A. K., Giri, R. R. und Chaudhuri, K. S. (1996) EOQ model for items with Weibull distribution deterioration, shortages and trended demand, in: International Journal of Systems Science, Vol. 27, S. 851-855.
- Janakiraman, J. und Honavar, V. (1993) Adaptive Learning Rate For Increasing Learning Speed In Backpropagation Networks, in: Wcnn'93 - Portland, World Congress On Neural Networks, Vol Iv, Lawrence Erlbaum Assoc Publ: Mahwah, S. 378-381.
- Janetzke, P. und Falk, J. (2006) Der Beitrag der Künstlichen Intelligenz zur betrieblichen Prognose, in: Prognoserechnung (Hrsg. Mertens, P. und Rässler, S.), Physica-Verlag: Heidelberg, S. 305-333.
- Jerez-Aragones, J. M., Gomez-Ruiz, J. A., Ramos-Jimenez, G., Munoz-Perez, J. und Alba-Conejo, E. (2003) A combined neural network and decision trees model for prognosis of breast cancer relapse, in: Artificial Intelligence In Medicine, Vol. 27, S. 45-63.
- Jiang, G. (2003) Travel-Prediction for Urban Arterial Road: A case on China, in: Proceedings of the 2003 IEEE Conference on Intelligent Transportation Systems, Vol. 2 IEEE, S. 1459-1462.
- Jiang, J.-H., Wang, J.-H., Xia, C. und Yu, R.-Q. (1996) Neural network learning to non-linear principal component analysis, in: Analytica Chimica Acta, Vol. 336, S. 209-222.
- Jiang, J. J., Zhong, M. S. und Klein, G. (2000) Marketing category forecasting: An alternative of BVAR - Artificial neural networks, in: Decision Sciences, Vol. 31, S. 789-812.
- Jiao, S. M., Han, P., Zhou, L. H. und Li, J. B. (2004) Recurrent neural network applied in dynamic process identification based on RPROP and chaos optimization coupling algorithm, in: Proceedings Of The 2004 International Conference On Machine Learning And Cybernetics, Vols 1-7, Ieee: New York, S. 3339-3343.
- Joeressen, A. (1996) Systematisierung der Datenvorverarbeitung für die Prognose, in: Proceedings of the Sitzung der Arbeitsgruppen 'Prognoseverfahren' und 'Fuzzy Sets und Neuronale Netze' der DGOR, Stralsund, Vol.
- Joneja, D. (1991) Multiechelon Assembly Systems with Nonstationary Demands - Heuristics and Worst Case Performance Bounds, in: Operations Research, Vol. 39, S. 512-518.
- Joo, D., Hong, T. und Han, I. (2003) The neural network models for IDS based on the asymmetric costs of false negative errors and false positive errors, in: Expert Systems With Applications, Vol. 25, S. 69-75.

- Juang, J. G. (2000) Effects of using different neural network structures and cost functions in locomotion control, in: *Proceedings Of The 2000 American Control Conference, Vols 1-6*, I E E E: New York, S. 3523-3527.
- Juang, J. G. (2006) Effects of using different neural network structures and cost functions in locomotion control, in: *Advances In Natural Computation, Pt 1, Vol. 4221*, Springer-Verlag Berlin: Berlin, S. 437-446.
- Kaastra, I. und Boyd, M. (1996) Designing a neural network for forecasting financial and economic time series, in: *Neurocomputing, Vol. 10*, S. 215-236.
- Kalogirou, S. A. (2001) Artificial neural networks in renewable energy systems applications: a review, in: *Renewable & Sustainable Energy Reviews, Vol. 5*, S. 373-401.
- Kamgar-Parsi, B. und Kamgar-Parsi, B. (1987) *An efficient model of neural networks for optimization*, University of Maryland: College Park, Md.
- Kamiyama, N., Taguchi, A., Watanabe, H., Pechranin, N., Yoshida, Y. und Sone, M. (1993) Tuning Of Learning Rate And Momentum On Backpropagation, in: *Wcnn'93 - Portland, World Congress On Neural Networks, Vol Iv*, Lawrence Erlbaum Assoc Publ: Mahwah, S. 119-122.
- Kamran, F., Harley, R. G., Burton, B. und Habetler, T. G. (1996) An on-line trained neural network with an adaptive learning rate for a wide range of power electronic applications, in: *Pesc 96 Record - 27th Annual Ieee Power Electronics Specialists Conference, Vols I And Ii*, I E E E: New York, S. 1499-1505.
- Kang, S. Y. (1991) An Investigation of the Use of Feedforward Neural Networks for Forecasting (Neural Networks), in: *DAI, 52, no. 08B*, (1991): 4313.
- Kantsila, A., Lehtokangas, M. und Saarinen, J. (2004) Complex RPROP-algorithm for neural network equalization of GSM data bursts, in: *Neurocomputing, Vol. 61*, S. 239-258.
- Kappen, B. und Gielen, S. (1997) *Neural networks: best practice in Europe: proceedings of the Stichting Neurale Netwerken Conference 1997*, Amsterdam, the Netherlands, 22 May 1997, World Scientific: Singapore; River Edge, N.J.
- Kappen, B., Gielen, S. und Stichting Neurale Netwerken. (1995) *Neural networks: artificial intelligence and industrial applications: proceedings of the Third Annual SNN Symposium on Neural Networks, Nijmegen, The Netherlands, 14-15 September 1995*, Springer: Berlin; New York.
- Kehagias, A. und Petridis, V. (1997) Predictive modular neural networks for time series classification, in: *Neural Networks, Vol. 10*, S. 31-49.
- Kehagias, A. und Petridis, V. (2002) Pattern Classification and Clustering - Predictive Modular Neural Networks for Unsupervised Segmentation of Switching Time Series: The Data Allocation Problem, in: *IEEE transactions on neural networks, Vol. 13*, S. 18.
- Kezunovic, M. (1997) A survey of neural net applications to protective relaying and fault analysis, in: *Engineering Intelligent Systems For Electrical Engineering And Communications, Vol. 5*, S. 185-192.
- Khotanzad, A. und Lu, J.-H. (1990) Non-parametric prediction of AR processes using neural networks, in: *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, 1990, Albuquerque, NM, USA, Vol. 5*, S. 2551-2554.
- Kim, T. H. und White, H. (2003) Estimation, Inference, and Specification Testing for Possibly Misspecified Quantile Regressions, in: *Maximum Likelihood Estimation of Misspecified Models: Twenty Years Later* (Hrsg. Fomby, T. und Hill, R. C.), Elsevier: New York, S. 107-132.
- Kimoto, T. und Kazuo, A. (1990) Stock Market Prediction Systems with Modular Neural Networks, in: *Proceedings of the Proceedings of the 1990 International Joint Conference on Neural Nets, Vol. 1*, S. 1-6.

- Kirkpatrick, S., Gelatt, C. D. und Vecchi, M. P. (1983) Optimization by Simulated Annealing, in: Science, Vol. 220, S. 671-680.
- Klappenbach, R. und Steinitz, W. (Hrsg.) (1967-1977) Wörterbuch der deutschen Gegenwartssprache, Akademie-Verlag 1967: Berlin.
- Klein, T. (1976) Lagerbewirtschaftung, Sortimentsgestaltung: ABC-Analysen als Entscheidungshilfen f. d. Unternehmensleitung, Forster: Zürich.
- Klemenc, J. und Fajdiga, M. (2002) A neural network approach to the simulation of load histories by considering the influence of a sequence of rainflow load cycles, in: International Journal of Fatigue, Vol. 24, S. 1109-1125.
- Klemm, H. und Mikut, M. (1972) Lagerhaltungsmodelle: Theorie und Anwendung, Verlag Die Wirtschaft: Berlin.
- Kline, D. M. und Berardi, V. L. (2005) Revisiting squared-error and cross-entropy functions for training neural network classifiers, in: Neural Computing & Applications, Vol. 14, S. 310-318.
- Klingst, A. (1979) Optimale Lagerhaltung: Wann und wieviel bestellen? Physica-Verl.: Würzburg [u.a.].
- Kluge, F. (2002) Etymologisches Wörterbuch der deutschen Sprache, [Elektronische Resource], 24. Aufl, de Gruyter: Berlin.
- Kneller, D. G., Cohen, F. E. und Langridge, R. (1990) Improvements In Protein Secondary Structure Prediction By An Enhanced Neural Network, in: Journal Of Molecular Biology, Vol. 214, S. 171-182.
- Koenker, R. (2005) Quantile regression, Cambridge Univ. Press: Cambridge u.a.
- Koenker, R. und Bassett, G. (1978) Regression Quantiles, in: Econometrica, Vol. 46, S. 33-50.
- Koenker, R. W. und Dorey, V. (1987) Computing Regression Quantiles, in: Applied Statistics-Journal of the Royal Statistical Society Series C, Vol. 36, S. 383-393.
- Kohonen, T. (1982) Analysis of a Simple Self-Organizing Process, in: Biological Cybernetics, Vol. 44, S. 135-140.
- Kohonen, T. (1985) Self-Organized Feature Maps, in: Journal of the Optical Society of America a-Optics Image Science and Vision, Vol. 2, S. P16-P16.
- Kohonen, T. (1989) Self-organization and associative memory, 3. ed, Springer: Berlin [u.a.].
- Kohonen, T. (2001) Self-organizing maps, 3rd, Springer: Berlin; New York.
- Kolarik, T. und Rudorfer, G. (1994) Time Series Forecasting Using Neural Networks, in: Proceedings of the Proceedings of the international conference on APL, Antwerp, Belgium, Vol., S. 86-94.
- Kong, J. H. L. und Martin, G. (1995) A backpropagation neural network for sales forecasting, in: 1995 IEEE International Conference On Neural Networks Proceedings, Vols 1-6, IEEE: New York, S. 1007-1011.
- Koopmann, J. (1988) Die Geschichte der Warenautomaten, in: Der Automatenvertrieb: theoretische, empirische und literarische Untersuchungen über eine "vergessene" Betriebsform des Einzelhandels (Hrsg. Schenk, H.-O.), Duisburg, S. 67-68.
- Kosiol, E. (1975) Zur Problematik der Planung in der Unternehmung, in: Unternehmungsplanung (Hrsg. Wild, J.), Rowohlt: Reinbeck, S. 38-57.
- Kotsialos, A., Papageorgiou, M. und Poulimenos, A. (2005) Long-term sales forecasting using Holt-Winters and neural network methods, in: Journal Of Forecasting, Vol. 24, S. 353-368.
- KPMG (2005) Trends im Handel 2005, Technical Report, (Hrsg. KPMG), Consumer Markets & Retail Reports: Köln, 73 Seiten.

- KPMG und EHI (2006) Trends im Handel 2010, Technical Report, (Hrsg. KPMG), Consumer Markets & Retail Reports: Köln, 78 Seiten.
- Krogh, A. und Hertz, J. A. (1992) A Simple Weight Decay Can Improve Generalization, in: *Advances In Neural Information Processing Systems 4*, Vol. 4, Morgan Kaufmann Pub Inc: San Mateo, S. 950-957.
- Krycha, K. A. und Wagner, U. (1999) Applications of artificial neural networks in management science: a survey, in: *Journal of Retailing and Consumer Services*, Vol. 6, S. 185-203.
- Kuo, R. J. (1997) An intelligent sales forecasting system through fuzzy neural network, in: *1997 IEEE International Conference On Neural Networks*, Vol. 1, IEEE: New York, S. 505-509.
- Kuo, R. J. (2001) A sales forecasting system based on fuzzy neural network with initial weights generated by genetic algorithm, in: *European Journal Of Operational Research*, Vol. 129, S. 496-517.
- Kuo, R. J., Wu, P. und Wang, C. P. (2002) An intelligent sales forecasting system through integration of artificial neural networks and fuzzy neural networks with fuzzy weight elimination, in: *Neural Networks*, Vol. 15, S. 909-925.
- Kuo, R. J. und Xue, K. C. (1998a) A decision support system for sales forecasting through fuzzy neural networks with asymmetric fuzzy weights, in: *Decision Support Systems*, Vol. 24, S. 105-126.
- Kuo, R. J. und Xue, K. C. (1998b) An intelligent sales forecasting system through integration of artificial neural network and fuzzy neural network, in: *Computers In Industry*, Vol. 37, S. 1-15.
- Kuo, R. J. und Xue, K. C. (1999) Fuzzy neural networks with application to sales forecasting, in: *Fuzzy Sets And Systems*, Vol. 108, S. 123-143.
- Küsters, U. (2006) Evaluation, Kombination und Auswahl betriebswirtschaftlicher Prognoseverfahren, in: *Prognoserechnung* (Hrsg. Mertens, P. und Rässler, S.), Physica-Verlag: Heidelberg, S. 367-404.
- Kwok, T. Y. und Yeung, D. Y. (1997) Objective functions for training new hidden units in constructive neural networks, in: *Ieee Transactions On Neural Networks*, Vol. 8, S. 1131-1148.
- Laguna, M. und Martái, R. (2002) Neural network prediction in a system for optimizing simulations, in: *IIE Transactions*, Vol. 34, S. 10.
- Lapedes, A. und Farber, R. (1988) How neural nets work, in: *Neural Information Processing Systems* (Hrsg. Anderson, D. Z.), American Institute of Physics: New York, S. 442-456.
- Lapedes, A., Farber, R. und Los Alamos National Laboratory. (1987) Nonlinear signal processing using neural networks: prediction and system modelling, Technical Report, Nr. LA-UR-87-2662, Los Alamos National Laboratory: Los Alamos, N.M., 36 Seiten.
- Lappalainen, H. (1998) Using an MDL-based cost function with neural networks, in: *Ieee World Congress On Computational Intelligence*, I E E E: New York, S. 2384-2389.
- Larson, C. M., Weigand, R. E. und Wright, J. S. (1976) *Basic retailing*, Prentice-Hall: Englewood Cliffs, N.J.
- Lau, H. S. und Lau, A. H. L. (2003) Nonrobustness of the normal approximation of lead-time demand in a (Q, R) system, in: *Naval Research Logistics*, Vol. 50, S. 149-166.
- Lau, H. S. und Zaki, A. (1982) The Sensitivity of Inventory Decisions to the Shape of Lead Time-Demand Distribution, in: *Iie Transactions*, Vol. 14, S. 265-271.
- Laudeman, I. V. An Application of Unsupervised Neural Networks and Fuzzy Clustering in the Identification of Structure in Personality Data, in: *DAI*, 56, no. 05B, (1994): 2922.

- Lee, H. L. und Billington, C. (1993) Material Management In Decentralized Supply Chains, in: *Operations Research*, Vol. 41, S. 835-847.
- Lee, H. L. und Nahmias, S. (1993) Single-Product, Single-Location Models, in: *Logistics of production and inventory* (Hrsg. Graves, S. C., Rinnooy Kan, A. H. G. und Zipkin, P. H.), North-Holland: Amsterdam [u.a.], S. 14-60.
- Lee, K. D., Booth, D. und Alam, P. (2005) A comparison of supervised and unsupervised neural networks in predicting bankruptcy of Korean firms, in: *Expert Systems With Applications*, Vol. 29, S. 1-16.
- Lee, S. C. und Kim, Y. H. (2002) An enhanced Lagrangian neural network for the ELD problems with piecewise quadratic cost functions and nonlinear constraints, in: *Electric Power Systems Research*, Vol. 60, S. 167-177.
- Lee, T. S. und Chen, I. F. (2005) A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines, in: *Expert Systems With Applications*, Vol. 28, S. 743-752.
- Lensing, M. und Sonnemann, K. (1995) *Materialwirtschaft und Einkauf: Organisation, Bedarfsermittlung, Beschaffungsplanung, Materialdisposition, ABC-Analyse, Einkaufspolitik, Beschaffungsmarketing, Bestellung, Vertrag, Terminsicherung*, Gabler: Wiesbaden.
- Leski, J. und Henzel, N. (2003) An epsilon-insensitive learning in neuro-fuzzy modeling, in: *Neural Networks And Soft Computing*, Physica-Verlag GmbH & Co: Heidelberg, S. 531-536.
- Levine, D. S. (2000) *Introduction to neural and cognitive modeling*, 2nd, Lawrence Erlbaum Associates Publishers: Mahwah, N.J.
- Lewandowski, R. (1974) *Prognose- und Informationssysteme und ihre Anwendungen*, de Gruyter: Berlin [u.a.].
- Li, E. Y. (1994) Artificial neural networks and their business applications, in: *Information & Management*, Vol. 27, S. 303-313.
- Li, M., Mehrotra, K., Mohan, C. und Ranka, S. (1990) Sunspot numbers forecasting using neural networks, in: *Proceedings of the 5th IEEE International Symposium on Intelligent Control*, Vol. 1, S. 524-529.
- Liang, X. B. und Wang, J. (2000) A recurrent neural network for nonlinear optimization with a continuously differentiable objective function and bound constraints, in: *Ieee Transactions On Neural Networks*, Vol. 11, S. 1251-1262.
- Liao, K. P. und Fildes, R. (2005) The accuracy of a procedural approach to specifying feedforward neural networks for forecasting, in: *Computers & Operations Research*, Vol. 32, S. 2151-2169.
- Lisboa, P. J. G., Edisbury, B. und Vellido, A. (2000) *Business applications of neural networks: the state-of-the-art of real-world applications*, World Scientific: Singapore; River Edge, N.J.
- Litterman, R. (1986) Forecasting with Bayesian vector autoregressions -- Five years of experience: Robert B. Litterman, *Journal of Business and Economic Statistics* 4 (1986) 25-38, in: *International Journal of Forecasting*, Vol. 2, S. 497-498.
- Liu, B. Y. und Ye, L. Y. (2004) Review on application of artificial neural networks in earthquake resistant and disaster prevention, in: *Proceedings Of The Eighth International Symposium On Structural Engineering For Young Experts*, Vols 1 And 2, Science Press Beijing: Beijing, S. 605-611.
- Lo, L. (1979) *Prognoseinformationen für kundendienstpolitische Entscheidungen: dargestellt an Beispielen des Fotomarktes*.
- Loderhose, B. (2006) *Ganzheitliche Optimierung der Supply Chain soll Kosen drücken*, Lebensmittelzeitung, 26.05.20006.

- Lohrbach, T. (1994) Einsatz von künstlichen neuronalen Netzen für ausgewählte betriebswirtschaftliche Aufgabenstellungen und Vergleich mit konventionellen Lösungsverfahren, Unitext-Verl.: Göttingen.
- Lorentz, J. und Yuh, J. (1996) A survey and experimental study of neural network AUV control, in: Proceedings Of The 1996 Symposium On Autonomous Underwater Vehicle Technology, IEEE: New York, S. 109-116.
- Luxhoj, J. T., Riis, J. O. und Stensballe, B. (1996) A hybrid econometric-neural network modeling approach for sales forecasting, in: International Journal Of Production Economics, Vol. 43, S. 175-192.
- Lyerly, S. B. (1952) The Average Spearman Rank Correlation Coefficient, in: Psychometrika, Vol. 17, S. 421-428.
- Maass, W. und Bishop, C. M. (1999) Pulsed neural networks, MIT Press: Cambridge, Mass.
- Magee, J. F. (1958) Production planning and inventory control, McGraw-Hill: New York [u.a.].
- Maglaveras, N., Stamkopoulos, T., Diamantaras, K., Pappas, C. und Strintzis, M. (1998) ECG pattern recognition and classification using non-linear transformations and neural networks: A review, in: International Journal Of Medical Informatics, Vol. 52, S. 191-208.
- Maier, H. R. und Dandy, G. C. (2000a) Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications, in: Environmental Modelling and Software, Vol. 15, S. 101-124.
- Maier, H. R. und Dandy, G. C. (2000b) Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications, in: Environmental Modelling & Software, Vol. 15, S. 101-124.
- Makridakis, S. (1993) Accuracy Measures - Theoretical And Practical Concerns, in: International Journal Of Forecasting, Vol. 9, S. 527-529.
- Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E. und Winkler, R. (1982) The Accuracy Of Extrapolation (Time-Series) Methods - Results Of A Forecasting Competition, in: Journal Of Forecasting, Vol. 1, S. 111-153.
- Makridakis, S., Chatfield, C., Hibon, M., Lawrence, M., Mills, T., Ord, K. und Simmons, L. F. (1993) The M2-competition: A real-time judgmentally based forecasting study, in: International Journal of Forecasting, Vol. 9, S. 5-22.
- Makridakis, S. und Hibon, M. (1979) Accuracy Of Forecasting - Empirical-Investigation, in: Journal Of The Royal Statistical Society Series A-Statistics In Society, Vol. 142, S. 97-145.
- Makridakis, S. und Hibon, M. (2000) The M3-Competition: results, conclusions and implications, in: International Journal Of Forecasting, Vol. 16, S. 451-476.
- Makridakis, S. G., Reschke, H. und Wheelwright, S. C. (1980) Prognosetechniken für Manager, Gabler: Wiesbaden.
- Makridakis, S. G., Wheelwright, S. C. und Hyndman, R. J. (1998) Forecasting: methods and applications, Wiley: New York.
- Mangasarian, O. L. und Musicant, D. R. (2000) Robust linear and support vector regression, in: IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 22, S. 950-955.
- Marquez, L., Hill, T., O'Connor, M. und Remus, W. (1992) Neural network models for forecast: a review, in: Proceedings of the System Sciences, 1992. Proceedings of the Twenty-Fifth Hawaii International Conference on, Vol. iv, S. 494-498 vol.4.
- Massey, F. J. (1951) The Kolmogorov-Smirnov Test For Goodness Of Fit, in: Journal Of The American Statistical Association, Vol. 46, S. 68-78.

- Mastorocostas, P. A. und Rekanos, I. T. (2005) Simulated annealing dynamic RPROP for training recurrent fuzzy systems, in: Fuzz-Ieee 2005: Proceedings Of The Ieee International Conference On Fuzzy Systems - Biggest Little Conference In The World, Ieee: New York, S. 1086-1091.
- Matsuda, S. (1998) "Optimal" neural representation for combinatorial optimization with linear cost function (extended abstract), in: Ieee World Congress On Computational Intelligence, I E E E: New York, S. 1657-1660.
- Mazanec, J. (1999) Marketing applications of neural networks, Elsevier Science: New York.
- McCulloch, W. S. und Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity (reprint in 1988), in: Neurocomputing: foundations of research (Hrsg. Anderson, J. A. und Rosenfeld, E.), MIT Press: Cambridge, Mass.
- McDowall, T. M. (1997) A robust partial least squares neural network
- McGregor, M. J., Flores, T. P. und Sternberg, M. J. E. (1989) Prediction Of Beta-Turns In Proteins Using Neural Networks, in: Protein Engineering, Vol. 2, S. 521-526.
- Medsker, L. R. und Jain, L. C. (2001) Recurrent neural networks design and applications, CRC Press: Boca Raton, FL.
- Meese, R. A. und Geweke, J. (1984) A comparison of autoregressive univariate forecasting procedures for macroeconomic time series, in: Journal of Business & Economic Statistics, S. 191-200.
- Meffert, H. (1980) Marketing: Einführung in die Absatzpolitik, 5. Aufl, Gabler: Wiesbaden.
- Meffert, H. (2000) Marketing, 9., überarb. und erw. Aufl, Gabler: Wiesbaden.
- Meffert, H. und Steffenhagen, H. (1977) Marketing-Prognosemodelle: quantitative Grundlagen des Marketing, Poeschel: Stuttgart.
- Meireles, M. R. G., Almeida, P. E. M. und Simoes, M. G. (2003) A comprehensive review for industrial applicability of artificial neural networks, in: Ieee Transactions On Industrial Electronics, Vol. 50, S. 585-601.
- Mentzer, J. T. und Krishnan, R. (1985) The Effect of the Assumption of Normality on Inventory Control / Customer Service, in: Journal of Business Logistics, Vol. 6, S. 101-123.
- Mertens, P. und Deistler, M. (1994) Prognoserechnung, 5., neu bearb. und erw. Aufl, Physica-Verl.: Heidelberg.
- Meuser, R. (1998) Vom Bestand zur Bestellung - Dispositionsverfahren im Überblick, in: Automatische Disposition - Bestandsaufnahme und Perspektiven (Hrsg. Atzberger, M. und Eurohandelsinstitut), DHI - Deutsches Handelsinstitut: Köln, S. 46-51.
- Michel, A. N. und Liu, D. (2002) Qualitative analysis and synthesis of recurrent neural networks, Marcel Dekker Inc.: New York.
- Mihalakakou, G., Flocas, H. A., Santamouris, M. und Helmis, C. G. (2002) Application of neural networks to the simulation of the heat island over Athens, Greece, using synoptic types as a predictor, in: Journal of Applied Meteorology, Vol. 41, S. 519-527.
- Mihalakakou, G., Santamouris, M. und Asimakopoulos, D. N. (2000) The total solar radiation time series simulation in Athens, using neural networks, in: Theoretical and applied Climatology, Vol. 66, S. 185-197.
- Miller, A. S., Blott, B. H. und Hames, T. K. (1992) Review Of Neural Network Applications In Medical Imaging And Signal-Processing, in: Medical & Biological Engineering & Computing, Vol. 30, S. 449-464.
- Minai, A. A. und Williams, R. D. (1990) Back-Propagation Heuristics - A Study Of The Extended Delta-Bar-Delta Algorithm, in: Ijcn International Joint Conference On Neural Networks, Vols 1-3, IEEE: New York, S. A595-A600.
- Minner, S. (2000) Strategic safety stocks in supply chains, Springer: Berlin [u.a.].

- Minsky, M. L. (1968) *Semantic information processing*, MIT Press: Cambridge, Mass.
- Minsky, M. L. und Papert, S. (1969) *Perceptrons; an introduction to computational geometry*, MIT Press: Cambridge, Mass.
- Minsky, M. L. und Papert, S. (1973) *Artificial intelligence*, Oregon State System of Higher Education; [distributor: Distribution Center: Eugene.
- Minsky, M. L. und Papert, S. (1988) *Perceptrons: an introduction to computational geometry*, Expanded, MIT Press: Cambridge, Mass.
- Monßen, N. (2001) *Die Entwicklung des Vending-Marktes*, Technical Report, (Hrsg. BDV), Presseinformationen des BDV, BDV: Köln, 25 Seiten.
- Montgomery, D. C., Johnson, L. A. und Gardiner, J. S. (1990) *Forecasting and time series analysis*, 2nd, McGraw-Hill: New York.
- Moody, J. (1998) *Forecasting the economy with neural nets: A survey of challenges and solutions*, in: *Neural Networks: Tricks Of The Trade*, Vol. 1524, Springer-Verlag Berlin: Berlin, S. 347-371.
- Moody, J. E. (1992) *The Effective Number of Parameters: An Analysis of Generalization and Regularization in Nonlinear Learning Systems*, in: *Advances in Neural Information Processing Systems 4* (Hrsg. Moody, J. E., Hanson, S. J. und Lippmann, R. P.), S. 847-854.
- Moon, M. A., Mentzer, J. T. und Smith, C. D. (2003) *Conducting a sales forecasting audit*, in: *International Journal of Forecasting*, Vol. 19, S. 5-25.
- Moors, J. J. A. und Strijbosch, L. W. G. (2002) *Exact fill rates for (R, s, S) inventory control with gamma distributed demand*, in: *Journal of the Operational Research Society*, Vol. 53, S. 1268-1274.
- Morse, P. M. und Kimball, G. E. (1951) *Methods of operations research*, (3. print.), Technology Press [u.a.]: New York.
- Moshiri, S. (1998) *Forecasting Inflation Using Econometric and Artificial Neural Network Models (Econometric Models)*, DAI, 59, no. 10A, (1998): 3897
- Moshiri, S. und Cameron, N. (2000) *Neural network versus econometric models in forecasting inflation*, in: *Journal of Forecasting*, Vol. 19, S. 201-217.
- Mosteller, F. und Tukey, J. (1980) *Robust Estimation Procedures*, Addison-Wesley: New York.
- Mozer, M. C. (1994) *Neural net architectures for temporal sequence processing*, in: *Time Series Prediction: Forecasting the Future and Understanding the Past* (Hrsg. Weigend, A. S. und Gershenfeld, N. A.), Addison-Wesley: Reading, MA, S. 243-264.
- Muckstadt, J. A. und Roundy, R. O. (1993) *Analysis of Multistage Production Systems*, in: *Logistics of production and inventory* (Hrsg. Graves, S. C., Rinnooy Kan, A. H. G. und Zipkin, P. H.), North-Holland: Amsterdam [u.a.], S. 59-131.
- Mujtaba, I. M. und Hussain, M. A. (2001) *Application of neural networks and other learning technologies in process engineering*, ICP: River Edge, NJ.
- Müller-Hagedorn, L. (1998) *Der Handel*, Kohlhammer: Stuttgart [u.a.].
- Muller, B. (1997) *Neural networks as a means of precise day-to-day sales prediction for milk products*, in: *Kieler Milchwirtschaftliche Forschungsberichte*, Vol. 49, S. 85-94.
- Müller, B., Reinhardt, J. und Strickland, M. T. (1995) *Neural networks: an introduction*, 2nd updated and corr., Springer: Berlin; New York.
- Müller, C. (2002) *Simulation sozialer Netzwerke mit neuronalen Netzen*.
- Mundie, D. B. und Massengill, L. W. (1991) *Weight Decay And Resolution Effects In Feedforward Artificial Neural Networks*, in: *Ieee Transactions On Neural Networks*, Vol. 2, S. 168-170.

- Nachtsheim, P. R. (1994) A 1st Order Adaptive Learning Rate Algorithm For Back Propagation Networks, in: 1994 Ieee International Conference On Neural Networks, Vol 1-7, I E E E: New York, S. 257-262.
- Naddor, E. (1971) Lagerhaltungssysteme, Deutsch: Frankfurt/M. [u.a.].
- Naddor, E. (1978) Sensitivity to distributions in inventory systems, in: Management Science, Vol. 24 INFORMS: Institute for Operations Research, S. 1769.
- Nahmias, S. (1982) Perishable Inventory-Theory - a Review, in: Operations Research, Vol. 30, S. 680-708.
- Nahmias, S. und Demmy, W. S. (1982) The Logarithmic Poisson Gamma-Distribution - a Model for Leadtime Demand, in: Naval Research Logistics, Vol. 29, S. 667-677.
- Nam, K. und Schaefer, T. (1995) Forecasting international airline passenger traffic using neural networks, in: Logistics and Transportation Review, Vol. 31, S. p. 239-251.
- Nam, K., Yi, J. und Prybutok, V. R. (1997) Predicting airline passenger volume, in: The Journal of Business Forecasting Methods & Systems, Vol. 16, S. 14-16.
- NAMA (2005) 2005 Concensus of the Industry, Vending Times: New York.
- NAMA (2006) Vending Industry Definitions, Vol. 2006 (Hrsg. NAMA) Chicago, <http://www.vending.org>
- Nayak, R., Jain, L. C. und Ting, B. K. H. (2001) Artificial neural networks in biomedical engineering: A review, in: Computational Mechanics, Vols 1 And 2, Proceedings - New Frontiers For The New Millennium, Elsevier Science Bv: Amsterdam, S. 887-892.
- Nelson, M., Hill, T., Remus, B. und O'Connor, M. (1994) Can neural networks applied to time series forecasting learn seasonal patterns: an empirical investigation, in: Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences, 1994. Vol.III: Information Systems: Decision Support and Knowledge-Based Systems, Vol. 3, S. 649-655.
- Nelson, M., Hill, T., Remus, W. und O'Connor, M. (1999) Time series forecasting using neural networks: Should the data be deseasonalized first? in: Journal of Forecasting, Vol. 18, S. 359-367.
- Nerlove, M., Grether, D. M. und Carvalho, J. L. (1995) Analysis of economic time series: a synthesis, Rev., 1. [Dr.]. Academic Press: San Diego.
- Neuneier, R. und Zimmermann, H.-G. (1998) How to Train Neural Networks, in: Neural networks: tricks of the trade (Hrsg. Orr, G. und Müller, K.-R.), Springer: Berlin; New York, S. 373-423.
- Neuralware (1991) Neural Computing: Pittsburgh.
- Newbold, P. und Granger, C. W.J. (1974) Experience With Forecasting Univariate Time Series And Combination Of Forecasts, in: Journal Of The Royal Statistical Society Series A-Statistics In Society, Vol. 137, S. 131-165.
- Ng, G. W. (1997) Application of neural networks to adaptive control of nonlinear systems, Wiley: Taunton, New York.
- Nieschlag, R., Dichtl, E. und Hörschgen, H. (2002) Marketing, 19., überarb. und erg. Aufl., Duncker & Humblot: Berlin.
- Nieschlag, R. und Kuhn, G. (1980) Binnenhandel und Binnenhandelspolitik, 3., Duncker & Humblot: Berlin.
- Nørgaard, M. (2000) Neural networks for modelling and control of dynamic systems: a practitioner's handbook, Springer: London; New York.
- Nunnari, G. und Cannavo, F. (2005) Modified cost functions for modelling air quality time series by using neural networks, in: Artificial Neural Networks: Formal Models And Their Applications - Icann 2005, Pt 2, Proceedings, Vol. 3697, Springer-Verlag Berlin: Berlin, S. 723-728.

- o.V. (1998) Neural nets help Texas detect Medicaid fraud, in: *Expert Systems*, Vol. 15, S. 217-217.
- Oeldorf, G. und Olfert, K. (2002) *Materialwirtschaft*, 10., durchges. und aktualisierte Aufl, Kiehl: Ludwigshafen (Rhein).
- Oh, K. J., Roh, T. H. und Moon, M. S. (2005) Developing time-based clustering neural networks to use change-point detection: Application to financial time series, in: *Asia-Pacific Journal Of Operational Research*, Vol. 22, S. 51-70.
- Oh, S. H. und Lee, S. Y. (1999) A new error function at hidden layers for fast training of multilayer perceptrons, in: *Ieee Transactions On Neural Networks*, Vol. 10, S. 960-964.
- Omidvar, O. und Elliott, D. L. (1997) *Neural systems for control*, Academic Press: San Diego.
- Ong, E. und Flitman, A. (1998) Using neural networks for sales forecasting: A comparison of methods, in: *Progress In Connectionist-Based Information Systems*, Vols 1 And 2, Springer: Singapore, S. 1305-1308.
- Ord, K., Hibon, M. und Makridakis, S. (2000a) *The M3-Competition*, North-Holland: Amsterdam.
- Ord, K., Hibon, M. und Makridakis, S. (2000b) The M3-Competition, in: *International Journal Of Forecasting*, Vol. 16, S. 433-436.
- Orra, M. und Serpen, G. (2005) An exploratory study for neural net forecasting of retail sales trends using economic indicators, in: *Proceedings Of The 8th Joint Conference On Information Sciences*, Vols 1-3, Joint Conference Information Sciences: Durham, S. 875-878.
- Osborn, D. R., Heravi, S. und Birchenhall, C. R. (1999) Seasonal unit roots and forecasts of two-digit European industrial production, in: *International Journal of Forecasting*, Vol. 15, S. 27-47.
- Osman, H. und Blostein, S. D. (1999) New cost function for backpropagation neural networks with application to SAR imagery classification, in: *Automatic Target Recognition IX*, Vol. 3718, Spie - Int Soc Optical Engineering: Bellingham, S. 110-117.
- Osowski, S., Bojarczak, P. und Stodolski, M. (1996) Fast second order learning algorithm for feedforward multilayer neural networks and its applications, in: *Neural Networks*, Vol. 9, S. 1583-1596.
- Osowski, S. und Siwek, K. (2002) Regularisation of neural networks for improved load forecasting in the power system, in: *Iee Proceedings-Generation Transmission and Distribution*, Vol. 149, S. 340-344.
- Ou, C. Y. und Chen, F. L. (2006) Development of the time-series forecasting model by an artificial neural network in the CVS ordering system, in: *Twelfth Issat International Conference Reliability And Quality In Design*, Proceedings, Int Soc Sci Appl Technol: Piscataway, S. 108-112.
- Ozaki, T. (1980) Non-Linear Time-Series Models For Non-Linear Random Vibrations, in: *Journal Of Applied Probability*, Vol. 17, S. 84-93.
- Ozaki, T. (1992) A Bridge Between Nonlinear Time-Series Models And Nonlinear Stochastic Dynamic-Systems - A Local Linearization Approach, in: *Statistica Sinica*, Vol. 2, S. 113-135.
- Ozaki, T. und Thomson, P. (2002) A non-linear dynamic model for multiplicative seasonal-trend decomposition, in: *Journal Of Forecasting*, Vol. 21, S. 107-124.
- Page, P. (2005) *Building Inventory's Crystal Ball*, in: *Traffic World*.
- Pankratz, A. (1991) *Forecasting with dynamic regression models*, Wiley: New York.
- Pant, P. N. und Starbuck, W. H. (1990) Innocents In The Forest - Forecasting And Research Methods, in: *Journal Of Management*, Vol. 16, S. 433-460.

- Park, D. C., Elsharkawi, M. A., Marks, R. J., Atlas, L. E. und Damborg, M. J. (1991) Electric-Load Forecasting Using An Artificial Neural Network, in: *Ieee Transactions On Power Systems*, Vol. 6, S. 442-449.
- Park, J. H., Kim, Y. S., Eom, I. K. und Lee, K. Y. (1993) Economic Load Dispatch For Piecewise Quadratic Cost Function Using Hopfield Neural-Network, in: *Ieee Transactions On Power Systems*, Vol. 8, S. 1030-1038.
- Parsian, A. (1990) On The Admissibility Of An Estimator Of A Normal-Mean Vector Under A Linex Loss Function, in: *Annals Of The Institute Of Statistical Mathematics*, Vol. 42, S. 657-669.
- Parsian, A. und Farsipour, N. S. (1993) On The Admissibility And Inadmissibility Of Estimators Of Scale-Parameters Using An Asymmetric Loss Function, in: *Communications In Statistics-Theory And Methods*, Vol. 22, S. 2877-2901.
- Parsian, A. und Farsipour, N. S. (1997) Estimation of parameters of exponential distribution in the truncated space using asymmetric loss function, in: *Statistical Papers*, Vol. 38, S. 423-443.
- Parsian, A. und Farsipour, N. S. (1999) Estimation of the mean of the selected population under asymmetric loss function, in: *Metrika*, Vol. 50, S. 89-107.
- Pascarella, S. und Bossa, F. (1989) Pronet - A Microcomputer Program For Predicting The Secondary Structure Of Proteins With A Neural Network, in: *Computer Applications In The Biosciences*, Vol. 5, S. 319-320.
- Patterson, P., Van Ispelen, P. und Kagan, A. (1997) Forecasting US commodity exports using intra-year sales data: A neural network approach, in: *American Journal Of Agricultural Economics*, Vol. 79, S. 1717-1718.
- Pearlmutter, B. (1990) Dynamic recurrent neural networks, Carnegie Mellon University Computer Science Dept.: Pittsburgh, Pa.
- Pegels, C. C. (1967) Exponential Forecasting: Some new Variations, in: *Management Science*, Vol. 15, S. 311-315.
- Pendharkar, P. C. (2005) A threshold-varying artificial neural network approach for classification and its application to bankruptcy prediction problem, in: *Computers & Operations Research*, Vol. 32, S. 2561-2582.
- Peng, T. M., Hubele, N. F. und Karady, G. G. (1990) Conceptual approach to the application of neural network for short-term load forecasting, in: *Proceedings of the Circuits and Systems, 1990., IEEE International Symposium on*, Vol., S. 2942-2945 vol.4.
- Perez, M. (2006) Artificial neural networks and bankruptcy forecasting: a state of the art, in: *Neural Computing & Applications*, Vol. 15, S. 154-163.
- Perkel, D. H. (1990) Computational Neuroscience: Scope and Structure, in: *Computational Neuroscience* (Hrsg. Schwartz, E. L.), MIT Press: Cambridge, Mass.
- Petranek, S., Krajca, V. und Horejsi, J. (1997) Neural network based principal components analysis for EEG pre-processing and analysis, in: *Electroencephalography and Clinical Neurophysiology*, Vol. 103, S. 115.
- Petridis, V. und Kehagias, A. (1996) Modular neural networks for MAP classification of time series and the partition algorithm, in: *Ieee Transactions On Neural Networks*, Vol. 7, S. 73-86.
- Pfohl, H.-C. (2004) *Logistiksysteme: betriebswirtschaftliche Grundlagen*, 7., korrigierte und aktualisierte Aufl, Springer: Berlin [u.a.].
- Phansalkar, V. V. und Sastry, P. S. (1994) Analysis Of The Backpropagation Algorithm With Momentum, in: *Ieee Transactions On Neural Networks*, Vol. 5, S. 505-506.
- Phillips, P. C. B. (1996) Econometric Model Determination, in: *Econometrica*, Vol. 64 The Econometric Society, S. 763.

- Pichler, S. (1989) Die direkte Produkt-Rentabilität und ihre Bedeutung für Hersteller und Handel, Poeschel: Stuttgart.
- Poggio, G. F. und Girosi, F. (1989) A Theory of networks for approximation and learning, Technical Report, Nr. 1440, (Hrsg. Laboratory, A. I.), AI Memo, MIT Artificial Intelligence Laboratory: Boston.
- Poggio, T. und Girosi, F. (1990) Networks for approximation and learning, in: Proceedings of the IEEE, Vol. 78, S. 1481-1497.
- Poh, H. L. und Jasic, T. (1995) Forecasting and analysis of marketing data using neural networks: a case of advertising and promotion impact, in: Proceedings of the 11th Conference on Artificial Intelligence for Applications, Vol., S. 224-230.
- Potvin, J.-Y., Robillard, C. und Centre for Research on Transportation (Montréal Quebec) (1993) Clustering for vehicle routing with a competitive neural network, Centre for Research on Transportation = Centre de recherche sur les transports: Montréal.
- Poulton, M. M. (2002) Neural networks as an intelligence amplification tool: A review of applications, in: Geophysics, Vol. 67, S. 979-993.
- Preßmar, D. B. (1978) Beschreibungssprachen für betriebliche Informationssysteme, in: Entwicklungstendenzen der Systemanalyse: 5. Wirtschaftsinformatik-Symposium d. IBM Deutschland GmbH, Bad Neuenahr, 10. - 12. Okt. 1978, Vol. Fachberichte und Referate // Science Research Associates GmbH, Stuttgart; 6 (Hrsg. Hansen, H. R.), Oldenbourg: München, S. 262-275.
- Preßmar, D. B. (1995) Computergestützte Planungssysteme, in: Kleines Lexikon der Informatik und Wirtschaftsinformatik (Hrsg. Szilahi-Szabo, M. G.), München Wien, S. 432-435.
- Qi, M. und Zhang, G. P. (2003) Trend time series Modeling and forecasting with neural networks, in: 2003 Ieee International Conference On Computational Intelligence For Financial Engineering, Proceedings, IEEE: New York, S. 331-337.
- Qian, N. (1999) On the momentum term in gradient descent learning algorithms, in: Neural Networks, Vol. 12, S. 145-151.
- Qiang, S., Gao, X. Z. und Zhuang, X. Y. (2002) Applications of neural networks in servo systems - A survey, in: Soft Computing And Industry - Recent Applications, Springer: Godalming, S. 237-254.
- Qiu, G., Varley, M. R. und Terrell, T. J. (1992) Accelerated Training Of Backpropagation Networks By Using Adaptive Momentum Step, in: Electronics Letters, Vol. 28, S. 377-379.
- Ragg, T., Menzel, W., Baum, W. und Wigbers, M. (2002) Bayesian learning for sales rate prediction for thousands of retailers, in: Neurocomputing, Vol. 43, S. 127-144.
- Raj, A. S., Ravi, R., Parthiban, T. und Radhakrishnan, G. (1999) Artificial neural network applications in electrochemistry - A review, in: Bulletin Of Electrochemistry, Vol. 15, S. 552-555.
- Rall, W. (1990) Some Historical Notes, in: Computational Neuroscience (Hrsg. Schwartz, E. L.), MIT Press: Cambridge, Mass., S. 1-7.
- Ramasubramanian, P. und Karman, A. (2004) Quickprop neural network short-term forecasting framework for a database intrusion prediction system, in: Artificial Intelligence And Soft Computing - Icaisc 2004, Vol. 3070, Springer-Verlag Berlin: Berlin, S. 847-852.
- Ramberg, J. S., Dudewicz, E. J., Tadikamalla, P. R. und Mykytka, E. F. (1979) Probability Distribution and Its Uses in Fitting Data, in: Technometrics, Vol. 21, S. 201-214.
- Ramsey, J. B. (1969) Tests For Specification Errors In Classical Linear Least-Squares Regression Analysis, in: Journal Of The Royal Statistical Society Series B-Statistical Methodology, Vol. 31, S. 350-&.

- Ramsey, W. M., Stich, S. P. und Rumelhart, D. E. (1991) *Philosophy and connectionist theory*, L. Erlbaum Associates: Hillsdale, N.J.
- Rangadass, V. (1996) *Nonlinear Time Series Forecasting: A Recurrent Neural Network Approach*, DAI, 58, no. 01B, (1996): 0284
- Reed, R. D. und Marks, R. J. (1999) *Neural smithing: supervised learning in feedforward artificial neural networks*, The MIT Press: Cambridge, Mass.
- Refenes, A. N., Bentz, Y., Bunn, D. W., Burgess, A. N. und Zapranis, A. D. (1997) *Financial time series modelling with discounted least squares backpropagation*, in: *Neurocomputing*, Vol. 14, S. 123-138.
- Rehkugler, H. und Poddig, T. (1991) *Artificial Neural Networks In Financial Analysis - A New Approach To Forecast Movements Of Stocks, Indexes And Currencies*, in: *Wirtschaftsinformatik*, Vol. 33, S. 365-374.
- Reid, D. J. (1969) *A comparative study of time series prediction techniques on economic data*, unpublished PhD thesis, University of Nottingham: Nottingham, UK
- Reid, D. J. (1972) *A comparison of forecasting techniques on economic time series*, in: *Forecasting in Action* (Hrsg. Bramson, M. J., Helps, I. G. und Watson-Grady, J. A. C. C.), Operations research Society: Birmingham, UK.
- Reilly, D. (2000) *The AUTOBOX system*, in: *International Journal Of Forecasting*, Vol. 16, S. 531-533.
- Remus, W. und O'Connor, M. (2001) *Neural networks for time-series forecasting*, in: *Principles of forecasting: a handbook for researchers and practitioners* (Hrsg. Armstrong, J. S.), Kluwer Academic: Boston; London, S. 245-258.
- Resh, M., Friedman, M. und Barbosa, L. C. (1976) *General Solution of Deterministic Lot Size Problem with Time-Proportional Demand*, in: *Operations Research*, Vol. 24, S. 718-725.
- Rezaie, M. G., Farbiz, F., Moghaddam, E. Z. und Hooshmand, A. (2004) *A novel approach for implementation of adaptive learning rate neural networks*, in: *22nd Norchip Conference, Proceedings*, Ieee: New York, S. 79-82.
- Rhee, H. S. und Oh, K. W. (1996) *A design and analysis of objective function-based unsupervised neural networks for fuzzy clustering*, in: *Neural Processing Letters*, Vol. 4, S. 83-95.
- Riedmiller, M. und Braun, H. (1993) *A Direct Adaptive Method For Faster Backpropagation Learning - The Rprop Algorithm*, in: *1993 IEEE International Conference On Neural Networks*, Vols 1-3, IEEE: New York, S. 586-591.
- Ritchken, P. H. und Sankar, R. (1984) *The Effect of Estimation Risk in Establishing Safety Stock Levels in an Inventory Model*, in: *Journal of the Operational Research Society*, Vol. 35, S. 1091-1099.
- Rittinghaus-Meyer, D. (1993) *Die Anwendung von neuronalen Netzen in der Marketingforschung*, Akademischer Verlag: München.
- Rizvi, S. A. und Nasrabadi, N. M. (1999) *Neural networks for image coding: A survey*, in: *Applications Of Artificial Neural Networks In Image Processing Iv*, Vol. 3647, *Spie - Int Soc Optical Engineering*: Bellingham, S. 46-57.
- Robert, C., Guilpin, C. und Limoge, A. (1998) *Review of neural network applications in sleep research*, in: *Journal Of Neuroscience Methods*, Vol. 79, S. 187-193.
- Robrade, A. D. (1991) *Dynamische Einprodukt-Lagerhaltungsmodelle bei periodischer Bestandsüberwachung*, Physica: Heidelberg.
- Robrade, A. D. und Zoller, K. (1987) *Dynamische Bestellmengen- und Losgrößenplanung*, Institut für Betriebliche Logistik und Organisation: Hamburg.

- Rojas, R. (1993) Theorie der neuronalen Netze: eine systematische Einführung, Springer-Verlag: Berlin; New York.
- Rosenblatt, F. (1958) The perceptron, a theory of statistical separability in cognitive systems. (Project PARA), United States Department of Commerce Office of Technical Services]: [Washington.
- Rosentreter, J. (1977) Prognosen in der industriellen Planung: eine Analyse von Prognoseverfahren und Anwendungsproblemen unter besonderer Berücksichtigung von Verfahren auf der Basis der exponentiellen Glättung, Deutsch: Zürich [u.a.].
- Rumelhart, D. E., Hinton, G. E. und Williams, R. J. (1994) Learning representations by back-propagating errors (from Nature 1986), in: Spie Milestone Series Ms, Vol. 96, S. 138.
- Rumelhart, D. E., McClelland, J. L. und University of California San Diego. PDP Research Group. (1986) Parallel distributed processing: explorations in the microstructure of cognition, MIT Press: Cambridge, Mass.
- Rumelhart, D. E. und Zipser, D. (1990) The Neurobiological Significance of new Learning Models, in: Computational neuroscience (Hrsg. Schwartz, E. L.), MIT Press: Cambridge, Mass., S. 192-200.
- Rygielski, C., Wang, J.-C. und Yen, D. C. (2002) Data mining techniques for customer relationship management, in: Technology in Society, Vol. 24, S. 483-502.
- Saad, E. W., Prokhorov, D. V. und Wunsch, D. C. (1998) Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks, in: Ieee Transactions On Neural Networks, Vol. 9, S. 1456-1470.
- Sabuncuoglu, I. (1998) Scheduling with neural networks: a review of the literature and new research directions, in: Production Planning & Control, Vol. 9, S. 2-12.
- Said, S. E. und Dickey, D. A. (1984) Testing For Unit Roots In Autoregressive-Moving Average Models Of Unknown Order, in: Biometrika, Vol. 71, S. 599-607.
- Saito, K. und Nakano, R. (2000) Second-order learning algorithm with squared penalty term, in: Neural Computation, Vol. 12, S. 709-729.
- Salzano, M. (1999) Neural networks for economic forecasting, in: Neural Nets - Wirn Vietri-99, Springer-Verlag, London Ltd: Godalming, S. 391-407.
- Sarle, W. S. (1994) Neural Networks and Statistical Models, in: Proceedings of the Proceedings of the Nineteenth Annual SAS Users Group International Conference, Cary, NC, Vol., S. 1538-1550.
- Schafer, A. M. und Zimmermann, H. G. (2006) Recurrent neural networks are universal approximators, in: Artificial Neural Networks - Icann 2006, Pt 1, Vol. 4131, Springer-Verlag Berlin: Berlin, S. 632-640.
- Schäfer, G. (1991) Beitrag zur Optimierung von Kommissionierfahrten in Hochregallagergasen mittels neuronaler Netze.
- Scheer, A.-W. (1983) Absatzprognosen, Springer: Berlin [u.a.].
- Scheer, A.-W. (2001) ARIS - Modellierungsmethoden, Metamodelle, Anwendungen, 4. Aufl, Springer: Berlin [u.a.].
- Scheer, A.-W. (2002) ARIS in der Praxis: Gestaltung, Implementierung und Optimierung von Geschäftsprozessen; mit 2 Tabellen, Springer: Berlin [u.a.].
- Schenk, H.-O. (1988) Der Automatenvertrieb: theoretische, empirische und literarische Untersuchungen über eine "vergessene" Betriebsform des Einzelhandels: Duisburg.
- Schierholt, K. und Dagli, C. H. (1996) Stock market prediction using different neural network classification architectures, in: Proceedings Of The Ieee/Iafe 1996 Conference On Computational Intelligence For Financial Engineering (Cifer), I E E E: New York, S. 72-78.
- Schlittgen, R. (1997a) Fitting of threshold models for time series.

- Schlittgen, R. (1997b) Fitting of threshold models for time series, Technical Report, Nr. 97, 13, Discussion paper of the Department of Economics, USC, University of California, 17 S. Seiten.
- Schlittgen, R. und Streitberg, B. (1987) *Zeitreihenanalyse*, 2., erw. Aufl, Oldenbourg: München [u.a.].
- Schmeiser, B. W. und Deutsch, S. J. (1977) Quantile Estimation from Grouped Data - Cell Midpoint, in: *Communications in Statistics Part B-Simulation and Computation*, Vol. 6, S. 221-234.
- Schnaars, S. P. (1986) A comparison of extrapolation models on yearly sales forecasts, in: *International Journal of Forecasting*, Vol. 2, S. 71-85.
- Schneeweiß, C. (1981) *Modellierung industrieller Lagerhaltungssysteme: Einführung und Fallstudien*, Springer: Berlin [u.a.].
- Schneeweiß, H. (1971) *Ökonometrie*, Physica-Verlag: Würzburg.
- Schöneburg (1990) Stock price prediction using neural networks: A project report, in: *Neurocomputing*, Vol. 2, S. 17-27.
- Schönfeld, P. (1969) *Methoden der Ökonometrie*, Vahlen: Berlin u.a.
- Schreiber, G. R. (1961) *A concise history of vending in the U.S.A*, Vend: Chicago.
- Schreiber, G. R. (1999) *A concise history of vending in the U.S.A*, Millennium, Sunrise Books: Lake Bluff, Ill.
- Schröder, M. (1978) Einführung in die kurzfristige Zeitreihenprognose und Vergleich der einzelnen Verfahren, in: *Prognoserechnung* (Hrsg. Mertens, P. und Bruckmann, G.), Physica-Verl.: Würzburg [u.a.], S. 21-60.
- Schulte, C. (1991) *Logistik: Wege zur Optimierung des Material- und Informationsflusses*, Vahlen: München.
- Schwartz, E. L. (1990a) *Computational neuroscience*, MIT Press: Cambridge, Mass.
- Schwartz, E. L. (1990b) Introduction to Computational Neuroscience, in: *Computational Neuroscience* (Hrsg. Schwartz, E. L.), MIT Press: Cambridge, Mass., S. I-XVII.
- Schwarze, J. (1973) Probleme der Fehlermessung bei quantitativen ökonomischen Prognosen, in: *Zeitschrift für die gesamte Staatswissenschaft*, Vol. 129 Verl. Neue Wirtschafts-Briefe: Herne [u.a.], S. 535-558.
- Schwarze, J. (1980) Statistische Kenngrößen zur Ex-post-Beurteilung von Prognosen (Prognosefehlermaße), in: *Angewandte Prognoseverfahren* (Hrsg. Schwarze, J. und Backert, K.), Verl. Neue Wirtschafts-Briefe: Herne [u.a.], S. 317-344.
- Schwarze, J. und Backert, K. (1980) *Angewandte Prognoseverfahren*, Verl. Neue Wirtschafts-Briefe: Herne [u.a.].
- Schwenker, F., Dietrich, C., Kestler, H. A., Riede, K. und Palm, G. (2003) Radial basis function neural networks and temporal fusion for the classification of bioacoustic time series, in: *Neurocomputing*, Vol. 51, S. 265-275.
- Seghouane, A. K. und Fleury, G. (2001) A cost function for learning feedforward neural networks subject to noisy inputs, in: *Isspa 2001: Sixth International Symposium On Signal Processing And Its Applications*, Vols 1 And 2, Proceedings, I E E E: New York, S. 386-389.
- Seifert, D. (2001) *Efficient Consumer Response: Supply Chain Management (SCM), Category Management (CM) und Collaborative Planning, Forecasting and Replenishment (CPFR) als neue Strategieansätze*, 2., erw. Aufl, Hampp: München [u.a.].
- Seifert, D. (2004) *Efficient Consumer Response: Supply Chain Management (SCM), Category Management (CM) und Collaborative Planning, Forecasting and Replenishment (CPFR) als neue Strategieansätze*, 3., erw. Aufl, Hampp: München [u.a.].

- Sejnowski, T. und Rosenberg, C. R. (1988) NETTalk: a parallel network that learns to read aloud. Technical Report JHU/EECS-86/01, John Hopkins University, 1986, in: Neuro-computing: foundations of research, Vol. 1 (Hrsg. Anderson, J. A. und Rosenfeld, E.), MIT Press: Cambridge, Mass., S. 117-122.
- Selmiae, R. und Lewis, F. L. (2000) Neural Network Approximation of piecewise continuous functions, in: Soft computing and intelligent systems: theory and applications (Hrsg. Sinha, N. K. und Gupta, M. M.), Academic Press: San Diego, S. xvii, 639 p.
- Sfetsos, A. und Siriopoulos, C. (2004) Combinatorial time series forecasting based on clustering algorithms and neural networks, in: Neural Computing & Applications, Vol. 13, S. 56-64.
- Shadbolt, J. und Taylor, J. G. (2002) Neural networks and the financial markets: predicting, combining, and portfolio optimisation, Springer: New York.
- Shaheed, M. H. (2005) Feedforward neural network based non-linear dynamic modelling of a TRMS using RPROP algorithm, in: Aircraft Engineering And Aerospace Technology, Vol. 77, S. 13-22.
- Shanmugasundaram, J., Prasad, M. V. N., Vadhavkar, S. und Gupta, A. (2002) Use of Recurrent Neural Networks for Strategic Data Mining of Sales, Technical Report, Nr. 4347-02, MIT Sloan Working Paper, Massachusetts Institute of Technology (MIT): Boston, 18 Seiten.
- Shapiro, S. S. und Wilk, M. B. (1965) An Analysis Of Variance Test For Normality (Complete Samples), in: Biometrika, Vol. 52, S. 591-&.
- Sharda, R. (1994) Neural Networks For The MS/OR Analyst - An Application Bibliography, in: Interfaces, Vol. 24, S. 116-130.
- Sharda, R. und Patil, R. B. (1990) Neural Networks As Forecasting Experts - An Empirical-Test, in: IJCNN'90 International Joint Conference On Neural Networks, Theory Track, Neural And Cognitive Sciences Track, Vol. 1 (Hrsg. IEEE), Lawrence Erlbaum Assoc Publ: Hillsdale, S. B491-B494.
- Sharda, R. und Patil, R. B. (1992a) Connectionist Approach To Time-Series Prediction - An Empirical-Test, in: Journal Of Intelligent Manufacturing, Vol. 3, S. 317-323.
- Sharda, R. und Patil, R. B. (1992b) Connectionist Approach To Time-Series Prediction - An Empirical Test, in: Journal Of Intelligent Manufacturing, Vol. 3, S. 317-323.
- Sharkey, A. J. C. (1999) Combining artificial neural nets: ensemble and modular multi-net systems, Springer: London; New York.
- Sharma, V. (2002) Simulation of sub-surface tile drain water quality (nitrates) using artificial neural networks, in: MAI, Vol. 40, S. 1265.
- Sherbrooke, C. C. (1992) Optimal inventory modeling of systems: multi-echelon techniques, Wiley: New York [u.a.].
- Sherbrooke, C. C. (2004) Optimal inventory modeling of systems: multi-echelon techniques, 2. ed., Kluwer Academic Publ.: Boston [u.a.].
- Shishkin, J. und Young, A. H. (1967) The X-11 variant of the Census II method seasonal adjustment program, Technical Report, Nr. 15, Technical Paper No. 15, Bureau of the Census.
- Shore, H. (1999) Optimal solutions for stochastic inventory models when the lead-time demand distribution is partially specified, in: International Journal of Production Economics, Vol. 59, S. 477-485.
- Sick, B. (2002) On-line and indirect tool wear monitoring in turning with artificial neural networks: A review of more than a decade of research, in: Mechanical Systems And Signal Processing, Vol. 16, S. 487-546.
- Silva, F. M. und Almeida, L. B. (1990) Acceleration Techniques For The Backpropagation Algorithm, in: Lecture Notes In Computer Science, Vol. 412, S. 110-119.

- Silver, E. (1978) Inventory Control under a Probabilistic Time-Varying, Demand Pattern, in: *AIIE Transactions*, Vol. 10, S. 371-379.
- Silver, E. A. (1980) Comments on the lognormal approximation of lead time demand, in: *Omega*, Vol. 8, S. 701.
- Silver, E. A. (1981a) Establishing Reorder Points in the (S,C,S) Coordinated Control-System under Compound Poisson Demand, in: *International Journal of Production Research*, Vol. 19, S. 743-750.
- Silver, E. A. (1981b) Operations-Research in Inventory Management - a Review and Critique, in: *Operations Research*, Vol. 29, S. 628-645.
- Silver, E. A. und Peterson, R. (1985) *Decision systems for inventory management and production planning*, 2., Wiley: New York [u.a.].
- Silver, E. A., Pyke, D. F. und Peterson, R. (1998) *Inventory management and production planning and scheduling*, 3. ed., Wiley: New York, NY [u.a.].
- Smith, D. E. (1974) Adaptive Response for Exponential Smoothing, in: *Operational Research Quarterly*, Vol. 25, S. 421-448.
- Smith, K. A. (1996) Solving Combinatorial Optimisation Problems Using Neural Networks (Traveling Salesman Problem), in: *DAI*, 57, no. 10B, (1996): 6547.
- Smith, K. A. (1999) Neural networks for combinatorial optimization: A review of more than a decade of research, in: *Inform's Journal On Computing*, Vol. 11, S. 15-34.
- Smith, K. A. und Gupta, J. N. D. (2000) Neural networks in business: techniques and applications for the operations researcher, in: *Computers & Operations Research*, Vol. 27, S. 1023-1044.
- Smith, K. A., Gupta, J. N. D. und NetLibrary Inc. (2002a) *Neural networks in business techniques and applications*, Idea Group Pub.: Hershey, PA, S. v, 258 p.
- Smith, K. A., Potvin, J. Y. und Kwok, T. (2002b) Neural network models for combinatorial optimization: A survey of deterministic, stochastic and chaotic approaches, in: *Control And Cybernetics*, Vol. 31, S. 183-216.
- Smola, A. J. und Schölkopf, B. (2004) A tutorial on support vector regression, in: *Statistics and Computing*, Vol. 14, S. 199-222.
- Somervuo, P. und Kohonen, T. (1999) Self-organizing maps and learning vector quantization for feature sequences, in: *Neural Processing Letters*, Vol. 10, S. 151-159.
- Song, J. S. und Zipkin, P. (1993) Inventory Control in a Fluctuating Demand Environment, in: *Operations Research*, Vol. 41, S. 351-370.
- Souza, B. A., Brito, N. S. D., Neves, W. L. A., Silva, K. M., Lima, R. B. V. und da Silva, S. S. B. (2004) Comparison between backpropagation and RPROP algorithms applied to fault classification in transmission lines, in: *2004 Ieee International Joint Conference On Neural Networks*, Vols 1-4, Proceedings, Ieee: New York, S. 2913-2918.
- Sox, C. R. (1997) Dynamic lot sizing with random demand and non-stationary costs, in: *Operations Research Letters*, Vol. 20, S. 155-164.
- Stahlbock, R. (2002) *Evolutionäre Entwicklung künstlicher neuronaler Netze zur Lösung betriebswirtschaftlicher Klassifikationsprobleme*, Verlag für Wissenschaft und Kultur (WiKu-Verlag) Dr.Stein: Berlin.
- Stahlbock, R. und Crone, S. F. (2005) Evolutionary neural classification for evaluation of retail stores and decision support, in: *Proceedings Of The International Joint Conference On Neural Networks (Ijcnnc)*, Vols 1-5, Ieee: New York, S. 1499-1504.
- Stahlbock, R., Lessmann, S. und Crone, S. F. (2005) Evolutionary neural classification approaches for strategic and operational decision support in retail store planning, in: *Icai '05: Proceedings Of The 2005 International Conference On Artificial Intelligence*, Vols 1 And 2, C S R E A Press: Athens, S. 60-66.

- Starr, M. K. und Miller, D. W. (1962) Inventory Control: theory and practice, Prentice-Hall: Englewood Cliffs, N.J.
- Statistisches Bundesamt Deutschland (2006a) Finanzen und Steuern. Fachserie 14, Reihe 8, Umsatzsteuer 2004, Steuerpflichtige und deren Lieferungen und Leistungen in ausgewählten Wirtschaftsbereichen; Einzelhandel (ohne Handel mit Kfz und Tankstellen), Technical Report, (Hrsg. Deutschland, S. B.): Stuttgart.
- Statistisches Bundesamt Deutschland (2006b) Umsatz Deutschland Einzelhandel, Vol. 2006 (Hrsg. Deutschland, S. B.) Statistisches Bundesamt Deutschland: Online im Internet, S. Umsatzzahlen des Einzelhandels in Deutschland, exkl. Kraftfahrzeuge und ohne Tankstellen (in Preisen von 2000).
- Steinberger, T. und Zinner, L. (1999) Forecasting under general Loss Function, Working Paper (nicht publiziert), 10
- Stelling, J. N. und Göllnitz, R. (2000) Direkte Produkt-Rentabilität (DPR): ein umstrittener Kostenrechnungsansatz im Konsumgüterbereich, Fh: Mittweida.
- Stickel, E. (1997) Gabler Wirtschaftsinformatik-Lexikon, Gabler: Wiesbaden.
- Stier, W. (2001) Methoden der Zeitreihenanalyse, Springer: Berlin [u.a].
- Stinchcombe, M. und White, H. (1990) Approximating And Learning Unknown Mappings Using Multilayer Feedforward Networks With Bounded Weights, in: Ijcn International Joint Conference On Neural Networks, Vols 1-3, I E E E: New York, S. C7-C16.
- Stölzle, W. und Helm, R. (2005) Mit Advanced Optimal Shelf Availability zu einer verbesserten Verfügbarkeit in den Regalen, Lebensmittelzeitung, Ed. 47, 07.10.2005, S. 81.
- Straube, F., Pfohl, H.-C., Günthner, W. A., Dangelmaier, W. und Bundesvereinigung Logistik (2005) Trends und Strategien in der Logistik: ein Blick auf die Agenda des Logistik-Managements 2010, Dt. Verkehrs-Verl.: Hamburg.
- Strecker, S. (1997) Künstliche Neuronale Netze - Aufbau und Funktionsweise, Technical Report, Nr. Arbeitspapiere WI Nr. 10/1997, Lehrstuhl für allg. BWL und Wirtschaftsinformatik: Mainz, 1-36 Seiten.
- Strijbosch, L. W. G. und Moors, J. J. A. (2005) The impact of unknown demand parameters on (R,S)-inventory control performance, in: European Journal of Operational Research, Vol. 162, S. 805-815.
- Strijbosch, L. W. G. und Moors, J. J. A. (2006) Modified normal demand distributions in (R, S)-inventory control, in: European Journal of Operational Research, Vol. 172, S. 201-212.
- Stührenberg, L. und Töpken, M. (1996) Prognostik im Spannungsfeld von Kreativität und Systematik: Integration der Szenario-Technik in ein quantitatives Prognosemodell am Beispiel der Bevölkerungsentwicklung in der Bundesrepublik Deutschland, Lang: Frankfurt am Main [u.a.].
- Suarez-Romero, J. A., Fontenla-Romero, O., Guijarro-Berdinas, B. und Alonso-Betanzos, A. (2003) A new learning method for single layer neural networks based on a regularized cost function, in: Computational Methods in Neural Modeling, Pt 1, Vol. 2686, SPRINGER-VERLAG BERLIN: Berlin, S. 270-277.
- Sudjianto, A. (1996) Artificial neural networks for nonlinear extensions of principal component analysis
- Südkamp, A. (1995) Einsatzmöglichkeiten quantitativer Prognoseverfahren im Rahmen der Betriebsergebnisplanung in der Eisen- und Stahlindustrie, Eul: Bergisch Gladbach [u.a.].
- Sutton, R. S. (1988) Learning to predict by the method of temporal differences, in: Machine Learning, Vol. 3, S. 9-44.

- Sutton, R. S. (1992) Adapting Bias By Gradient Descent - An Incremental Version Of Delta-Bar-Delta, in: Aaai-92 Proceedings: Tenth National Conference On Artificial Intelligence, Amer: Menlo, PK, S. 171-176.
- Suykens, J. A. K. und Vandewalle, J. (2000) Recurrent least squares support vector machines, in: Ieee Transactions on Circuits and Systems I-Fundamental Theory and Applications, Vol. 47, S. 1109-1114.
- Suykens, J. A. K., Vandewalle, J. und Moor, B. L. R. d. (1996) Artificial neural networks for modelling and control of non-linear systems, Kluwer Academic Publishers: Boston.
- Suzuki, Y. (2000) Soft computing in industrial applications, Springer: London; New York.
- Sztandera, L. A., Frank, C., Garg, A. und Raheja, A. (2003) Statistical and neural network forecasts of apparel sales, in: Intelligent Information Processing And Web Mining, Physica-Verlag Gmbh & Co: Heidelberg, S. 321-329.
- Tadikamalla, P. R. (1978a) Applications of Weibull Distribution in Inventory Control, in: Journal of the Operational Research Society, Vol. 29, S. 77-83.
- Tadikamalla, P. R. (1978b) Eoq Inventory Model for Items with Gamma Distributed Deterioration, in: Aiie Transactions, Vol. 10, S. 100-103.
- Tadikamalla, P. R. (1979) The lognormal approximation to the lead time demand in inventory control, in: Omega, Vol. 7, S. 553.
- Tadikamalla, P. R. (1984) A Comparison of Several Approximations to the Lead Time Demand Distribution, in: Omega-International Journal of Management Science, Vol. 12, S. 575-581.
- Tai, P., Ryaciotakiboussalis, H. A. und Hollaway, D. (1991) Neural Network Implementation To Control-Systems - A Survey Of Algorithms And Techniques, in: Conference Record Of The Twenty-Fifth Asilomar Conference On Signals, Systems & Computers, Vols 1 And 2, IEEE, Computer Soc.Press: Los Alamitos, S. 1123-1127.
- Tai, P., Ryaciotakiboussalis, H. A. und Tai, K. (1990) The Application Of Neural Networks To Control-Systems - A Survey, in: Twenty-Fourth Asilomar Conference On Signals, Systems & Computers, Vols 1 And 2, Maple Press: San Jose, S. 148-151.
- Takeuchi, I., Bengio, Y. und Kanamori, T. (2002) Robust regression with asymmetric heavy-tail noise distributions, in: Neural Computation, Vol. 14, S. 2469-2496.
- Tam, K. Y. und Kiang, M. Y. (1992) Managerial Applications Of Neural Networks - The Case Of Bank Failure Predictions, in: Management Science, Vol. 38, S. 926-947.
- Tang, O. und Teunter, R. (2006) Economic lot scheduling problem with returns, in: Production and Operations Management, Vol. 15, S. 488-497.
- Tang, X., Zhou, Z. und Shi, Y. (2003) The variable weighted functions of combined forecasting*1, in: Computers & Mathematics with Applications, Vol. 45, S. 723-730.
- Tang, Z. Y., De Almeida, C. und Fishwick, P. A. (1991) Time-Series Forecasting Using Neural Networks vs Box-Jenkins Methodology, in: Simulation, Vol. 57, S. 303-310.
- Tang, Z. Y. und Fishwick, P. A. (1993a) Feed-forward Neural Nets as Models for Time Series Forecasting, Technical Report, Nr. TR-91008, Computer and Information Science, University of Florida: Gainesville.
- Tang, Z. Y. und Fishwick, P. A. (1993b) Feed-forward Neural Nets as Models for Time Series Forecasting, in: ORSA Journal on Computing, Vol. 5, S. 374-386.
- Tashman, L. J. (2000) Out-of-sample tests of forecasting accuracy: an analysis and review, in: International Journal of Forecasting, Vol. 16, S. 437-450.
- Tay, F. E. H. und Cao, L. (2001) Application of support vector machines in financial time series forecasting, in: Omega, Vol. 29, S. 309-317.

- Taylor, J. W. (2000) A quantile regression neural network approach to estimating the conditional density of multiperiod returns, in: *Journal Of Forecasting*, Vol. 19, S. 299-311.
- Tchaban, T., Griffin, J. P. und Taylor, M. J. (1997) A comparison between single and combined backpropagation neural networks in the prediction of turnover, in: *Proceedings of the Kes '97, First International Conference On Knowledge-Based Intelligent Electronic Systems*, New York, Vol. 2 IEEE, S. 347-354.
- Tchaban, T., Griffin, J. P. und Taylor, M. J. (1998) A comparison between single and combined backpropagation neural networks in the prediction of turnover, in: *Engineering Applications Of Artificial Intelligence*, Vol. 11, S. 41-47.
- Telagi, M. S. (1996) Survey of control systems in manufacturing field and implementation of multilayer back propagation neural [i.e. neural] network for robot manipulation
- Tempelmeier, H. (1992) *Material-Logistik: Grundlagen der Bedarfs- und Losgrößenplanung in PPS-Systemen*, 2., neubearb. und erw. Aufl, Springer: Berlin [u.a.].
- Tempelmeier, H. (2003) *Material-Logistik: Modelle und Algorithmen für die Produktionsplanung und -steuerung und das Supply Chain Management; mit 148 Tabellen*, 5., neubearb. Aufl, Springer: Berlin [u.a.].
- Teräsvirta, T. und Lukkonen, R. (1983) Choosing between linear and threshold autoregressive models, *Elinkeinoelämän tutkimuslaitos: Helsinki, Finland*.
- Terui, N. und van Dijk, H. K. (2002) Combined forecasts from linear and nonlinear time series models, in: *International Journal of Forecasting*, Vol. 18, S. 421-438.
- Theera-Umpon, N. und Gader, P. D. (2000) Training neural networks to count white blood cells via a minimum counting error objective function, in: *15th International Conference On Pattern Recognition*, Vol 2, *Proceedings - Pattern Recognition And Neural Networks*, Ieee Computer Soc: Los Alamitos, S. 299-302.
- Theil, H. (1966) *Applied economic forecasting*, Elsevier: Amsterdam.
- Thibault, J., Vanbreusegem, V. und Cheruy, A. (1990) Online Prediction Of Fermentation Variables Using Neural Networks, in: *Biotechnology And Bioengineering*, Vol. 36, S. 1041-1048.
- Thiesing, F. M. (1998) *Analyse und Prognose von Zeitreihen mit neuronalen Netzen*, Shaker: Aachen.
- Thiesing, F. M., Middelberg, U. und Vornberger, O. (1995a) Abverkaufsprognose mit paralleler Backpropagation, in: *Proceedings of the Fuzzy-Neuro-Systeme '95*, TH Darmstadt, Vol., S. 373-380.
- Thiesing, F. M., Middelberg, U. und Vornberger, O. (1995b) A Neural Network Approach for Predicting the Sale of Articles in Supermarkets, in: *Proceedings of the 3rd European Congress on Intelligent Techniques and Soft Computing*, Aachen, Germany, Vol. 1, S. 315-320.
- Thiesing, F. M., Middelberg, U. und Vornberger, O. (1995c) Parallel Back-Propagation for Sales Prediction on Transputer Systems, in: *Proceedings of the Proceedings World Transputer Congress '95*, Harrogate, UK, Vol., S. 318-331.
- Thiesing, F. M., Middelberg, U. und Vornberger, O. (1995d) Short Term Prediction of Sales in Supermarkets, in: *Proceedings of the ICNN'95*, Perth, WA, Vol. 2, S. 1028-1031.
- Thiesing, F. M. und Vornberger, O. (1995) Time Series Prediction by Neural Networks, in: *Power Explorer User Report*, Parsytec GmbH, S. 33-36.
- Thiesing, F. M. und Vornberger, O. (1997a) Forecasting sales using neural networks, in: *Computational Intelligence - Theory And Applications*, Vol. 1226, Springer-Verlag Berlin: Berlin 33, S. 321-328.

- Thiesing, F. M. und Vornberger, O. (1997b) Forecasting Sales Using Neural Networks Computational Intelligence, in: Proceedings of the 5th International Conference Fuzzy Days, Dortmund, Germany, Vol. Springer LNCS 1226, S. 321-328.
- Thiesing, F. M. und Vornberger, O. (1997c) Sales forecasting using neural networks, in: Proceedings of the 1997 IEEE International Conference On Neural Networks, New York, Vol. 1 IEEE, S. 2125-2128.
- Thiesing, F. M. und Vornberger, O. (1998) Abverkaufsprognose im Supermarkt mit Neuronalen Netzen., in: Betriebswirtschaftliche Anwendungen des Soft Computing (Hrsg. Biethahn, J., Hönerloh, A., Kuhl, J., Leisewitz, M. C., Nissen, V. und Tietze, M.), Vieweg, S. 115-126.
- Thomae, D. A. und Vandembout, D. E. (1990) Encoding Logical Constraints Into Neural Network Cost-Functions, in: Ijcn International Joint Conference On Neural Networks, Vols 1-3, I E E E: New York, S. C863-C868.
- Thomassey, S., Happiette, M. und Castelain, J.-M. (2005) A global forecasting support system adapted to textile distribution, in: International Journal of Production Economics, Vol. 96, S. 81.
- Thomassey, S., Happiette, M. und Castelain, J. M. (2002) Three complementary sales forecasting models for textile distributors, in: Proceedings of the 2002 IEEE International Conference on Systems, Man and Cybernetics, Vol. 6, S. 5 pp. vol.6.
- Tietz, B. (1992) Dynamik im Handel, Dt. Fachverl.: Frankfurt am Main.
- Tietz, B. (1993a) Marketing, 3., überarb. und erw. Aufl, Werner: Düsseldorf.
- Tietz, B. (1993b) Zukunftsstrategien für Handelsunternehmen, Dt. Fachverl.: Frankfurt am Main.
- Tietz, B. (1994) Zukunft im Handel, Dt. Fachverl.: Frankfurt am Main.
- Tijms, H. C. (1994) Stochastic models: an algorithmic approach, Wiley: Chichester [u.a.].
- Tijms, H. C. und Groenevelt, H. (1984) Simple Approximations for the Reorder Point in Periodic and Continuous Review (S,S) Inventory Systems with Service Level Constraints, in: European Journal of Operational Research, Vol. 17, S. 175-190.
- Timothy, H. L. und Sung-Chang, J. (1999) Forecasting creditworthiness: Logistic vs. artificial neural net, in: The Journal of Business Forecasting Methods & Systems, Vol. 18, S. 28.
- Tollenaere, T. (1990) Supersab - Fast Adaptive Back Propagation With Good Scaling Properties, in: Neural Networks, Vol. 3, S. 561-573.
- Tong, H. (1983) Threshold models in non-linear time series analysis, Springer-Verlag: New York.
- Tong, H. (1990) Non-linear time series: a dynamical system approach, Oxford University Press: Oxford.
- Toporowski, W. (1996) Logistik im Handel: optimale Lagerstruktur und Bestellpolitik einer Filialunternehmung, Physica-Verl.: Heidelberg.
- Toporowski, W. (1998) Grundlagen der Bestellpunkt- und Bestellzyklusverfahren, in: WISU, Vol. 27, S. 1142-1154.
- Toporowski, W. (1999a) Bestellmengenpolitiken bei stochastischer, stationärer Nachfrage (I), in: WISU, Vol. 28, S. 197-204.
- Toporowski, W. (1999b) Bestellmengenpolitiken bei stochastischer, stationärer Nachfrage (II), in: WISU, Vol. 28, S. 325-333.
- Treadgold, N. K. und Gedeon, T. D. (1997) A Cascade network algorithm employing Progressive RPROP, in: Biological And Artificial Computation: From Neuroscience To Technology, Vol. 1240, Springer-Verlag Berlin: Berlin 33, S. 733-742.

- Trident Press International. (2003) The new international Webster's encyclopedic dictionary of the English language, 2003, Trident Press International: Naples, Fla.
- Trigg, D. W. und Leach, A. G. (1967) Exponential Smoothing with an Adaptive Response Rate, in: *Operational Research Quarterly*, Vol. 18, S. 53-59.
- Trux, W. R. (1972) Einkauf und Lagerdisposition mit Datenverarbeitung: Bedarf, Bestand, Bestellung, *Wirtschaftlichkeit*, 2., erw. Aufl, Verl. Moderne Industrie: München.
- Tseng, F.-M., Yu, H.-C. und Tzeng, G.-H. (2001) Combining neural network model with seasonal time series ARIMA model, in: *Technological Forecasting and Social Change*, Vol. 69, S. 71-87.
- Tsuji, T., Fukuda, O., Kaneko, M. und Ito, K. (2000) Pattern classification of time-series EMC signals using neural networks, in: *International Journal Of Adaptive Control And Signal Processing*, Vol. 14, S. 829-848.
- Tukey, J. W. (1960, reprinted in 1998) A survey of sampling from contaminated distributions, in: *The collected works of John W. Tukey* (Hrsg. Tukey, J. W. und Jones, L. V.), Chapman & Hall/CRC: Boca Raton.
- Tyworth, J. E. und O'Neill, L. (1997) Robustness of the normal approximation of lead-time demand in a distribution setting, in: *Naval Research Logistics*, Vol. 44, S. 165-186.
- Ueda, H., Anzai, Y., Ohta, M., Yoneda, S. und Ogihara, A. (1993) Associative Neural Network Models Based On A Measure Of Manhattan Length, in: *Ieice Transactions On Fundamentals Of Electronics Communications And Computer Sciences*, Vol. E76A, S. 277-283.
- Vankayala, V. S. S. und Rao, N. D. (1993) Artificial Neural Networks And Their Applications To Power-Systems - A Bibliographical Survey, in: *Electric Power Systems Research*, Vol. 28, S. 67-79.
- Vapnik, V. N. (1995) *The nature of statistical learning theory*, Springer: New York.
- Vapnik, V. N. (1998) *Statistical learning theory*, Wiley: New York [u.a.].
- Vapnik, V. N. (1999) An Overview of Statistical Learning Theory, in: *IEE Transactions on Neural Networks*, S. 988-999.
- Vapnik, V. N. und Chervonenkis, A. I. A. (1979) *Theorie der Zeichenerkennung*, Akademie-Verlag: Berlin.
- Vargas, V. A. und Metters, R. (1996) Adapting lot-sizing techniques to stochastic demand through production scheduling policy, in: *Iie Transactions*, Vol. 28, S. 141-148.
- Varian, H. R. (1975) A Bayesian approach to real estate assessment, in: *Studies in Bayesian econometrics and statistics in Honor of Leonard J. Savage* (Hrsg. Fienberg, S. E. und Zellner, A.), Amsterdam, S. 195-208.
- Vellido, A., Lisboa, P. J. G. und Vaughan, J. (1999) Neural networks in business: a survey of applications (1992-1998), in: *Expert Systems with Applications*, Vol. 17, S. 51-70.
- Venter, G. und Sobieszczanski-Sobieski, J. (2003) Particle swarm optimization, in: *Aiaa Journal*, Vol. 41, S. 1583-1589.
- Venugopal, V. und Baets, W. (1994) Neural networks & their applications in marketing management, in: *Journal of Systems Management*, Vol. 45, S. 16-21.
- Viaene, S., Dedene, G. und Derrig, R. A. (2005) Auto claim fraud detection using Bayesian learning neural networks, in: *Expert Systems With Applications*, Vol. 29, S. 653-666.
- Vieweg, H.-G. (1999) *Wirtschaftsentwicklung Unterhaltungsautomaten 1999 und Ausblick 2000*, Technical Report, (Hrsg. VDAI), VDAI Jahresberichte, Verband der Deutschen Automatenindustrie: Köln.
- Vieweg, H.-G. (2006) *IFO-Wirtschaftsstudie - Wirtschaftsentwicklung Unterhaltungsautomaten 2006 und Ausblick 2007*, Technical Report, (Hrsg. IFO), IFO Wirtschaftsstudien Branchenforschung, IFO: München, 64 Seiten.

- Vollmer, U. und Strey, A. (1999) Experimental study on the precision requirements of RBF, RPROP and BPTT training, in: Ninth International Conference On Artificial Neural Networks (Icann99), Vols 1 And 2, Inst Electrical Engineers: Stevenage Herts, S. 239-244.
- von Kerssenbrock-Krosigk, D. (1998) Die Mechanische Erbauung - Automaten vor dem Automatenzeitalter, in: Automatenwelten: Freizeitzeugen des Jahrhunderts (Hrsg. Hornbostel, W., Jockel, N. und Museum für Kunst und Gewerbe Hamburg), Prestel: München New York, S. 159 p.
- von Neumann, J. (1943) First draft of a report on the EDVAC, (nicht publiziert)
- von Neumann, J. (1958) The Computer and the Brain, Yale University Press: Yale.
- Voss, S. (1996) Dynamic tabu search strategies for the traveling purchaser problem, in: Annals of Operations Research, Vol. 63, S. 253-275.
- Voß, S. und Gutenschwager, K. (2001) Informationsmanagement, Springer: Berlin [u.a.].
- Vrahatis, M. N., Magoulas, G. D. und Plagianakos, V. P. (2000) Globally convergent modification of the quickprop method, in: Neural Processing Letters, Vol. 12, S. 159-169.
- Wagner, H. und Whitin, T. M. (1958) Dynamic Version of the Economic Lot Size Model, in: Management Science, Vol. 5, S. 89-96.
- Wahl, C. (1999) Bestandsmanagement in Distributionssystemen mit dezentraler Disposition.
- Wang, M.-C. und Subba Rao, S. (1992) Estimating reorder points and other management science applications by bootstrap procedure, in: European Journal of Operational Research, Vol. 56, S. 332.
- Wang, S., Chung, K. F. L., Zhaohong, D. und Dewen, H. (2007) Robust fuzzy clustering neural network based on [var epsilon]-insensitive loss function, in: Applied Soft Computing, Vol. 7, S. 577.
- Wang, X. G., Sun, W. D., Tang, Z., Tamura, H. und Ishii, M. (2003) A RPROP method with Weight Perturbation, in: Sice 2003 Annual Conference, Vols 1-3, Ieee: New York, S. 634-637.
- Wang, X. X. und Brown, D. J. (2003) Incremental neural network construction by using correlation as cost function, in: Design And Application Of Hybrid Intelligent Systems, Vol. 104, I O S Press: Amsterdam, S. 22-31.
- Wang, Y. J. und Lin, C. T. (1998) A second-order learning algorithm for multilayer networks based on block Hessian matrix, in: Neural Networks, Vol. 11, S. 1607-1622.
- Waters, C. D. J. (2003) Inventory control and management, 2. ed, Wiley: Chichester [u.a.].
- Weber, D. und Breitenbach, J. (1997) Neural network objective functions for detection problems, in: Comsig '97 - Proceedings Of The 1997 South African Symposium On Communications And Signal Processing, I E E E: New York, S. 43-46.
- Weber, K. (1991) Wirtschaftsprognostik, in: Die Unternehmung, Vol. 45, S. 65-74.
- Wechsler, W. (1978) Delphi-Methode: Gestaltung und Potential für betriebliche Prognoseprozesse, 1. Aufl, Florentz: München.
- Wedding, I., Donald K. und Cios, K. J. (1996) Time series forecasting by combining RBF networks, certainty factors, and the Box-Jenkins model, in: Neurocomputing, Vol. 10, S. 149-168.
- Wedekind, H. (1968) Ein Vorhersagemodell für sporadische Nachfragemengen bei der Lagerhaltung, in: Ablauf und Planungsforschung, Vol. 1, S. 1-11.
- Wei, H., Lai, K. K., Nakamori, Y. und Wang, S. Y. (2004) Forecasting foreign exchange rates with artificial neural networks: A review, in: International Journal Of Information Technology & Decision Making, Vol. 3, S. 145-165.

- Weigend, A. S. und Bhansali, R. J. (1994) Paradigm Change In Prediction, in: *Philosophical Transactions Of The Royal Society Of London Series A-Mathematical Physical And Engineering Sciences*, Vol. 348, S. 405-420.
- Weigend, A. S., Huberman, B. A. und Rumelhart, D. E. (1990) Predicting the future: a connectionist approach, in: *International Journal of Neural Systems*, Vol. 1, S. 193-209.
- Weigend, A. S. und Nix, A. N. (1994) Predictions with Confidence Intervals (Local Error Bars), in: *Proceedings of the International Conference on Neural Information Processing (ICONIP)*, Seoul, Vol. (Hrsg. Amari, S.) Springer, S. 847-852.
- Weigend, A. S., Rumelhart, D. E. und Huberman, B. A. (1991) Generalization by Weight-Elimination With Application to Forecast, in: *Advances in Neural Information Processing Systems 3* (Hrsg. Lipmann, R. P., Moody, J. E. und Touretzky, D. S.), Morgan Kaufmann: San Mateo, CA, S. 875-882.
- Weigend, A. S. und Srivastava, A. N. (1995) Predicting Conditional-Probability Distributions - a Connectionist Approach, in: *International Journal of Neural Systems*, Vol. 6, S. 109-118.
- Weiss, A. A. (1996) Estimating Time Series Models Using the Relevant Cost Function, in: *Journal of Applied Econometrics*, Vol. 11 John Wiley & Sons, S. 539.
- Werbos, P. (1974) Beyond Regression: new tools for prediction and analysis in the behavioural sciences, unpublished PhD thesis, Harvard University: Massachusetts
- Werbos, P. (1988a) Backpropagation: Past and future, in: *Proceedings of the ICNN'88 - International Conference on Neural Networks*, Vol. 1 (Hrsg. IEEE) IEEE Press, S. 343-353.
- Werbos, P. J. (1988b) Generalization Of Backpropagation With Application To A Recurrent Gas Market Model, in: *Neural Networks*, Vol. 1, S. 339-356.
- Werbos, P. J. (1990) Backpropagation Through Time - What It Does And How To Do It, in: *Proceedings Of The Ieee*, Vol. 78, S. 1550-1560.
- Werbos, P. J. (1994) The roots of backpropagation: from ordered derivatives to neural networks and political forecasting, J. Wiley & Sons: New York.
- Werbos, P. J. (1996) From backpropagation to neurocontrol, in: *Sixth, Seventh, And Eighth Workshops On Virtual Intelligence - Academic/Industrial/Nasa/Defense Technical Interchange And Tutorials*, Vol. 2878, Spie - Int Soc Optical Engineering: Bellingham, S. 476-536.
- Werner, T. (2000) Die Klassifikation von Zeitreihen zur saisonalen Analyse von Absatzzahlen, Eul: Lohmar [u.a.].
- West, D. (2000) Neural network credit scoring models, in: *Computers & Operations Research*, Vol. 27, S. 1131-1152.
- Wheelwright, S. C. und Makridakis, S. G. (1985) *Forecasting methods for management*, 4th, Wiley: New York.
- White, H. (1988) Economic prediction using Neural Networks: The Case of IBM daily stock returns, in: *Proceedings of the Proceedings IEEE International Conference on Neural Networks*, Vol. 2, S. 451-458.
- White, H. (1989a) Learning in artificial neural networks: A statistical perspective, in: *Neural Computation*, Vol. 1, S. 425-464.
- White, H. (1989b) Some Asymptotic Results for Learning in Single Hidden-Layer Feedforward Network Models, in: *Journal of the American Statistical Association*, Vol. 84, S. 1003-1013.
- White, H. (1990) Connectionist Nonparametric Regression - Multilayer Feedforward Networks Can Learn Arbitrary Mappings, in: *Neural Networks*, Vol. 3, S. 535-549.

- White, H. (1992a) Nonparametric Estimation of Conditional Quantiles using Neural Networks, in: Proceedings of the 22nd Symposium on the Interface, Statistics and Many Parameters: Curves, Images, Spatial Models, Vol. (Hrsg. Page, C. und LePage, R.) Springer Verlag: New York.
- White, H. (1992b) Nonparametric Estimation of Conditional Quantiles using Neural Networks, in: Artificial neural networks: approximation and learning theory (Hrsg. White, H., Gallant, A. R., Hornik, K., Stinchcombe, M. und Wooldrich, J.), Blackwell: Oxford [u.a.], S. 2-24.
- White, H. (2006) Approximate Nonlinear Forecasting Methods, in: Handbook of economic forecasting (Hrsg. Elliott, G., Granger, C. W. J. und Timmermann, A.), Elsevier: Boston, Mass., S. 460-512.
- Whitin, T. M. (1953) The theory of inventory management, Princeton University Press: Princeton.
- Whybark, D. C. und Williams, J. G. (1976) Material requirements planning under uncertainty, Vol. 7, S. 595-606.
- Widmann, G. (2001) Künstliche neuronale Netze und ihre Beziehungen zur Statistik.
- Widrow, B. (2003) Persönliches Gespräch im Rahmen der IJCNN 2003, (persönliche Kommunikation mit Crone, S. F.), Portland, Oregon, 22.07.2003
- Widrow, B. und Hoff, M. (1960) Adaptive Switching Circuits, in: Proceedings of the IRE WESCON, Convention Record, Vol. 4, S. 96-104.
- Widrow, B. und Lehr, M. A. (1990) 30 Years Of Adaptive Neural Networks - Perceptron, Madaline, And Backpropagation, in: Proceedings Of The Ieee, Vol. 78, S. 1415-1442.
- Widrow, B., Rumelhart, D. E. und Lehr, M. A. (1994) Neural Networks: Applications in Industry, Business, and Science, in: Communications of the ACM, Vol. 37, S. 93.
- Widrow, B. und Smith, F. W. (1963) Pattern-Recognising Control Systems, in: Proceedings of the Computer and Information Sciences Symposium, Washington DC, Vol. 1 Spartan Books.
- Wiedmann, K.-P. und Buckler, F. (2001) Neuronale Netze im Marketing-Management: praxisorientierte Einführung in modernes data mining, 1. Aufl., Gabler: Wiesbaden.
- Wiener, N. (1948) Cybernetics, J. Wiley: New York.
- Wiener, N. und Schadâe, J. P. (1965) Cybernetics of the nervous system, Elsevier Pub. C.: Amsterdam, New York.
- Wiggins, V. L., Engquist, S. K., Looper, L. T. und Armstrong Laboratory (U.S.) (1992) Neural network applications a literature review, Armstrong Laboratory Air Force Materiel Command: Brooks Air Force Base, Tex.
- Wilbert, R. (1995) Interpretation Neuronaler Netze in den Sozialwissenschaften, in: ZfB, Vol. 65 P. Lang: Frankfurt am Main; New York, S. 769-783.
- Wilcox, G., Poliac, M. und Liebman, M. N. (1989) Prediction Of 3-Dimensional Protein-Structure From Sequence Using Neural Networks (Machine Learning, Neural Networks), in: Abstracts Of Papers Of The American Chemical Society, Vol. 198, S. 13.
- Wild, J. (1974) Grundlagen der Unternehmensplanung, Rowohlt: Reinbek bei Hamburg.
- Wilkinson, J. T. (1993) Explaining agricultural bank failures: a neural network approach
- Wilson, H. R. und Cowan, J. D. (1972) Excitatory And Inhibitory Interactions In Localized Populations Of Model Neurons, in: Biophysical Journal, Vol. 12, S. 1-&.
- Wilson, R. W. (1934) A Scientific Routine fro Stock Control, in: Havard Business Review, Vol. 13, S. 116-128.
- Winters, P. R. (1960) Forecasting Sales by Exponentially Weighted Moving Averages, in: Management Science, Vol. 6, S. 324-342.

- Wiskott, L. und Sejnowski, T. (1997) Objective functions for neural map formation, in: Proceedings Of The 4th Joint Symposium On Neural Computation, Vol 7, University California San Diego: La Jolla, S. 242-248.
- Wöhe, G. und Döring, U. (1990) Einführung in die allgemeine Betriebswirtschaftslehre, 17., überarb. und erw. Aufl, Vahlen: München.
- Wong, B. K., Bodnovich, T. A. und Selvi, Y. (1995) A Bibliography Of Neural-Network Business Applications Research - 1988 To September 1994, in: Expert Systems, Vol. 12, S. 253-262.
- Wong, B. K., Bodnovich, T. A. und Selvi, Y. (1997) Neural network applications in business: A review and analysis of the literature (1988-1995), in: Decision Support Systems, Vol. 19, S. 301-320.
- Wong, B. K., Lai, V. S. und Lam, J. (2000) A bibliography of neural network business applications research: 1994-1998, in: Computers & Operations Research, Vol. 27, S. 1045-1076.
- Wong, B. K. und Selvi, Y. (1998) Neural network applications in finance: A review and analysis of literature (1990-1996), in: Information & Management, Vol. 34, S. 129-139.
- Wu, X., Westervelt, J. D., Construction Engineering Research Laboratory. und United States. Army. Corps of Engineers. (1991) Engineering applications of neural computing, a state-of-the-art survey, US Army Corps of Engineers Construction Engineering Research Laboratory: Springfield, VA.
- Wu, Y.-J. (1998) Exchange rate forecasting: an application of radial basis function neural networks
- Yager, R. R. (1982) Fuzzy Prediction Based On Regression-Models, in: Information Sciences, Vol. 26, S. 45-63.
- Yang, H. Z., Li, D. J., Tao, Z. L. und Zhang, S. Z. (2002) Variable-step BP training algorithm with an adaptive momentum term, in: Proceedings Of The 4th World Congress On Intelligent Control And Automation, Vols 1-4, I E E E: New York, S. 1961-1965.
- Yasdi, R. (2000) A literature survey on applications of neural networks for human-computer interaction, in: Neural Computing & Applications, Vol. 9, S. 245-258.
- Ye, Q., Liang, B. und Li, Y. J. (2004) Amnestic neural network for classification: Application on stock trend prediction, in: 2005 International Conference On Services Systems And Services Management, Vols 1 And 2, Proceedings, Ieee: New York, S. 1031-1034.
- Yip, D. H. F., Hines, E. L. und Yu, W. W. H. (1997) Application of artificial neural networks in sales forecasting, in: 1997 IEEE International Conference On Neural Networks, Vol. 2, IEEE: New York, S. 2121-2124.
- Yohda, M., Saito-Arita, M., Okada, A., Suzuki, R. und Kakemoto, Y. (2002) Demand forecasting by the neural network with discrete Fourier transform, in: 2002 IEEE International Conference On Data Mining, IEEE Computer Society: Los Alamitos, S. 779-782.
- Yu, C. C. und Liu, B. D. (2002) A backpropagation algorithm with adaptive learning rate and momentum coefficient, in: Proceeding Of The 2002 International Joint Conference On Neural Networks, Vols 1-3, I E E E: New York, S. 1218-1223.
- Yu, X. H. und Chen, G. A. (1997) Efficient backpropagation learning using optimal learning rate and momentum, in: Neural Networks, Vol. 10, S. 517-527.
- Zabel, T. (2000) Langfristige Absatzprognosen mit Neuronalen Netzen in der Automobilindustrie und Abverkaufsprognose im Supermarkt, Technical Report, (Hrsg. Lippe, W.), Arbeitspapiere des Fachbereichs Informatik, Universität Münster: Münster, 19 Seiten.
- Zadeh, A. G., Fahim, A. und El-Gindy, M. (1997) Neural network and fuzzy logic applications to vehicle systems: literature survey.

- Zaghw, A. und Dong, W. M. (1994) An Automated Approach For Selecting The Learning Rate And Momentum In Back-Propagation Networks, in: 1994 Ieee International Conference On Neural Networks, Vol 1-7, I E E E: New York, S. 464-469.
- Zekic-Susac, M., Kolaric, D. und Piasevoli, T. (2005) Combining neural networks and on-line analytical processing in sale prediction of footwear and leather goods, in: Proceedings Of The 10th International Conference On Operational Research - Koi 2004, Univ Osijek, Dept Mathematics: Osijek, S. 131-142.
- Zell, A. (2000) Simulation Neuronaler Netze, R. Oldenbourg Verlag: Munich.
- Zellner, A. (1986a) Bayesian-Estimation and Prediction Using Asymmetric Loss Functions, in: Journal of the American Statistical Association, Vol. 81, S. 446-451.
- Zellner, A. (1986b) A tale of forecasting 1001 series: The Bayesian knight strikes again, in: International Journal of Forecasting, Vol. 2, S. 491-494.
- Zhang, G., Patuwo, B. E. und Hu, M. Y. (1998a) Forecasting with artificial neural networks: The state of the art, in: International Journal of Forecasting, Vol. 14, S. 35-62.
- Zhang, G., Y. Hu, M., Eddy Patuwo, B. und C. Indro, D. (1999) Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis, in: European Journal of Operational Research, Vol. 116, S. 16-32.
- Zhang, G. P. (2001) An investigation of neural networks for linear time-series forecasting, in: Computers & Operations Research, Vol. 28, S. 1183-1202.
- Zhang, G. P. (2004) Neural networks in business forecasting, Idea Group: Hershey, Pa.
- Zhang, G. P., Patuwo, B. E. und Hu, M. Y. (1998b) Nonlinear time series forecasting with artificial neural networks, in: Proceedings of the Decision Sciences Institute 1998, Atlanta, Vol. 2 Decision Sciences Inst, S. 1023-1025.
- Zhang, G. P., Patuwo, B. E. und Hu, M. Y. (1998c) Nonlinear time series forecasting with artificial neural networks, in: Decision Sciences Institute 1998 Proceedings, Vols 1-3, Decision Sciences Inst: Atlanta, S. 1023-1025.
- Zhang, G. P., Patuwo, B. E. und Hu, M. Y. (2001) A simulation study of artificial neural networks for nonlinear time-series forecasting, in: Computers & Operations Research, Vol. 28, S. 381-396.
- Zhang, G. P. und Qi, M. (2005) Neural network forecasting for seasonal and trend time series, in: European Journal Of Operational Research, Vol. 160, S. 501-514.
- Zhang, H. C. und Huang, S. H. (1995) Applications Of Neural Networks In Manufacturing - A State-Of-The-Art Survey, in: International Journal Of Production Research, Vol. 33, S. 705-728.
- Zhang, Z. und Friedrich, K. (2003) Artificial neural networks applied to polymer composites: a review, in: Composites Science and Technology, Vol. 63, S. 2029-2044.
- Zhao, B. (1996) A generalized momentum method and its application to artificial neural networks, Shaker: Aachen
- Zhao, L., Collopy, F. und Kennedy, M. (2003) The Problem of Neural Networks in Business Forecasting - An Attempt to Reproduce the Hill, O'Connor and Remus Study, Technical Report, SPROUTS - Working papers on Information Environments, Systems and Organisations, Case Wetsern Reserve University: Cleveland, 11 Seiten.
- Zhou, L. R., Han, P., Jiao, S. M. und Lin, B. H. (2002) Feedforward neural networks using RPROP algorithm and its application in system identification, in: 2002 International Conference On Machine Learning And Cybernetics, Vols 1-4, Proceedings, I E E E: New York, S. 2041-2044.
- Zimmerer, T. (1997) Künstliche neuronale Netze versus ökonomische und zeitreihenanalytische Verfahren zur Prognose ökonomischer Zeitreihen, P. Lang: Frankfurt am Main; New York.

- Zimmermann, H. G. (2006a) Personal conversation between H.-G. Zimmermann and S.F. Crone at ISF06, (persönliche Kommunikation mit Crone, S. F.), Santander, Spain, 22.07.2006
- Zimmermann, H. G. (2006b) System Identification & Forecasting by Neural Networks - Principles, Techniques, Applications, Siemens Corporate Technology: Santander, Spain, S. 76.
- Zimmermann, H. G., Grothmann, R. und Neuneier, R. (2002) Multi-agent FX-market modeling by neural networks.
- Zimmermann, H. G., Grothmann, R., Schafer, A. M. und Tietz, C. (2005) Dynamical consistent recurrent neural networks, in: Proceedings Of The International Joint Conference On Neural Networks (Ijcn), Vols 1-5, Ieee: New York, S. 1537-1541.
- Zimmermann, H. G. und Neuneier, R. (2000a) Combining state space reconstruction and forecasting by neural networks.
- Zimmermann, H. G. und Neuneier, R. (2000b) Modeling dynamical systems by recurrent neural networks, in: Data Mining Ii, Vol. 2, Wit Press: Southampton, S. 557-566.
- Zimmermann, H. G., Neuneier, R. und Grothmann, R. (2001) Multiagent modeling of multiple FX-markets by neural networks, in: Ieee Transactions On Neural Networks, Vol. 12, S. 735-743.
- Zipser und Rumelhart, D. E. (1990) The Neurobiological Significance of new Learning Models, in: Computational Neuroscience (Hrsg. Schwartz, E. L.), MIT Press: Cambridge, Mass., S. 192-200.
- Zorriassatine, F. und Tannock, J. D. T. (1998) A review of neural networks for statistical process control, in: Journal Of Intelligent Manufacturing, Vol. 9, S. 209-224.
- Zou, H. und Yang, Y. (2004) Combining time series models for forecasting, in: International Journal of Forecasting, Vol. 20, S. 69-84.
- Zuperl, U. und Cus, F. (2003) Optimization of cutting conditions during cutting by using neural networks, in: Robotics and Computer-Integrated Manufacturing, Vol. 19, S. 189-199.

8 Anhang

8.1 Tabellen der Zielfunktionswerte der MLPs zur Disposition

Tabelle 8.1. MSE auf Trainings-, Validierungs- und Testdatensatz je Zeitreihe

		MLP. <i>SE</i>			MLP. <i>LLC</i> ₁			MLP. <i>LLC</i> ₂			MLP. <i>LLC</i> ₃		
		Train	Valid	Test	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
MSE	MW stat.	103,6	98,3	107,1	201,9	195,0	197,9	399,6	390,2	388,2	715,2	701,0	694,3
	MW sais.	104,2	106,4	116,8	209,3	210,2	214,2	409,6	410,5	409,8	797,4	787,2	789,8
	MW total	103,9	102,3	111,9	205,6	202,6	206,0	404,6	400,4	399,0	756,3	744,1	742,0
	StAb stat.	213,6	199,3	228,3	419,8	383,4	404,6	828,8	776,5	791,5	1508,9	1433,1	1441,2
	StAb sais.	216,1	213,2	243,8	438,6	420,8	439,9	868,6	839,1	854,6	1686,2	1621,3	1648,9
	StAb total	208,4	200,2	229,2	416,5	390,6	410,1	823,6	784,3	799,1	1552,8	1485,1	1503,1
MAE	MW stat.	5,47	5,46	5,52	8,02	8,11	8,04	12,10	12,22	12,04	16,92	17,01	16,84
	MW sais.	5,36	5,56	5,63	7,97	8,24	8,18	11,89	12,10	12,02	17,45	17,52	17,48
	MW total	5,42	5,51	5,58	7,99	8,18	8,11	11,99	12,16	12,03	17,18	17,26	17,16
	StAb stat.	6,21	6,24	6,49	9,07	8,83	8,85	13,56	13,34	13,11	19,13	18,83	18,63
	StAb sais.	6,10	6,32	6,51	9,05	9,10	9,11	13,56	13,48	13,40	20,15	19,82	19,82
	StAb total	5,97	6,10	6,31	8,79	8,70	8,72	13,15	13,01	12,86	19,06	18,76	18,66
MLLC ₁	MW stat.	3,48	3,44	3,55	1,91	1,94	1,97	2,24	2,29	2,26	3,01	3,03	3,00
	MW sais.	3,22	3,32	3,41	1,84	1,98	2,03	2,19	2,28	2,30	3,09	3,11	3,13
	MW total	3,35	3,38	3,48	1,87	1,96	2,00	2,22	2,28	2,28	3,05	3,07	3,06
	StAb stat.	3,76	3,85	4,09	2,14	2,10	2,22	2,52	2,49	2,45	3,42	3,34	3,31
	StAb sais.	3,46	3,67	3,85	2,00	2,16	2,31	2,47	2,52	2,56	3,57	3,50	3,56
	StAb total	3,51	3,65	3,86	2,01	2,07	2,19	2,42	2,43	2,43	3,39	3,32	3,34
MLLC ₂	MW stat.	3,18	3,14	3,25	0,99	1,01	1,06	0,76	0,79	0,79	0,92	0,92	0,92
	MW sais.	2,90	2,98	3,08	0,92	1,04	1,11	0,74	0,80	0,83	0,93	0,95	0,97
	MW total	3,04	3,06	3,17	0,95	1,03	1,08	0,75	0,80	0,81	0,93	0,94	0,94
	StAb stat.	3,39	3,50	3,74	1,11	1,09	1,22	0,86	0,85	0,84	1,06	1,01	1,00
	StAb sais.	3,07	3,27	3,46	0,95	1,12	1,28	0,80	0,87	0,93	1,07	1,04	1,12
	StAb total	3,14	3,28	3,49	1,00	1,07	1,22	0,81	0,84	0,86	1,03	1,00	1,03
MLLC ₃	MW stat.	3,08	3,04	3,15	0,67	0,69	0,74	0,25	0,28	0,28	0,20	0,20	0,20
	MW sais.	2,79	2,86	2,97	0,60	0,72	0,79	0,23	0,30	0,33	0,19	0,20	0,23
	MW total	2,93	2,95	3,06	0,64	0,70	0,77	0,24	0,29	0,31	0,20	0,20	0,22
	StAb stat.	3,27	3,37	3,61	0,75	0,75	0,88	0,29	0,30	0,29	0,24	0,21	0,21
	StAb sais.	2,93	3,13	3,32	0,59	0,76	0,93	0,23	0,31	0,37	0,22	0,20	0,28
	StAb total	3,02	3,16	3,37	0,66	0,73	0,88	0,25	0,29	0,33	0,22	0,20	0,24

Tabelle 8.2. MSE auf Trainings-, Validierungs- und Testdatensatz je Zeitreihe

	MLP SE			MLP LLC ₁			MLP LLC ₂			MLP LLC ₃		
	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
N _{niedrig}	26,43	27,55	26,48	52,96	55,19	57,98	105,32	105,92	113,73	167,51	167,69	179,02
N _{mittel}	163,51	159,86	150,61	314,67	359,40	309,43	593,54	665,91	585,13	931,95	1030,34	917,82
N _{hoch}	656,13	612,35	702,64	1289,56	1171,54	1244,79	2552,32	2383,41	2439,70	4658,08	4419,11	4453,71
NB _{niedrig}	2,54	2,45	2,21	4,12	4,05	3,69	11,53	11,53	11,01	19,98	19,95	19,44
NB _{mittel}	10,97	9,89	11,72	21,16	20,11	19,39	53,37	53,43	49,70	97,93	97,36	94,98
NB _{hoch}	59,19	58,52	55,40	106,15	113,86	114,85	227,34	237,12	240,74	459,06	471,83	482,18
P _{niedrig}	2,30	2,39	2,61	4,05	4,64	4,95	9,07	10,46	11,01	15,44	16,61	17,28
P _{mittel}	4,41	5,03	4,63	10,57	10,94	10,41	19,09	18,72	17,94	36,94	35,48	34,78
P _{hoch}	7,07	7,31	7,57	14,10	15,41	15,57	25,11	26,13	25,46	49,90	50,92	49,55
MW	103,62	98,37	107,10	201,93	195,02	197,90	399,63	390,29	388,27	715,20	701,03	694,31
Std.Abw.	213,61	199,31	228,35	419,80	383,44	404,68	828,86	776,59	791,54	1508,96	1433,17	1441,24
SN _{niedrig}	26,34	33,01	28,77	55,07	59,41	57,21	95,73	98,29	99,44	198,92	200,79	203,63
SN _{mittel}	165,85	176,14	172,55	303,33	354,75	327,52	544,42	631,72	571,72	1120,50	1224,51	1147,29
SN _{hoch}	663,33	655,17	750,53	1350,96	1291,97	1354,30	2679,41	2583,14	2636,18	5194,09	4986,29	5079,27
SNB _{niedrig}	2,54	2,38	2,49	4,69	4,56	4,65	11,55	11,17	11,40	23,34	22,98	23,26
SNB _{mittel}	10,94	11,33	15,53	22,27	20,91	20,61	57,87	57,31	55,88	110,23	109,78	109,02
SNB _{hoch}	55,26	63,20	65,40	114,20	125,38	127,74	237,79	251,27	252,74	424,15	434,21	439,55
SP _{niedrig}	2,46	2,75	2,95	4,49	5,79	5,90	8,66	10,20	10,53	15,53	17,56	17,75
SP _{mittel}	4,72	5,84	5,15	13,71	13,47	13,78	20,80	19,94	19,89	36,09	34,17	34,94
SP _{hoch}	7,14	7,90	8,41	15,63	16,39	16,66	30,61	31,47	30,71	54,33	54,90	54,10
MW	104,29	106,41	116,86	209,37	210,29	214,26	409,65	410,50	409,83	797,46	787,24	789,87
Std.Abw.	216,10	213,21	243,83	438,69	420,89	439,99	868,62	839,10	854,63	1686,24	1621,37	1648,93
MW	103,95	102,39	111,98	205,65	202,65	206,08	404,64	400,40	399,05	756,33	744,14	742,09
Std.Abw.	208,44	200,26	229,22	416,55	390,66	410,17	823,64	784,38	799,17	1552,86	1485,14	1503,14

Tabelle 8.3. MAE auf Trainings-, Validierungs- und Testdatensatz je Zeitreihe

	MLP SE			MLP LLC ₁			MLP LLC ₂			MLP LLC ₃		
	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
N _{niedrig}	4,18	4,22	4,00	5,93	5,99	6,34	9,00	9,01	9,53	11,88	11,86	12,38
N _{mittel}	10,34	10,08	9,64	14,50	16,05	14,67	21,20	22,87	21,53	27,68	29,39	27,84
N _{hoch}	20,01	20,24	21,24	29,55	28,28	28,80	44,52	43,27	42,87	63,25	61,81	61,51
NB _{niedrig}	1,18	1,17	1,11	1,74	1,75	1,64	3,10	3,11	3,06	4,24	4,22	4,19
NB _{mittel}	2,53	2,41	2,68	4,01	3,92	3,80	6,65	6,62	6,26	9,31	9,31	9,05
NB _{hoch}	6,02	5,84	5,84	8,87	9,14	9,24	13,45	13,94	14,01	20,08	20,55	20,73
P _{niedrig}	1,21	1,22	1,25	1,70	1,84	1,89	2,68	2,89	3,02	3,63	3,77	3,87
P _{mittel}	1,68	1,80	1,72	2,78	2,80	2,74	3,92	3,80	3,74	5,68	5,54	5,48
P _{hoch}	2,12	2,13	2,23	3,11	3,26	3,25	4,38	4,48	4,39	6,54	6,61	6,49
MW	5,47	5,46	5,52	8,02	8,11	8,04	12,10	12,22	12,04	16,92	17,01	16,84
Std.Abw.	6,21	6,24	6,49	9,07	8,83	8,85	13,56	13,34	13,11	19,13	18,83	18,63
SN _{niedrig}	4,08	4,34	4,09	5,85	6,04	6,18	8,16	8,21	8,61	12,73	12,72	13,14
SN _{mittel}	10,11	10,30	9,75	13,45	15,26	14,14	19,43	21,42	20,24	29,84	31,41	30,31
SN _{hoch}	19,72	20,50	21,35	29,87	29,58	29,97	44,94	44,18	44,27	66,29	64,79	65,12
SNB _{niedrig}	1,19	1,17	1,15	1,80	1,84	1,83	3,02	2,99	3,02	4,49	4,48	4,51
SNB _{mittel}	2,49	2,48	2,87	3,92	3,88	3,79	6,65	6,66	6,43	9,34	9,38	9,24
SNB _{hoch}	5,75	5,98	6,12	8,83	9,32	9,44	13,38	13,91	14,01	18,57	19,04	19,17
SP _{niedrig}	1,20	1,24	1,25	1,78	1,99	2,03	2,59	2,78	2,87	3,63	3,81	3,88
SP _{mittel}	1,67	1,86	1,78	2,95	2,91	2,91	3,94	3,82	3,84	5,30	5,18	5,18
SP _{hoch}	2,07	2,17	2,34	3,24	3,35	3,33	4,87	4,98	4,89	6,81	6,85	6,77
MW	5,36	5,56	5,63	7,97	8,24	8,18	11,89	12,10	12,02	17,45	17,52	17,48
Std.Abw.	6,10	6,32	6,51	9,05	9,10	9,11	13,56	13,48	13,40	20,15	19,82	19,82
MW	5,42	5,51	5,58	7,99	8,18	8,11	11,99	12,16	12,03	17,18	17,26	17,16
Std.Abw.	5,97	6,10	6,31	8,79	8,70	8,72	13,15	13,01	12,86	19,06	18,76	18,66

Tabelle 8.4 LLC1 auf Trainings-, Validierungs- und Testdatensatz je Zeitreihe

	MLP SE			MLP LLC ₁			MLP LLC ₂			MLP LLC ₃		
	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
N _{niedrig}	2,45	2,50	2,17	1,34	1,39	1,44	1,64	1,66	1,76	2,10	2,11	2,21
N _{mittel}	6,97	6,18	6,51	3,33	3,64	3,39	3,91	4,22	4,05	4,94	5,25	5,00
N _{hoch}	11,91	12,55	13,29	7,02	6,81	7,28	8,28	8,08	8,01	11,29	10,95	10,92
NB _{niedrig}	0,90	0,88	0,86	0,47	0,48	0,42	0,58	0,58	0,56	0,76	0,75	0,75
NB _{mittel}	1,67	1,60	1,90	1,03	0,98	1,05	1,23	1,23	1,20	1,66	1,66	1,60
NB _{hoch}	4,30	4,06	3,93	2,26	2,30	2,26	2,52	2,69	2,61	3,55	3,70	3,67
P _{niedrig}	0,79	0,74	0,73	0,44	0,46	0,46	0,49	0,54	0,58	0,64	0,67	0,70
P _{mittel}	0,96	1,09	1,07	0,59	0,64	0,63	0,72	0,71	0,70	1,00	0,98	0,97
P _{hoch}	1,39	1,37	1,47	0,70	0,77	0,79	0,82	0,86	0,85	1,16	1,17	1,15
MW	3,48	3,44	3,55	1,91	1,94	1,97	2,24	2,29	2,26	3,01	3,03	3,00
Std.Abw.	3,76	3,85	4,09	2,14	2,10	2,22	2,52	2,49	2,45	3,42	3,34	3,31
SN _{niedrig}	2,32	2,56	2,19	1,32	1,47	1,39	1,51	1,60	1,64	2,25	2,28	2,35
SN _{mittel}	6,57	5,93	6,16	3,31	3,64	3,55	3,62	4,03	3,94	5,29	5,59	5,43
SN _{hoch}	11,00	12,02	12,60	6,53	7,03	7,55	8,17	8,25	8,43	11,73	11,43	11,70
SNB _{niedrig}	0,80	0,80	0,79	0,44	0,46	0,46	0,57	0,56	0,57	0,80	0,79	0,80
SNB _{mittel}	1,63	1,65	1,98	1,04	1,01	1,09	1,25	1,25	1,22	1,66	1,67	1,65
SNB _{hoch}	3,66	3,70	3,74	2,15	2,33	2,34	2,50	2,67	2,65	3,28	3,43	3,42
SP _{niedrig}	0,91	0,87	0,87	0,41	0,47	0,48	0,48	0,53	0,56	0,64	0,68	0,71
SP _{mittel}	0,95	1,12	1,07	0,62	0,65	0,65	0,75	0,72	0,73	0,94	0,92	0,92
SP _{hoch}	1,16	1,21	1,33	0,71	0,76	0,78	0,88	0,92	0,91	1,20	1,22	1,20
MW	3,22	3,32	3,41	1,84	1,98	2,03	2,19	2,28	2,30	3,09	3,11	3,13
Std.Abw.	3,46	3,67	3,85	2,00	2,16	2,31	2,47	2,52	2,56	3,57	3,50	3,56
MW	3,35	3,38	3,48	1,87	1,96	2,00	2,22	2,28	2,28	3,05	3,07	3,06
Std.Abw.	3,51	3,65	3,86	2,01	2,07	2,19	2,42	2,43	2,43	3,39	3,32	3,34

Tabelle 8.5 LLC2 auf Trainings-, Validierungs- und Testdatensatz je Zeitreihe

	MLP SE			MLP LLC ₁			MLP LLC ₂			MLP LLC ₃		
	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
N _{niedrig}	2,19	2,24	1,90	0,65	0,69	0,70	0,53	0,56	0,59	0,63	0,64	0,68
N _{mittel}	6,47	5,60	6,04	1,65	1,77	1,69	1,31	1,42	1,42	1,53	1,62	1,56
N _{hoch}	10,69	11,39	12,10	3,64	3,58	4,05	2,83	2,79	2,77	3,48	3,30	3,31
NB _{niedrig}	0,85	0,84	0,82	0,28	0,28	0,24	0,20	0,20	0,19	0,24	0,23	0,23
NB _{mittel}	1,54	1,48	1,79	0,58	0,54	0,64	0,41	0,42	0,44	0,50	0,50	0,49
NB _{hoch}	4,04	3,80	3,64	1,26	1,27	1,21	0,87	0,99	0,90	1,07	1,17	1,10
P _{niedrig}	0,73	0,67	0,65	0,24	0,25	0,25	0,16	0,18	0,21	0,19	0,21	0,23
P _{mittel}	0,85	0,98	0,98	0,27	0,31	0,32	0,24	0,25	0,25	0,30	0,29	0,29
P _{hoch}	1,28	1,26	1,36	0,34	0,40	0,42	0,29	0,31	0,31	0,35	0,35	0,35
MW	3,18	3,14	3,25	0,99	1,01	1,06	0,76	0,79	0,79	0,92	0,92	0,92
Std.Abw.	3,39	3,50	3,74	1,11	1,09	1,22	0,86	0,85	0,84	1,06	1,01	1,00
SN _{niedrig}	2,06	2,29	1,91	0,64	0,78	0,67	0,52	0,60	0,59	0,67	0,71	0,73
SN _{mittel}	6,04	5,27	5,62	1,78	1,89	1,95	1,25	1,41	1,49	1,60	1,70	1,69
SN _{hoch}	9,69	10,74	11,29	3,02	3,64	4,18	2,64	2,85	3,04	3,53	3,41	3,67
SNB _{niedrig}	0,75	0,75	0,74	0,24	0,26	0,25	0,21	0,19	0,20	0,24	0,24	0,24
SNB _{mittel}	1,50	1,53	1,85	0,60	0,58	0,69	0,44	0,43	0,44	0,50	0,51	0,51
SNB _{hoch}	3,34	3,36	3,39	1,14	1,28	1,27	0,86	0,98	0,94	0,98	1,09	1,05
SP _{niedrig}	0,87	0,82	0,81	0,20	0,25	0,25	0,16	0,19	0,22	0,19	0,21	0,23
SP _{mittel}	0,84	1,01	0,96	0,27	0,31	0,31	0,27	0,26	0,26	0,29	0,28	0,28
SP _{hoch}	1,02	1,06	1,18	0,33	0,38	0,40	0,28	0,31	0,32	0,36	0,37	0,37
MW	2,90	2,98	3,08	0,92	1,04	1,11	0,74	0,80	0,83	0,93	0,95	0,97
Std.Abw.	3,07	3,27	3,46	0,95	1,12	1,28	0,80	0,87	0,93	1,07	1,04	1,12
MW	3,04	3,06	3,17	0,95	1,03	1,08	0,75	0,80	0,81	0,93	0,94	0,94
Std.Abw.	3,14	3,28	3,49	1,00	1,07	1,22	0,81	0,84	0,86	1,03	1,00	1,03

Tabelle 8.6 MLLC₃ auf Trainings-, Validierungs- und Testdatensatz je Zeitreihe

	MLP SE			MLP LLC ₁			MLP LLC ₂			MLP LLC ₃		
	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
N _{niedrig}	2,10	2,15	1,80	0,42	0,46	0,44	0,15	0,18	0,19	0,13	0,14	0,16
N _{mittel}	6,29	5,40	5,87	1,07	1,13	1,11	0,41	0,45	0,52	0,35	0,38	0,38
N _{hoch}	10,27	11,00	11,69	2,47	2,47	2,94	0,96	0,97	0,97	0,80	0,67	0,69
NB _{niedrig}	0,84	0,83	0,81	0,21	0,22	0,18	0,07	0,07	0,06	0,06	0,05	0,05
NB _{mittel}	1,49	1,44	1,75	0,43	0,39	0,50	0,13	0,15	0,18	0,11	0,11	0,10
NB _{hoch}	3,95	3,71	3,54	0,92	0,92	0,85	0,31	0,41	0,31	0,22	0,30	0,22
P _{niedrig}	0,70	0,64	0,62	0,18	0,18	0,18	0,05	0,06	0,09	0,04	0,05	0,06
P _{mittel}	0,82	0,95	0,94	0,15	0,20	0,21	0,07	0,09	0,09	0,06	0,06	0,06
P _{hoch}	1,24	1,22	1,32	0,22	0,27	0,29	0,10	0,12	0,13	0,07	0,07	0,07
MW	3,08	3,04	3,15	0,67	0,69	0,74	0,25	0,28	0,28	0,20	0,20	0,20
Std.Abw.	3,27	3,37	3,61	0,75	0,75	0,88	0,29	0,30	0,29	0,24	0,21	0,21
SN _{niedrig}	1,97	2,19	1,81	0,40	0,54	0,42	0,17	0,26	0,23	0,13	0,17	0,17
SN _{mittel}	5,86	5,04	5,43	1,26	1,29	1,41	0,43	0,52	0,65	0,33	0,37	0,40
SN _{hoch}	9,24	10,30	10,84	1,82	2,48	3,03	0,75	1,00	1,19	0,71	0,65	0,91
SNB _{niedrig}	0,73	0,73	0,72	0,17	0,19	0,18	0,08	0,06	0,08	0,05	0,05	0,05
SNB _{mittel}	1,45	1,48	1,80	0,45	0,43	0,55	0,16	0,16	0,17	0,11	0,11	0,12
SNB _{hoch}	3,23	3,24	3,26	0,80	0,92	0,90	0,30	0,40	0,35	0,19	0,28	0,24
SP _{niedrig}	0,86	0,80	0,79	0,13	0,17	0,17	0,05	0,08	0,10	0,04	0,05	0,07
SP _{mittel}	0,81	0,97	0,92	0,15	0,19	0,19	0,10	0,10	0,10	0,06	0,06	0,06
SP _{hoch}	0,97	1,01	1,13	0,20	0,24	0,27	0,07	0,10	0,11	0,07	0,08	0,08
MW	2,79	2,86	2,97	0,60	0,72	0,79	0,23	0,30	0,33	0,19	0,20	0,23
Std.Abw.	2,93	3,13	3,32	0,59	0,76	0,93	0,23	0,31	0,37	0,22	0,20	0,28
MW	2,93	2,95	3,06	0,64	0,70	0,77	0,24	0,29	0,31	0,20	0,20	0,22
Std.Abw.	3,02	3,16	3,37	0,66	0,73	0,88	0,25	0,29	0,33	0,22	0,20	0,24

8.2 Graphen der PQ-Diagramme für Poisson- und Negative Binomialverteilung

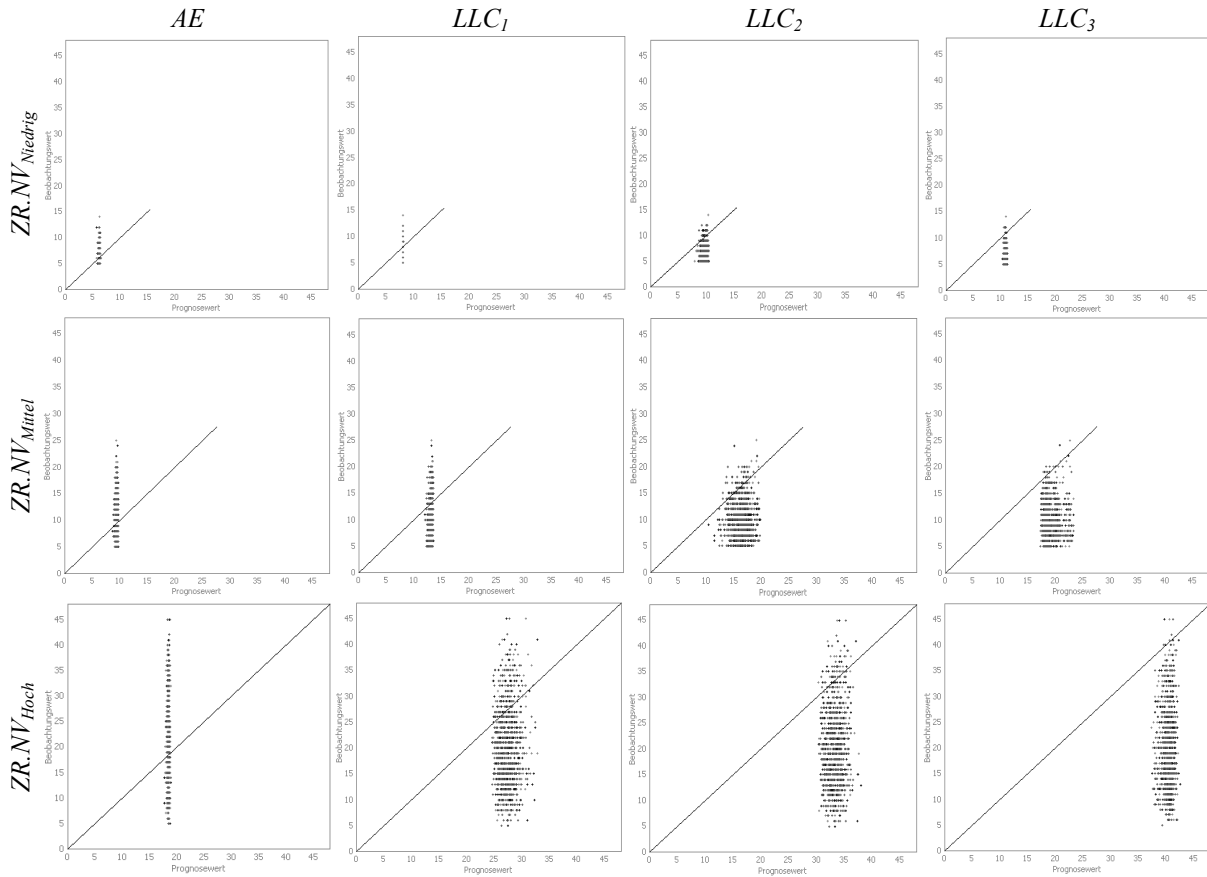


Abbildung 8.1. PQ-Streudiagramme der Prognosewerte auf Trainings-, Validierungs- und Testmenge für 4 Zielfunktionen und drei stationäre Zeitreihen mit negativ binomial verteilten Zufallszahlen

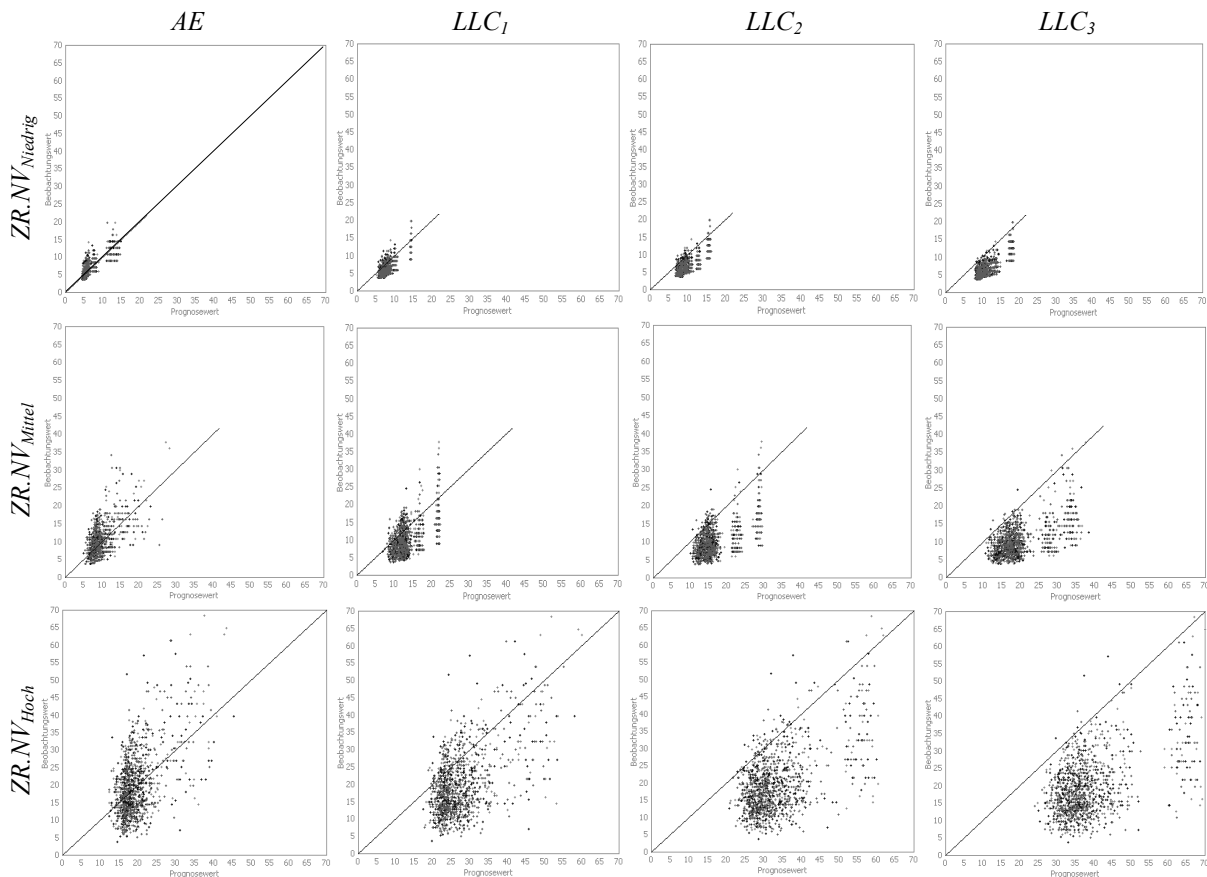


Abbildung 8.2. PQ-Streudiagramme für saisonale Zeitreihen mit negativ binomial verteilten Zufallszahlen

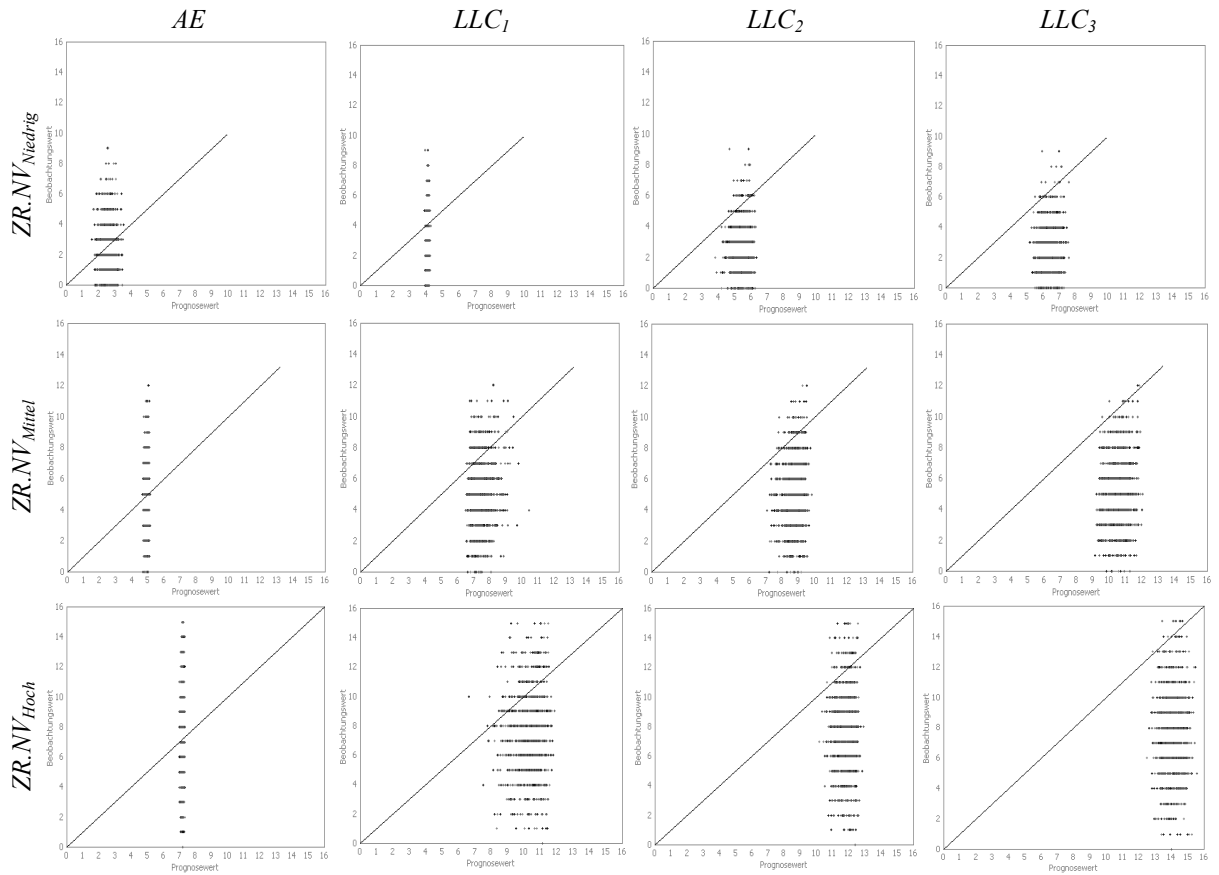


Abbildung 8.3. PQ-Streudiagramme der Prognosewerte auf Trainings-, Validierungs- und Testmenge für 4 Zielfunktionen und drei stationären Zeitreihen mit poisson verteilten Zufallszahlen

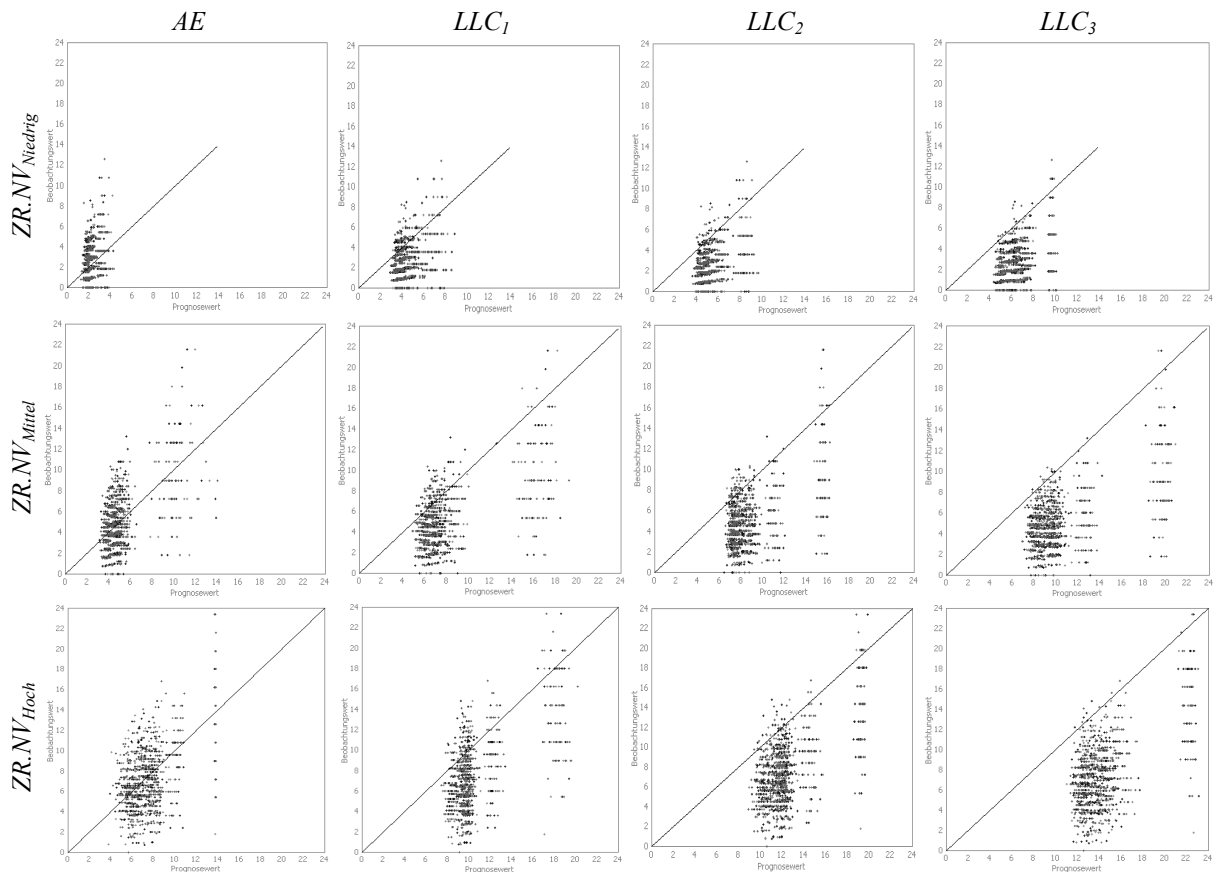


Abbildung 8.4. PQ-Streudiagramme für drei saisonalen Zeitreihen mit poisson verteilten Zufallszahlen

8.3 Graphen des vorhergesagten Bestellmengen aller Verfahren zur Disposition

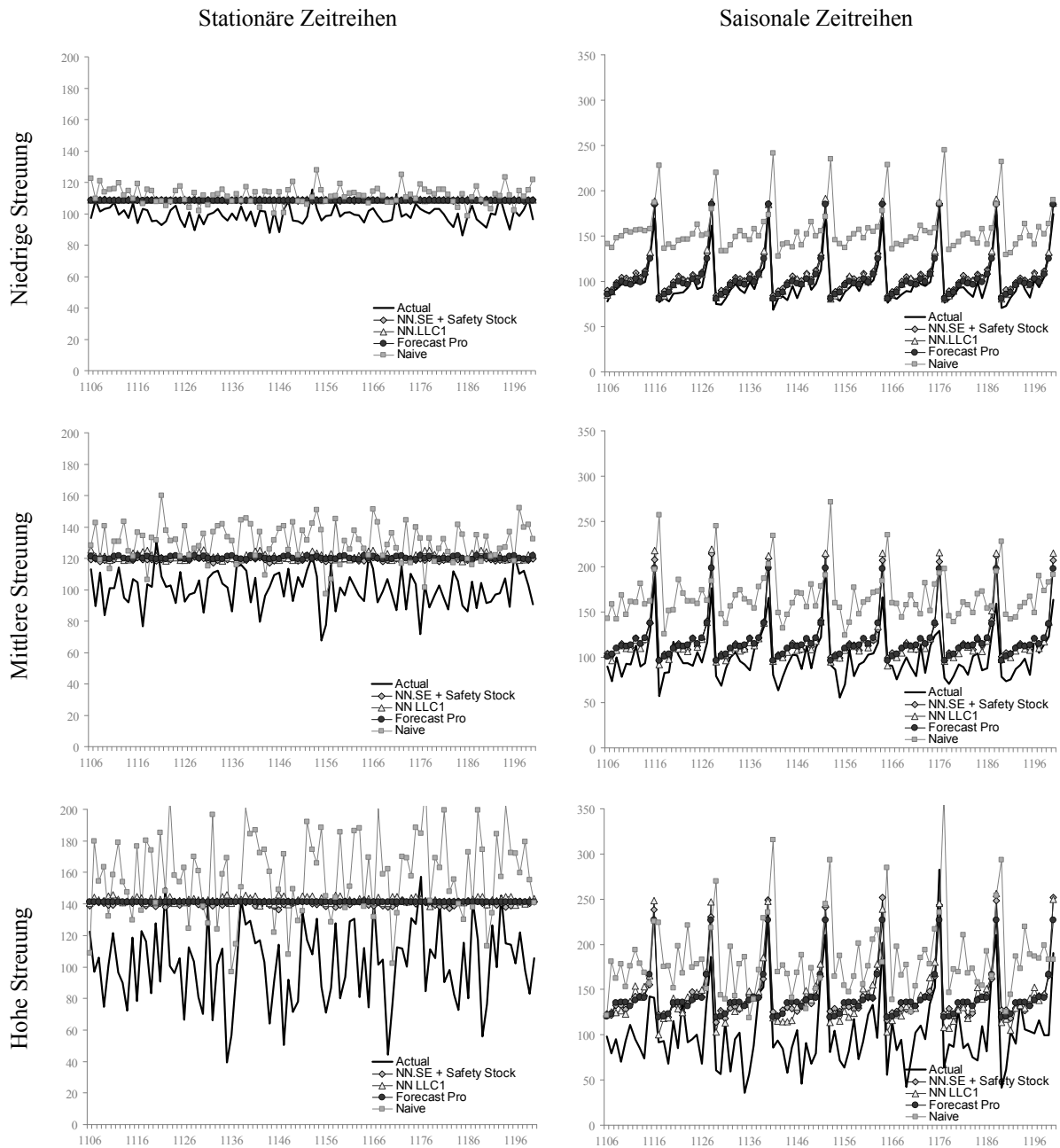


Abbildung 8.5. Ausschnitte der Zeitreihen der Normalverteilung der Beobachtungswerte und Bestellmengen der NN trainiert mit LLC3, ForecastPro sowie NN trainiert mit SE und dem Naiven Verfahren unter konventioneller Berechnung der Sicherheitsbestände

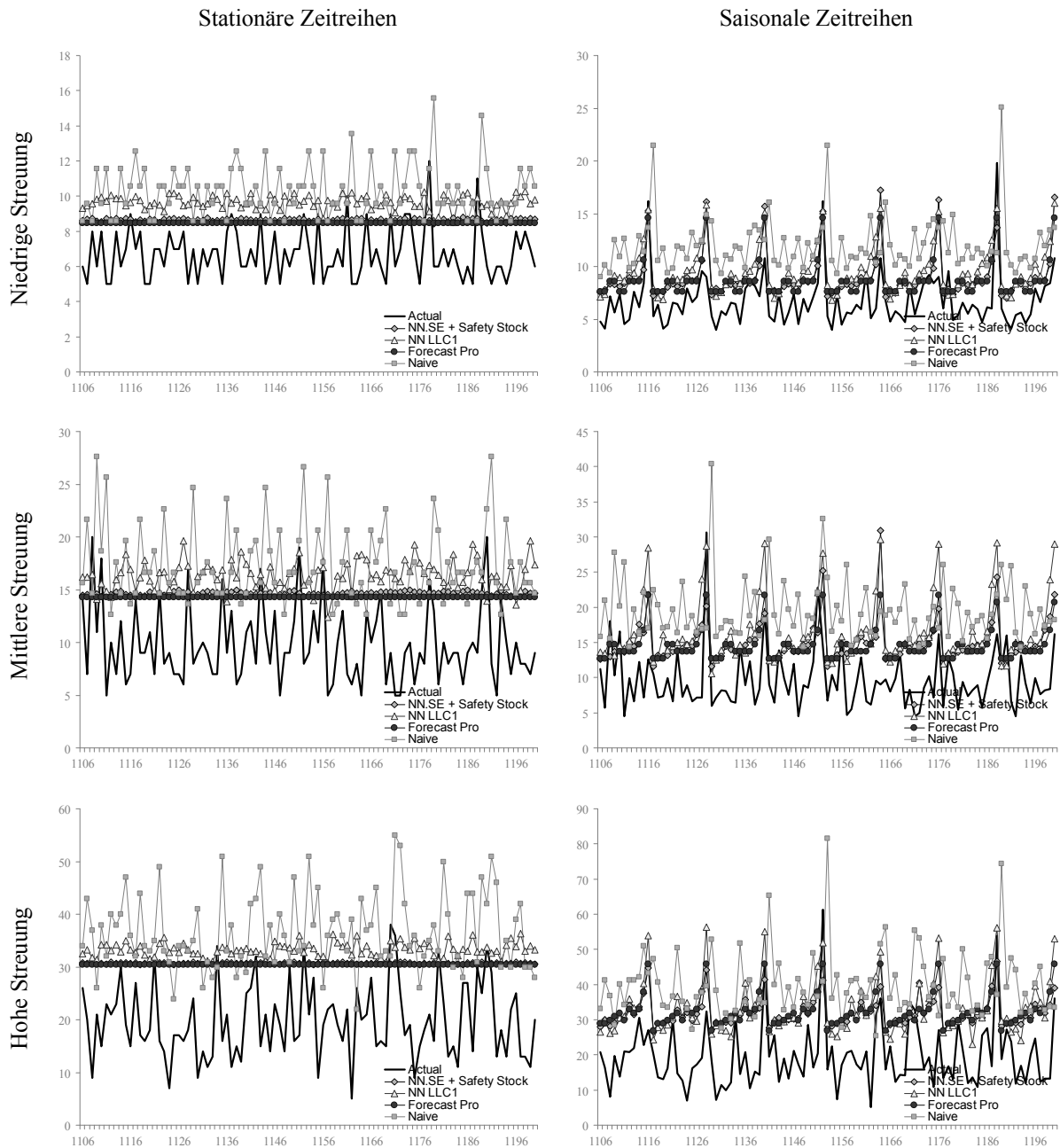


Abbildung 8.6. Ausschnitte der Zeitreihen mit Negativer Binomialverteilung der NN trainiert mit LLC3, ForecastPro sowie NN trainiert mit SE und dem Naiven Verfahren unter konventioneller Berechnung der Sicherheitsbestände

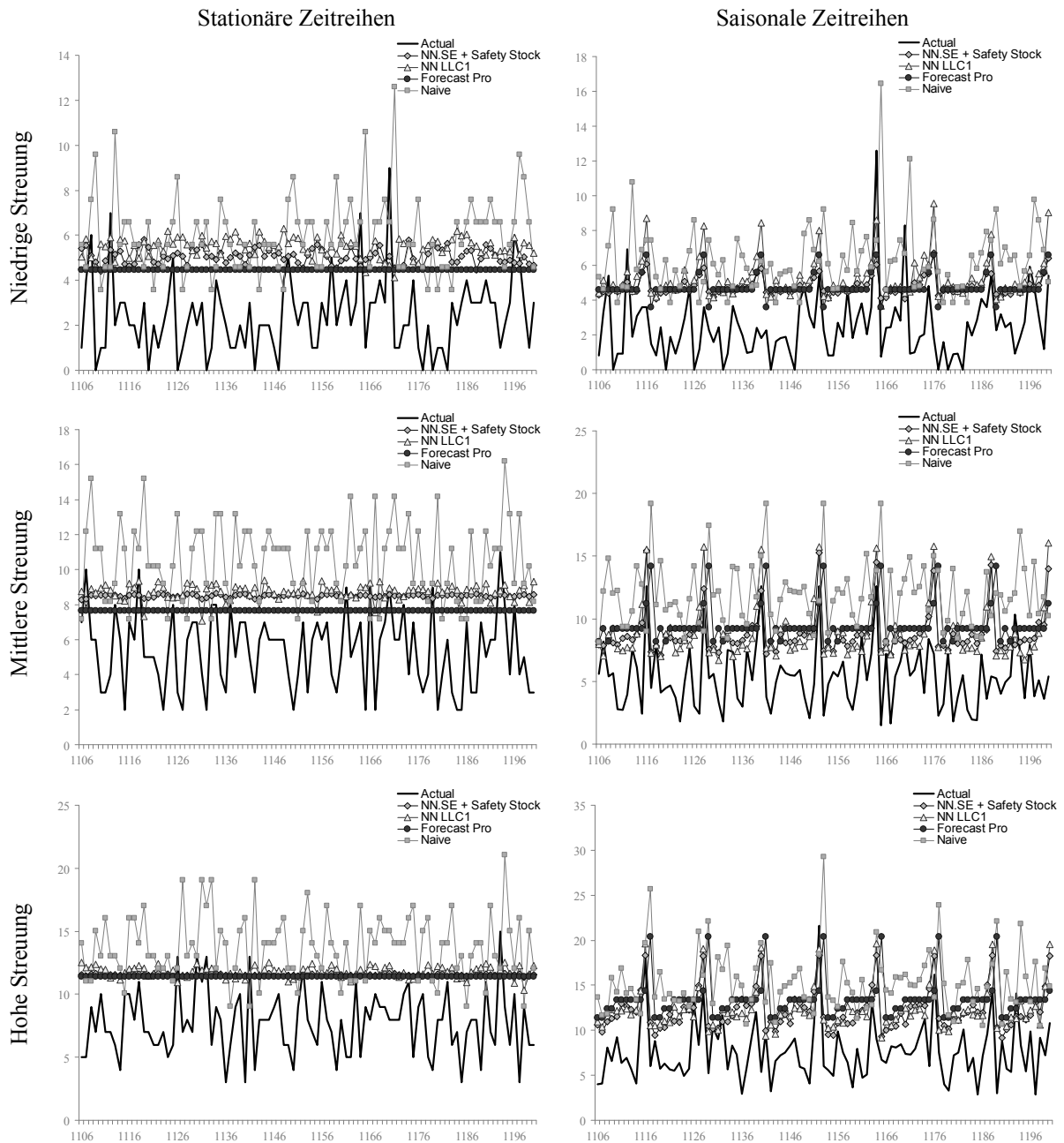


Abbildung 8.7. Ausschnitte der Zeitreihen mit Poisson-Verteilung der NN trainiert mit LLC3, ForecastPro sowie NN trainiert mit SE und dem Naiven Verfahren unter konventioneller Berechnung der Sicherheitsbestände

8.4 Darstellung ausgewählter Zeitreihen der Absatzstelle AU-1

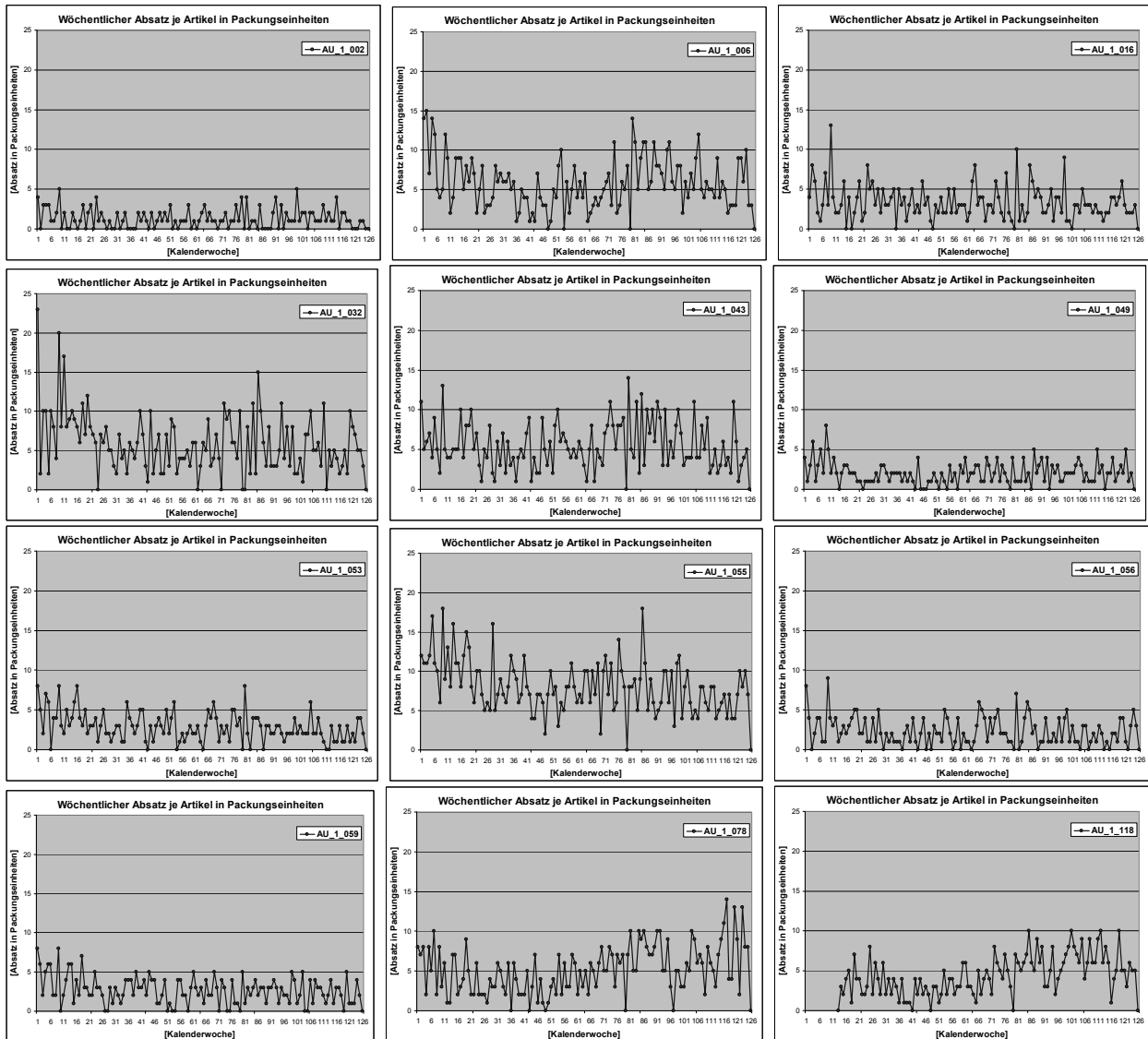


Abbildung 8.8. Ausgewählte Zeitreihen der Absatzstelle AU-1

8.5 Ergebnistabellen der Bedarfsprognose

8.5.1 Ergebnisse der Bedarfsprognose je Zeitreihe

Tabelle 8.7. MAE und Rang nach MAE je Zeitreihe und Verfahren an Absatzstelle AU-1

MAE	S1	S2	S3	S4	S5	S6	S7
NF1	(1,41 / 1,72) 0,95	(2,72 / 3,47) 2,48	(2,57 / 2,67) 1,00	(3,85 / 4,42) 3,05	(3,01 / 4,36) 3,29	(1,40 / 1,75) 1,71	(1,94 / 1,72) 1,52
NF2	(1,49 / 1,50) 1,14	(3,18 / 3,31) 3,48	(2,38 / 2,69) 1,33	(4,04 / 4,67) 3,62	(3,04 / 4,19) 2,81	(1,47 / 1,19) 1,48	(2,12 / 1,81) 1,43
MA	(1,13 / 1,36) 1,00	(2,47 / 2,86) 2,62	(2,12 / 2,33) 1,24	(2,99 / 3,78) 2,76	(2,37 / 3,31) 2,48	(1,18 / 1,28) 1,43	(1,69 / 1,50) 1,33
S.ES	(1,19 / 1,31) 0,95	(2,35 / 2,75) 2,71	(2,03 / 2,19) 1,19	(3,15 / 3,67) 2,33	(2,50 / 3,22) 2,38	(1,18 / 1,19) 1,38	(1,69 / 1,44) 1,38
DT.ES	(1,17 / 1,25) 1,19	(2,39 / 2,86) 3,52	(1,77 / 2,08) 1,52	(2,57 / 3,50) 2,10	(2,25 / 3,06) 3,38	(1,17 / 1,44) 1,43	(1,49 / 1,22) 1,48
EXP.ES	(1,17 / 1,25) 1,19	(2,32 / 2,78) 3,52	(1,80 / 2,08) 1,52	(2,90 / 3,36) 2,52	(2,28 / 3,08) 3,38	(1,16 / 1,31) 1,43	(1,71 / 1,33) 1,48
EXP.ARIMA	(1,07 / 1,06) 1,19	(2,43 / 2,72) 3,24	(1,77 / 2,08) 1,52	(3,01 / 3,39) 3,14	(2,26 / 2,94) 2,67	(1,29 / 1,31) 1,43	(1,71 / 1,72) 2,19
EXP.ES&ARIMA	(1,17 / 1,25) 1,19	(2,43 / 2,72) 3,24	(1,80 / 2,08) 1,52	(3,01 / 3,39) 3,14	(2,26 / 2,94) 2,67	(1,16 / 1,31) 1,43	(1,71 / 1,33) 1,48
mv.EXP.ES	(1,17 / 1,25) 1,19	(2,19 / 2,72) 2,57	(1,78 / 1,97) 1,33	(2,57 / 3,25) 1,76	(2,07 / 3,00) 2,38	(1,16 / 1,31) 1,43	(1,71 / 1,33) 1,48
MLP	(0,92 / 1,00) 1,00	(2,10 / 2,44) 2,86	(1,49 / 1,94) 1,00	(2,08 / 3,17) 2,48	(2,44 / 2,50) 3,24	(0,98 / 1,08) 1,00	(1,31 / 1,19) 1,48
mv.MLP	(1,08 / 0,97) 1,00	(2,13 / 2,47) 2,90	(1,23 / 1,75) 1,00	(2,05 / 3,14) 2,19	(2,34 / 2,44) 3,00	(0,90 / 1,03) 1,14	(1,31 / 1,22) 1,43
Verfahren	S8	S9	S10	S11	S12	mwMAE	Rank mwMAE
NF1	(3,18 / 4,08) 2,24	(1,90 / 2,00) 1,57	(1,87 / 1,86) 1,90	(2,85 / 2,47) 4,05	(2,05 / 2,28) 2,62	(2,40 / 2,73) 2,20	(11 / 11) 9
NF2	(2,71 / 4,03) 2,48	(1,97 / 2,14) 1,43	(1,90 / 1,33) 1,71	(2,28 / 2,50) 3,48	(2,17 / 2,00) 2,57	(2,40 / 2,61) 2,25	(10 / 10) 10
MA	(2,85 / 3,33) 2,19	(1,76 / 1,81) 1,33	(1,76 / 1,53) 1,48	(2,26 / 2,39) 3,67	(1,61 / 2,14) 2,33	(2,02 / 2,30) 1,99	(9 / 9) 5
S.ES	(2,62 / 3,19) 2,14	(1,65 / 1,83) 1,38	(1,69 / 1,53) 1,48	(2,13 / 2,17) 3,43	(1,63 / 1,89) 2,33	(1,98 / 2,20) 1,92	(8 / 8) 4
DT.ES	(2,39 / 2,97) 1,90	(1,67 / 1,69) 1,57	(1,49 / 1,33) 1,33	(2,10 / 2,22) 2,95	(1,51 / 1,78) 2,95	(1,83 / 2,12) 2,11	(4 / 7) 6
EXP.ES	(2,49 / 2,92) 2,19	(1,67 / 1,69) 1,57	(1,65 / 1,36) 1,33	(2,09 / 2,17) 2,95	(1,53 / 1,81) 2,90	(1,90 / 2,09) 2,17	(5 / 6) 8
EXP.ARIMA	(2,57 / 2,97) 2,86	(1,67 / 1,69) 1,57	(1,55 / 1,31) 1,48	(2,01 / 2,03) 3,05	(1,50 / 1,81) 2,90	(1,90 / 2,09) 2,27	(6 / 5) 11
EXP.ES&ARIMA	(2,49 / 2,92) 2,19	(1,67 / 1,69) 1,57	(1,65 / 1,36) 1,33	(2,09 / 2,17) 2,95	(1,53 / 1,81) 2,90	(1,91 / 2,08) 2,13	(7 / 4) 7
mv.EXP.ES	(2,23 / 2,78) 1,52	(1,67 / 1,69) 1,57	(1,65 / 1,36) 1,33	(2,09 / 2,17) 2,95	(1,53 / 1,81) 2,90	(1,82 / 2,05) 1,87	(3 / 3) 1
MLP	(2,46 / 2,75) 2,19	(1,44 / 1,39) 1,19	(1,46 / 1,28) 1,33	(2,07 / 1,94) 3,19	(2,45 / 1,81) 1,86	(1,77 / 1,88) 1,90	(2 / 2) 3
mv.MLP	(2,21 / 2,92) 1,95	(1,54 / 1,47) 1,52	(1,62 / 1,17) 1,19	(1,98 / 1,94) 3,33	(2,31 / 1,89) 1,90	(1,73 / 1,87) 1,88	(1 / 1) 2
Rang des MAE	S1	S2	S3	S4	S5	S6	S7
NF1	(10 / 11) 1	(10 / 11) 1	(11 / 10) 1	(10 / 10) 8	(10 / 11) 9	(10 / 11) 11	(10 / 9) 10
NF2	(11 / 10) 6	(11 / 10) 9	(10 / 11) 6	(11 / 11) 11	(11 / 10) 6	(11 / 3) 10	(11 / 11) 3
MA	(4 / 9) 3	(9 / 8) 3	(9 / 9) 5	(6 / 9) 7	(7 / 9) 3	(7 / 5) 4	(4 / 8) 1
S.ES	(9 / 8) 1	(5 / 6) 4	(8 / 8) 4	(9 / 8) 4	(9 / 8) 1	(7 / 3) 3	(4 / 7) 2
DT.ES	(5 / 4) 7	(6 / 8) 10	(3 / 4) 8	(3 / 7) 2	(2 / 6) 10	(6 / 10) 4	(3 / 2) 5
EXP.ES	(5 / 4) 7	(4 / 7) 10	(6 / 4) 8	(5 / 4) 6	(5 / 7) 10	(3 / 6) 4	(6 / 4) 5
EXP.ARIMA	(2 / 3) 7	(7 / 3) 7	(3 / 4) 8	(7 / 5) 9	(3 / 3) 4	(9 / 6) 4	(6 / 9) 11
EXP.ES&ARIMA	(5 / 4) 7	(7 / 3) 7	(6 / 4) 8	(7 / 5) 9	(3 / 3) 4	(3 / 6) 4	(6 / 4) 5
mv.EXP.ES	(5 / 4) 7	(3 / 3) 2	(5 / 3) 6	(3 / 3) 1	(1 / 5) 1	(3 / 6) 4	(6 / 4) 5
MLP	(1 / 2) 3	(1 / 1) 5	(2 / 2) 1	(2 / 2) 5	(8 / 2) 8	(2 / 2) 1	(1 / 1) 5
mv.MLP	(3 / 1) 3	(2 / 2) 6	(1 / 1) 1	(1 / 1) 3	(6 / 1) 7	(1 / 1) 2	(1 / 2) 3
Verfahren	S8	S9	S10	S11	S12	Sum of Ranks MAE	Rank of SoR MAE
NF1	(11 / 11) 9	(10 / 10) 6	(10 / 11) 11	(11 / 10) 11	(8 / 11) 6	(121 / 126) 84	(10 / 11) 9
NF2	(9 / 10) 10	(11 / 11) 4	(11 / 4) 10	(10 / 11) 9	(9 / 9) 5	(126 / 111) 89	(11 / 10) 11
MA	(10 / 9) 5	(9 / 8) 2	(9 / 9) 7	(9 / 9) 10	(6 / 10) 3	(89 / 102) 53	(9 / 9) 5
S.ES	(8 / 8) 4	(3 / 9) 3	(8 / 9) 7	(8 / 4) 8	(7 / 7) 3	(85 / 85) 44	(8 / 8) 4
DT.ES	(3 / 6) 2	(4 / 3) 6	(2 / 4) 2	(7 / 8) 1	(2 / 1) 11	(46 / 63) 68	(4 / 7) 7
EXP.ES	(5 / 3) 5	(4 / 3) 6	(5 / 6) 2	(4 / 4) 1	(3 / 2) 7	(55 / 54) 71	(6 / 6) 8
EXP.ARIMA	(7 / 6) 11	(4 / 3) 6	(3 / 3) 7	(1 / 2) 5	(1 / 2) 7	(54 / 50) 86	(5 / 5) 10
EXP.ES&ARIMA	(5 / 3) 5	(4 / 3) 6	(5 / 6) 2	(4 / 4) 1	(3 / 2) 7	(58 / 47) 65	(7 / 4) 6
mv.EXP.ES	(2 / 2) 1	(4 / 3) 6	(5 / 6) 2	(4 / 4) 1	(3 / 2) 7	(44 / 45) 43	(3 / 3) 1
MLP	(4 / 1) 5	(1 / 1) 1	(1 / 2) 2	(3 / 1) 6	(11 / 2) 1	(37 / 19) 43	(2 / 1) 1
mv.MLP	(1 / 3) 3	(2 / 2) 5	(4 / 1) 1	(1 / 1) 7	(10 / 7) 2	(33 / 23) 43	(1 / 2) 1

Tabelle 8.8. MAPE und Rang nach MAPE je Zeitreihe und Verfahren an Absatzstelle AU-1

MAPE je Zeitreihe	S1	S2	S3	S4	S5	S6	S7
NF1	(104,8% / 99,2%) 71,0%	(79,2% / 66,1%) 59,6%	(93,4% / 126,6%) 41,7%	(89,9% / 99,0%) 70,0%	(91,6% / 80,4%) 104,2%	(86,6% / 97,9%) 101,9%	(78,2% / 73,3%) 95,6%
NF2	(115,0% / 97,0%) 98,0%	(90,8% / 59,2%) 84,0%	(87,3% / 108,1%) 59,9%	(101,1% / 115,2%) 86,3%	(94,1% / 85,1%) 91,2%	(87,1% / 78,3%) 90,4%	(85,5% / 70,9%) 85,7%
MA	(97,7% / 95,0%) 105,6%	(78,6% / 54,7%) 63,7%	(85,3% / 109,6%) 56,7%	(80,8% / 89,1%) 68,3%	(82,8% / 63,9%) 94,2%	(80,2% / 75,4%) 89,0%	(75,4% / 66,6%) 86,1%
S.ES	(101,4% / 97,3%) 99,2%	(80,4% / 53,9%) 68,4%	(83,6% / 111,3%) 55,2%	(84,8% / 92,5%) 60,7%	(88,7% / 64,9%) 86,0%	(84,9% / 81,8%) 90,5%	(77,2% / 69,4%) 90,5%
DT.ES	(101,8% / 100,3%) 108,7%	(77,4% / 54,7%) 96,5%	(74,7% / 105,6%) 74,6%	(61,1% / 99,9%) 54,7%	(84,3% / 59,9%) 124,2%	(80,3% / 94,3%) 97,9%	(69,2% / 58,8%) 110,7%
EXP.ES	(101,8% / 100,3%) 108,7%	(81,2% / 52,6%) 96,5%	(75,8% / 105,6%) 74,6%	(75,9% / 89,8%) 72,1%	(87,8% / 59,1%) 124,2%	(81,5% / 86,4%) 97,9%	(83,9% / 60,4%) 110,7%
EXP.ARIMA	(98,5% / 94,7%) 108,7%	(84,4% / 52,6%) 86,0%	(74,7% / 105,6%) 74,6%	(77,8% / 100,3%) 89,7%	(88,6% / 52,0%) 98,3%	(91,1% / 83,4%) 97,9%	(84,8% / 86,8%) 157,1%
EXP.ES&ARIMA	(101,8% / 100,3%) 108,7%	(84,4% / 52,6%) 86,0%	(75,8% / 105,6%) 74,6%	(77,8% / 100,3%) 89,7%	(88,6% / 52,0%) 98,3%	(81,5% / 86,4%) 97,9%	(83,9% / 60,4%) 110,7%
mv.EXP.ES	(101,8% / 100,3%) 108,7%	(67,0% / 44,1%) 69,0%	(75,7% / 100,0%) 65,1%	(59,8% / 80,8%) 46,6%	(62,5% / 52,1%) 88,8%	(81,5% / 86,4%) 97,9%	(83,9% / 60,4%) 110,7%
MLP	(98,1% / 89,2%) 96,0%	(92,7% / 48,7%) 68,2%	(58,0% / 81,5%) 48,0%	(61,3% / 83,5%) 68,0%	(103,3% / 48,6%) 119,0%	(69,7% / 61,4%) 60,5%	(66,4% / 58,5%) 110,7%
mv.MLP	(104,1% / 90,3%) 101,6%	(90,3% / 50,9%) 74,2%	(50,9% / 69,5%) 47,2%	(50,5% / 79,1%) 59,5%	(94,0% / 50,4%) 117,7%	(73,8% / 68,5%) 75,8%	(64,8% / 62,0%) 106,0%
Verfahren	S8	S9	S10	S11	S12	mwMAPE	Rank mwMAPE
NF1	(45,7% / 60,6%) 41,4%	(102,0% / 101,9%) 107,1%	(78,7% / 96,3%) 107,1%	(98,9% / 49,3%) 78,8%	(84,7% / 50,0%) 71,0%	(86,2% / 83,4%) 79,1%	(10, / 11) 5
NF2	(39,3% / 62,0%) 42,7%	(110,8% / 122,2%) 87,5%	(82,8% / 74,4%) 110,6%	(86,2% / 49,0%) 68,4%	(90,4% / 40,2%) 87,4%	(89,2% / 80,1%) 82,7%	(11, / 10) 7
MA	(43,8% / 51,2%) 38,9%	(96,5% / 99,3%) 87,4%	(85,1% / 89,0%) 87,4%	(90,8% / 46,2%) 80,0%	(70,3% / 47,8%) 69,3%	(80,6% / 74,0%) 77,2%	(5, / 6) 3
S.ES	(41,1% / 50,8%) 41,3%	(95,0% / 110,2%) 89,8%	(82,4% / 93,3%) 92,1%	(89,1% / 42,7%) 74,5%	(76,0% / 43,0%) 73,9%	(82,0% / 75,9%) 76,8%	(9, / 9) 2
DT.ES	(37,0% / 48,0%) 37,6%	(99,7% / 107,7%) 109,0%	(71,3% / 86,0%) 92,4%	(88,3% / 44,0%) 63,5%	(77,4% / 41,4%) 87,1%	(76,9% / 75,0%) 88,1%	(2, / 7) 8
EXP.ES	(41,3% / 46,3%) 46,8%	(99,3% / 107,7%) 109,0%	(81,5% / 86,5%) 92,4%	(91,1% / 42,9%) 63,5%	(76,4% / 40,6%) 87,1%	(81,5% / 73,2%) 90,3%	(7, / 4) 10
EXP.ARIMA	(41,3% / 49,8%) 59,1%	(99,3% / 107,7%) 109,0%	(71,0% / 86,9%) 99,9%	(85,4% / 40,6%) 66,7%	(74,2% / 40,6%) 87,1%	(80,9% / 75,1%) 94,5%	(6, / 8) 11
EXP.ES&ARIMA	(41,3% / 46,3%) 46,8%	(99,3% / 107,7%) 109,0%	(81,5% / 86,5%) 92,4%	(91,1% / 42,9%) 63,5%	(76,4% / 40,6%) 87,1%	(82,0% / 73,5%) 88,7%	(8, / 5) 9
mv.EXP.ES	(33,4% / 44,3%) 29,1%	(99,3% / 107,7%) 109,0%	(81,5% / 86,5%) 92,4%	(91,1% / 42,9%) 63,5%	(76,4% / 40,6%) 87,1%	(76,2% / 70,5%) 80,7%	(1, / 3) 6
MLP	(38,1% / 43,0%) 46,8%	(76,5% / 75,1%) 84,4%	(70,6% / 82,7%) 94,0%	(97,6% / 39,6%) 67,3%	(128,0% / 42,0%) 55,8%	(80,0% / 62,8%) 76,6%	(4, / 1) 1
mv.MLP	(34,8% / 44,6%) 41,7%	(84,6% / 88,8%) 106,0%	(70,3% / 74,4%) 81,7%	(95,5% / 40,4%) 71,3%	(121,3% / 42,2%) 60,6%	(77,9% / 63,4%) 78,6%	(3, / 2) 4
Rang des MAPE	S1	S2	S3	S4	S5	S6	S7
NF1	(10 / 7) 1	(4 / 11) 1	(11 / 11) 1	(10 / 7) 7	(8 / 10) 7	(9 / 11) 11	(6 / 10) 4
NF2	(11 / 5) 3	(10 / 10) 7	(10 / 8) 6	(11 / 11) 9	(10 / 11) 3	(10 / 4) 4	(11 / 9) 1
MA	(1 / 4) 6	(3 / 9) 2	(9 / 9) 5	(8 / 4) 6	(2 / 8) 4	(3 / 3) 3	(4 / 7) 2
S.ES	(4 / 6) 4	(5 / 7) 4	(8 / 10) 4	(9 / 6) 4	(7 / 9) 1	(8 / 5) 5	(5 / 8) 3
DT.ES	(5 / 8) 7	(2 / 8) 10	(3 / 4) 8	(3 / 8) 2	(3 / 7) 10	(4 / 10) 6	(3 / 2) 6
EXP.ES	(5 / 8) 7	(6 / 6) 10	(6 / 4) 8	(5 / 5) 8	(4 / 6) 10	(5 / 7) 6	(7 / 3) 6
EXP.ARIMA	(3 / 3) 7	(7 / 4) 8	(3 / 4) 8	(6 / 9) 10	(5 / 3) 5	(11 / 6) 6	(10 / 11) 11
EXP.ES&ARIMA	(5 / 8) 7	(7 / 4) 8	(6 / 4) 8	(6 / 9) 10	(5 / 3) 5	(5 / 7) 6	(7 / 3) 6
mv.EXP.ES	(5 / 8) 7	(1 / 1) 5	(5 / 3) 7	(2 / 2) 1	(1 / 5) 2	(5 / 7) 6	(7 / 3) 6
MLP	(2 / 1) 2	(11 / 2) 3	(2 / 2) 3	(4 / 3) 5	(11 / 1) 9	(1 / 1) 1	(2 / 1) 6
mv.MLP	(9 / 2) 5	(9 / 3) 6	(1 / 1) 2	(1 / 1) 3	(9 / 2) 8	(2 / 2) 2	(1 / 6) 5
Verfahren	S8	S9	S10	S11	S12	Sum of Ranks MAPE	Rank of SoR MAPE
NF1	(11 / 10) 5	(10 / 4) 6	(5 / 11) 10	(11 / 11) 10	(8 / 11) 4	(103 / 114) 67	(10 / 11) 6
NF2	(5 / 11) 7	(11 / 11) 3	(10 / 11) 11	(2 / 10) 7	(9 / 1) 11	(110 / 92) 72	(11 / 9) 8
MA	(10 / 9) 3	(4 / 3) 2	(11 / 9) 2	(5 / 9) 11	(1 / 10) 3	(61 / 84) 49	(4 / 8) 1
S.ES	(6 / 8) 4	(3 / 10) 4	(9 / 10) 3	(4 / 4) 9	(3 / 9) 5	(71 / 92) 50	(9 / 9) 2
DT.ES	(3 / 6) 2	(9 / 5) 7	(4 / 4) 4	(3 / 8) 1	(7 / 6) 6	(49 / 76) 69	(2 / 7) 7
EXP.ES	(8 / 4) 8	(5 / 5) 7	(6 / 5) 4	(6 / 5) 1	(4 / 2) 6	(67 / 60) 81	(7 / 5) 10
EXP.ARIMA	(7 / 7) 11	(5 / 5) 7	(3 / 8) 9	(1 / 3) 5	(2 / 2) 6	(63 / 65) 93	(6 / 6) 11
EXP.ES&ARIMA	(8 / 4) 8	(5 / 5) 7	(6 / 5) 4	(6 / 5) 1	(4 / 2) 6	(70 / 59) 76	(8 / 4) 9
mv.EXP.ES	(1 / 2) 1	(5 / 5) 7	(6 / 5) 4	(6 / 5) 1	(4 / 2) 6	(48 / 48) 53	(1 / 3) 3
MLP	(4 / 1) 8	(1 / 1) 1	(2 / 3) 8	(10 / 1) 6	(11 / 7) 1	(61 / 24) 53	(4 / 1) 3
mv.MLP	(2 / 3) 6	(2 / 2) 5	(1 / 2) 1	(9 / 2) 8	(10 / 8) 2	(56 / 34) 53	(3 / 2) 3

Tabelle 8.9. MAE und Rang nach MAE je Zeitreihe und Verfahren an Absatzstelle AU-2

MAE	S1	S2	S3	S4	S5	S6	S7
NF1	(2,72 / 4,11) 3,10	(2,50 / 2,89) 3,33	(2,81 / 3,33) 2,62	(3,12 / 3,22) 3,52	(1,54 / 1,25) 1,10	(1,79 / 2,33) 1,52	(2,63 / 3,33) 2,14
NF2	(2,84 / 4,17) 3,19	(2,34 / 3,42) 2,90	(2,93 / 3,81) 2,76	(2,99 / 2,94) 2,71	(1,41 / 1,33) 1,10	(1,84 / 2,19) 1,67	(3,13 / 3,31) 2,52
MA	(2,19 / 3,28) 2,29	(2,09 / 2,36) 2,24	(2,50 / 2,97) 2,24	(2,74 / 2,58) 2,48	(1,40 / 1,11) 1,05	(1,65 / 1,78) 1,43	(2,41 / 2,86) 2,05
S.ES	(2,15 / 3,25) 2,05	(1,96 / 2,36) 2,33	(2,35 / 2,81) 1,95	(2,49 / 2,36) 2,43	(1,26 / 1,00) 1,05	(1,60 / 1,89) 1,43	(2,25 / 2,78) 1,76
DT.ES	(1,99 / 2,75) 1,86	(1,78 / 2,11) 2,05	(2,07 / 2,83) 1,95	(2,26 / 2,14) 2,24	(1,22 / 0,97) 1,38	(1,64 / 1,61) 1,24	(2,28 / 2,64) 1,90
EXP.ES	(1,99 / 2,78) 1,86	(1,78 / 2,08) 2,05	(2,07 / 2,83) 1,95	(2,30 / 2,14) 2,29	(1,29 / 0,94) 1,38	(1,67 / 1,94) 1,29	(2,35 / 2,64) 1,90
EXP.ARIMA	(2,01 / 2,89) 2,10	(1,78 / 2,00) 1,90	(2,07 / 2,83) 1,95	(2,29 / 2,44) 2,05	(1,25 / 0,94) 1,38	(1,65 / 1,94) 1,29	(2,20 / 2,53) 1,90
EXP.ES&ARIMA	(1,99 / 2,78) 1,86	(1,78 / 2,08) 2,05	(2,07 / 2,83) 1,95	(2,30 / 2,14) 2,29	(1,29 / 0,94) 1,38	(1,67 / 1,94) 1,29	(2,35 / 2,64) 1,90
mv.EXP.ES	(2,01 / 2,47) 1,95	(1,78 / 2,08) 2,05	(1,84 / 2,44) 1,95	(2,07 / 1,83) 2,52	(1,29 / 0,94) 1,38	(1,67 / 1,94) 1,29	(2,35 / 2,64) 1,90
MLP	(2,21 / 2,86) 1,90	(1,82 / 1,86) 2,10	(2,07 / 2,53) 1,86	(2,33 / 2,22) 2,00	(1,18 / 0,86) 1,29	(1,16 / 1,64) 1,38	(1,80 / 2,53) 1,86
mv.MLP	(1,89 / 2,58) 2,14	(1,70 / 1,81) 1,81	(1,98 / 2,64) 1,86	(2,28 / 2,11) 2,05	(1,21 / 0,94) 1,38	(1,13 / 1,53) 1,29	(1,95 / 2,36) 1,90
Verfahren	S8	S9	S10	S11	S12	S13	Rank mwMAE
NF1	(2,28 / 2,31) 2,29	(2,85 / 4,39) 3,86	(1,41 / 0,67) 1,29	(1,95 / 2,81) 1,62	(2,63 / 2,14) 1,90	(3,24 / 3,19) 3,57	(10 / 10) 11
NF2	(2,35 / 3,08) 2,24	(2,81 / 4,08) 3,29	(1,51 / 0,81) 1,38	(1,75 / 2,44) 1,48	(2,56 / 2,25) 1,52	(3,29 / 3,56) 3,48	(11 / 11) 10
MA	(2,01 / 2,11) 1,76	(2,47 / 3,06) 2,62	(1,28 / 0,86) 1,48	(1,50 / 2,53) 1,62	(2,22 / 1,86) 1,52	(2,63 / 2,72) 3,14	(9 / 9) 9
S.ES	(2,00 / 2,25) 1,76	(2,38 / 3,06) 2,48	(1,25 / 0,81) 1,33	(1,48 / 2,22) 1,52	(2,07 / 1,67) 1,52	(2,37 / 2,69) 2,76	(8 / 8) 6
DT.ES	(1,80 / 2,36) 1,62	(2,16 / 2,83) 3,67	(1,14 / 0,67) 1,29	(1,76 / 1,94) 2,43	(1,99 / 1,53) 1,29	(2,41 / 2,81) 2,33	(5 / 4) 8
EXP.ES	(1,80 / 2,36) 1,62	(2,14 / 2,86) 2,38	(1,09 / 0,67) 1,52	(1,46 / 2,06) 2,57	(2,07 / 1,64) 1,29	(2,41 / 2,81) 2,33	(4 / 5) 7
EXP.ARIMA	(1,80 / 2,36) 1,62	(2,38 / 2,94) 2,48	(1,28 / 0,72) 1,29	(1,34 / 1,97) 1,76	(1,97 / 1,81) 1,90	(2,55 / 2,58) 2,71	(6 / 7) 5
EXP.ES&ARIMA	(1,80 / 2,36) 1,62	(2,14 / 2,86) 2,38	(1,32 / 0,72) 1,29	(1,44 / 2,14) 1,81	(2,07 / 1,64) 1,29	(2,41 / 2,81) 2,33	(7 / 6) 3
mv.EXP.ES	(1,80 / 2,36) 1,62	(1,90 / 2,64) 2,38	(1,32 / 0,72) 1,29	(1,44 / 2,14) 1,81	(2,07 / 1,64) 1,29	(2,09 / 2,47) 2,33	(3 / 3) 4
MLP	(1,75 / 2,22) 1,57	(2,48 / 2,53) 2,38	(0,75 / 0,72) 1,33	(1,82 / 2,06) 1,48	(1,56 / 1,39) 1,38	(1,93 / 2,39) 2,19	(2 / 2) 1
mv.MLP	(1,69 / 1,97) 1,52	(2,28 / 2,61) 2,48	(0,74 / 0,72) 1,24	(1,57 / 1,89) 1,33	(1,46 / 1,50) 1,57	(2,10 / 2,28) 2,52	(1 / 1) 2
Rang des MAE	S1	S2	S3	S4	S5	S6	S7
NF1	(10 / 10) 10	(11 / 10) 11	(10 / 10) 10	(11 / 11) 11	(11 / 10) 3	(10 / 11) 10	(10 / 11) 10
NF2	(11 / 11) 11	(10 / 11) 10	(11 / 11) 11	(10 / 10) 10	(10 / 11) 3	(11 / 10) 11	(11 / 10) 11
MA	(8 / 9) 9	(9 / 8) 8	(9 / 9) 9	(9 / 9) 8	(9 / 9) 1	(5 / 4) 8	(9 / 9) 9
S.ES	(7 / 8) 6	(8 / 8) 9	(8 / 4) 3	(8 / 7) 7	(5 / 8) 1	(3 / 5) 8	(4 / 8) 1
DT.ES	(2 / 3) 1	(2 / 7) 3	(4 / 5) 3	(2 / 3) 4	(3 / 7) 6	(4 / 2) 1	(5 / 4) 3
EXP.ES	(2 / 4) 1	(2 / 4) 3	(4 / 5) 3	(5 / 3) 5	(6 / 2) 6	(7 / 6) 2	(6 / 4) 3
EXP.ARIMA	(5 / 7) 7	(2 / 3) 2	(4 / 5) 3	(4 / 8) 2	(4 / 2) 6	(6 / 6) 2	(3 / 2) 3
EXP.ES&ARIMA	(2 / 4) 1	(2 / 4) 3	(4 / 5) 3	(5 / 3) 5	(6 / 2) 6	(7 / 6) 2	(6 / 4) 3
mv.EXP.ES	(5 / 1) 5	(2 / 4) 3	(1 / 1) 3	(1 / 1) 9	(6 / 2) 6	(7 / 6) 2	(6 / 4) 3
MLP	(9 / 6) 4	(7 / 2) 7	(3 / 2) 1	(7 / 6) 1	(1 / 1) 5	(2 / 3) 7	(1 / 2) 2
mv.MLP	(1 / 2) 8	(1 / 1) 1	(2 / 3) 1	(3 / 2) 2	(2 / 2) 6	(1 / 1) 2	(2 / 1) 3
Verfahren	S8	S9	S10	S11	S12	S13	Rank of SoR MAE
NF1	(10 / 5) 11	(11 / 11) 11	(10 / 1) 2	(11 / 11) 5	(11 / 10) 10	(10 / 10) 11	(11 / 10) 11
NF2	(11 / 11) 10	(10 / 10) 9	(11 / 9) 9	(8 / 9) 2	(10 / 11) 6	(11 / 11) 10	(10 / 11) 10
MA	(9 / 2) 8	(8 / 8) 8	(7 / 11) 10	(6 / 10) 5	(9 / 9) 6	(9 / 6) 9	(9 / 9) 9
S.ES	(8 / 4) 8	(7 / 8) 5	(5 / 9) 7	(5 / 8) 4	(8 / 7) 6	(4 / 5) 8	(8 / 8) 8
DT.ES	(3 / 6) 3	(4 / 4) 10	(4 / 1) 2	(9 / 2) 10	(4 / 3) 1	(5 / 7) 2	(3 / 4) 5
EXP.ES	(3 / 6) 3	(2 / 5) 1	(3 / 1) 11	(4 / 4) 11	(5 / 4) 1	(5 / 7) 2	(4 / 5) 6
EXP.ARIMA	(3 / 6) 3	(6 / 7) 5	(6 / 4) 2	(1 / 3) 7	(3 / 8) 10	(8 / 4) 7	(5 / 7) 7
EXP.ES&ARIMA	(3 / 6) 3	(2 / 5) 1	(8 / 4) 2	(2 / 6) 8	(5 / 4) 1	(5 / 7) 2	(7 / 6) 1
mv.EXP.ES	(3 / 6) 3	(1 / 3) 1	(8 / 4) 2	(2 / 6) 8	(5 / 4) 1	(2 / 3) 2	(2 / 3) 4
MLP	(2 / 3) 2	(9 / 1) 1	(2 / 4) 7	(10 / 4) 2	(2 / 1) 5	(1 / 2) 1	(6 / 2) 2
mv.MLP	(1 / 1) 1	(5 / 2) 5	(1 / 4) 1	(7 / 1) 1	(1 / 2) 9	(3 / 1) 6	(1 / 1) 3

Tabelle 8.10. MAPE und Rang nach MAPE je Zeitreihe und Verfahren an Absatzstelle AU-2

MAPE je Zeitreihe	S1	S2	S3	S4	S5	S6	S7
NF1	(68,4% / 71,6%) 47,1%	(82,1% / 99,3%) 137,2%	(60,6% / 118,0%) 60,2%	(54,5% / 42,4%) 54,5%	(102,1% / 84,5%) 104,0%	(83,6% / 125,5%) 71,8%	(85,4% / 80,7%) 51,1%
NF2	(75,2% / 88,1%) 45,9%	(76,2% / 119,3%) 111,4%	(60,0% / 99,9%) 50,1%	(51,4% / 40,9%) 41,3%	(77,6% / 89,4%) 101,6%	(89,9% / 118,4%) 91,0%	(78,1% / 76,4%) 55,4%
MA	(60,9% / 60,5%) 37,9%	(77,5% / 79,2%) 101,8%	(58,6% / 86,7%) 53,4%	(49,2% / 36,6%) 40,2%	(97,2% / 84,1%) 109,5%	(82,9% / 102,5%) 77,9%	(76,6% / 81,6%) 50,9%
S.ES	(63,2% / 66,1%) 34,3%	(73,9% / 85,8%) 107,2%	(55,5% / 85,8%) 47,4%	(45,5% / 33,5%) 40,3%	(91,3% / 78,5%) 117,5%	(85,7% / 108,4%) 81,0%	(73,8% / 81,6%) 47,1%
DT.ES	(62,1% / 64,9%) 28,9%	(68,5% / 79,6%) 105,5%	(50,8% / 83,2%) 49,3%	(42,8% / 32,0%) 40,2%	(88,3% / 74,9%) 128,6%	(90,8% / 95,0%) 67,6%	(79,9% / 78,4%) 53,2%
EXP.ES	(62,1% / 65,6%) 28,9%	(68,5% / 73,7%) 105,5%	(50,8% / 83,2%) 49,3%	(43,8% / 32,1%) 41,9%	(98,4% / 78,5%) 128,6%	(95,0% / 117,5%) 89,5%	(80,8% / 78,2%) 53,2%
EXP.ARIMA	(64,6% / 61,6%) 28,1%	(68,5% / 69,8%) 90,1%	(50,8% / 83,2%) 49,3%	(44,3% / 33,2%) 30,6%	(91,9% / 78,5%) 128,6%	(94,6% / 117,5%) 89,5%	(75,9% / 73,9%) 53,2%
EXP.ES&ARIMA	(62,1% / 65,6%) 28,9%	(68,5% / 73,7%) 105,5%	(50,8% / 83,2%) 49,3%	(43,8% / 32,1%) 41,9%	(98,4% / 78,5%) 128,6%	(95,0% / 117,5%) 89,5%	(80,8% / 78,2%) 53,2%
mv.EXP.ES	(61,6% / 54,6%) 27,7%	(68,5% / 73,7%) 105,5%	(43,1% / 72,1%) 49,3%	(37,7% / 33,9%) 45,6%	(98,4% / 78,5%) 128,6%	(95,0% / 117,5%) 89,5%	(80,8% / 78,2%) 53,2%
MLP	(79,1% / 69,1%) 28,9%	(72,8% / 66,6%) 97,7%	(44,6% / 69,7%) 39,5%	(46,9% / 32,5%) 35,8%	(93,2% / 71,1%) 123,8%	(66,7% / 97,8%) 77,1%	(62,1% / 70,5%) 45,6%
mv.MLP	(57,2% / 52,0%) 34,4%	(68,9% / 65,2%) 91,5%	(41,7% / 73,5%) 41,0%	(44,3% / 33,0%) 36,8%	(94,3% / 78,5%) 128,6%	(67,1% / 90,0%) 69,2%	(68,9% / 65,4%) 52,4%
Verfahren	S8	S9	S10	S11	S12	S13	Rank mwMAPE
NF1	(79,7% / 64,8%) 106,5%	(69,3% / 75,0%) 57,0%	(97,8% / 67,4%) 87,4%	(76,5% / 73,6%) 39,6%	(104,1% / 84,3%) 95,0%	(83,6% / 63,6%) 68,4%	(11, / 10) 11
NF2	(86,6% / 93,7%) 104,7%	(65,7% / 71,4%) 53,1%	(95,6% / 79,9%) 81,9%	(65,6% / 84,1%) 42,8%	(101,9% / 96,7%) 89,8%	(80,6% / 82,6%) 67,1%	(10, / 11) 9
MA	(78,8% / 61,0%) 89,2%	(59,1% / 54,0%) 41,2%	(117,7% / 97,9%) 111,6%	(67,1% / 74,1%) 51,1%	(89,9% / 86,5%) 75,1%	(68,4% / 55,9%) 62,1%	(9, / 5) 4
S.ES	(82,8% / 73,8%) 94,8%	(60,5% / 55,8%) 41,0%	(120,3% / 113,2%) 99,5%	(68,4% / 69,3%) 50,6%	(89,5% / 79,4%) 84,4%	(67,3% / 59,5%) 56,7%	(6, / 9) 3
DT.ES	(78,7% / 81,5%) 91,1%	(55,4% / 52,2%) 62,4%	(108,3% / 110,4%) 87,5%	(87,3% / 65,0%) 96,8%	(86,9% / 69,0%) 69,3%	(70,2% / 63,1%) 44,2%	(4, / 4) 8
EXP.ES	(78,7% / 81,5%) 91,1%	(57,9% / 50,7%) 39,1%	(108,0% / 110,4%) 96,9%	(70,9% / 66,7%) 100,4%	(93,7% / 76,8%) 69,3%	(70,2% / 63,1%) 44,2%	(7, / 7) 10
EXP.ARIMA	(78,7% / 87,5%) 91,1%	(66,5% / 51,2%) 31,7%	(119,2% / 110,4%) 87,5%	(64,3% / 68,5%) 73,8%	(84,6% / 90,1%) 108,3%	(71,2% / 60,5%) 59,5%	(5, / 8) 7
EXP.ES&ARIMA	(78,7% / 81,5%) 91,1%	(57,9% / 50,7%) 39,1%	(120,1% / 110,4%) 87,5%	(63,1% / 66,3%) 74,5%	(93,7% / 76,8%) 69,3%	(70,2% / 63,1%) 44,2%	(8, / 6) 5
mv.EXP.ES	(78,7% / 81,5%) 91,1%	(36,8% / 48,2%) 39,1%	(120,1% / 110,4%) 87,5%	(63,1% / 66,3%) 74,5%	(93,7% / 76,8%) 69,3%	(58,7% / 48,8%) 44,2%	(2, / 3) 6
MLP	(72,8% / 77,2%) 87,4%	(73,5% / 46,5%) 35,1%	(113,7% / 110,4%) 96,7%	(95,7% / 69,6%) 56,8%	(69,0% / 64,3%) 78,8%	(55,8% / 51,6%) 40,7%	(3, / 2) 2
mv.MLP	(65,7% / 68,0%) 77,6%	(61,4% / 50,6%) 36,5%	(109,8% / 110,4%) 85,2%	(86,6% / 66,4%) 55,4%	(66,4% / 70,6%) 86,5%	(63,9% / 53,7%) 48,6%	(1, / 1) 1
Rang des MAPE	S1	S2	S3	S4	S5	S6	S7
NF1	(9 / 10) 11	(11 / 10) 11	(11 / 11) 11	(11 / 11) 11	(11 / 10) 2	(4 / 11) 3	(11 / 9) 4
NF2	(10 / 11) 10	(9 / 11) 10	(10 / 10) 9	(10 / 10) 7	(1 / 11) 1	(6 / 10) 11	(6 / 4) 11
MA	(2 / 3) 9	(10 / 7) 4	(9 / 9) 10	(9 / 9) 5	(7 / 9) 3	(3 / 4) 5	(5 / 11) 3
S.ES	(7 / 8) 7	(8 / 9) 9	(8 / 8) 3	(7 / 7) 6	(3 / 3) 4	(5 / 5) 6	(3 / 10) 2
DT.ES	(4 / 5) 3	(1 / 8) 5	(4 / 4) 4	(2 / 1) 4	(2 / 2) 6	(7 / 2) 1	(7 / 8) 6
EXP.ES	(4 / 6) 3	(1 / 4) 5	(4 / 4) 4	(3 / 2) 8	(8 / 3) 6	(9 / 6) 7	(8 / 5) 6
EXP.ARIMA	(8 / 4) 2	(1 / 3) 1	(4 / 4) 4	(5 / 6) 1	(4 / 3) 6	(8 / 6) 7	(4 / 3) 6
EXP.ES&ARIMA	(4 / 6) 3	(1 / 4) 5	(4 / 4) 4	(3 / 2) 8	(8 / 3) 6	(9 / 6) 7	(8 / 5) 6
mv.EXP.ES	(3 / 2) 1	(1 / 4) 5	(2 / 2) 4	(1 / 8) 10	(8 / 3) 6	(9 / 6) 7	(8 / 5) 6
MLP	(11 / 9) 6	(7 / 2) 3	(3 / 1) 1	(8 / 4) 2	(5 / 1) 5	(1 / 3) 4	(1 / 2) 1
mv.MLP	(1 / 1) 8	(6 / 1) 2	(1 / 3) 2	(6 / 5) 3	(6 / 3) 6	(2 / 1) 2	(2 / 1) 5
Verfahren	S8	S9	S10	S11	S12	S13	Rank of SoR MAPE
NF1	(9 / 2) 11	(10 / 11) 10	(2 / 1) 3	(8 / 9) 1	(11 / 8) 10	(11 / 10) 11	(11 / 10) 10
NF2	(11 / 11) 10	(8 / 11) 9	(1 / 2) 1	(4 / 11) 2	(10 / 11) 9	(10 / 11) 10	(10 / 11) 11
MA	(8 / 1) 3	(5 / 8) 8	(7 / 3) 11	(5 / 10) 4	(6 / 9) 5	(5 / 4) 9	(8 / 8) 8
S.ES	(10 / 4) 9	(6 / 9) 7	(11 / 11) 10	(6 / 7) 3	(5 / 7) 7	(4 / 5) 7	(9 / 9) 9
DT.ES	(3 / 6) 4	(2 / 7) 11	(4 / 4) 4	(10 / 1) 10	(4 / 2) 1	(6 / 7) 2	(3 / 4) 3
EXP.ES	(3 / 6) 4	(3 / 4) 4	(3 / 4) 9	(7 / 5) 11	(7 / 4) 1	(6 / 7) 2	(4 / 6) 7
EXP.ARIMA	(3 / 10) 4	(9 / 6) 1	(8 / 4) 4	(3 / 6) 7	(3 / 10) 11	(9 / 6) 8	(6 / 7) 4
EXP.ES&ARIMA	(3 / 6) 4	(3 / 4) 4	(9 / 4) 4	(1 / 2) 8	(7 / 4) 1	(6 / 7) 2	(4 / 4) 4
mv.EXP.ES	(3 / 6) 4	(1 / 2) 4	(9 / 4) 4	(1 / 2) 8	(7 / 4) 1	(2 / 1) 2	(2 / 3) 4
MLP	(2 / 5) 2	(11 / 1) 2	(6 / 4) 8	(11 / 8) 6	(2 / 1) 6	(1 / 2) 1	(6 / 2) 1
mv.MLP	(1 / 3) 1	(7 / 3) 3	(5 / 4) 2	(9 / 4) 5	(1 / 3) 8	(3 / 3) 6	(1 / 1) 2

Tabelle 8.11. MAE und Rang nach MAE je Zeitreihe und Verfahren an Absatzstelle IN-1

MAE	S1	S2	S3	S4	S5	S6
NF1	(1,02 / 1,78) 1,10	(50,66 / 31,11) 36,81	(5,18 / 2,92) 3,13	(6,81 / 2,06) 1,76	(23,78 / 11,31) 8,90	(5,50 / 5,17) 6,67
NF2	(0,70 / 1,50) 1,14	(47,31 / 23,64) 38,71	(4,57 / 2,89) 3,73	(5,46 / 2,33) 1,90	(20,72 / 10,44) 7,71	(5,21 / 6,19) 7,14
MA	(0,87 / 1,58) 1,05	(37,93 / 23,11) 24,71	(3,81 / 2,28) 2,60	(4,76 / 1,61) 1,14	(16,94 / 8,92) 5,76	(4,01 / 5,00) 4,29
S.ES	(0,98 / 1,50) 0,95	(34,34 / 21,03) 23,00	(3,75 / 2,22) 2,93	(4,69 / 1,72) 1,24	(15,56 / 8,64) 5,76	(3,87 / 4,89) 4,05
DT.ES	(0,91 / 1,17) 0,81	(31,33 / 18,39) 31,86	(3,30 / 2,33) 2,53	(4,62 / 1,61) 1,43	(14,10 / 7,86) 5,48	(3,61 / 5,19) 4,14
EXP.ES	(0,91 / 1,17) 0,81	(30,87 / 24,22) 28,19	(3,46 / 2,53) 2,47	(3,89 / 2,03) 1,14	(15,45 / 8,58) 7,33	(3,59 / 4,86) 4,29
EXP.ARIMA	(0,98 / 1,28) 0,86	(30,17 / 21,39) 17,62	(3,30 / 2,92) 3,13	(3,65 / 2,08) 1,33	(14,26 / 8,94) 6,05	(3,71 / 4,81) 4,62
EXP.ES&ARIMA	(0,91 / 1,17) 0,81	(31,07 / 22,58) 26,33	(3,48 / 2,56) 2,47	(4,37 / 2,06) 1,43	(15,35 / 8,89) 6,14	(3,70 / 4,97) 4,19
mv.EXP.ES	(0,91 / 1,17) 0,81	(31,07 / 22,58) 26,33	(3,48 / 2,56) 2,47	(4,37 / 2,06) 1,43	(15,35 / 8,89) 6,14	(3,70 / 4,97) 4,19
MLP	(0,90 / 1,14) 1,14	(34,87 / 16,03) 21,90	(3,92 / 2,14) 2,87	(2,82 / 1,72) 1,00	(8,66 / 7,86) 6,90	(3,70 / 4,58) 4,76
mv.MLP	(0,82 / 1,08) 1,00	(27,49 / 13,42) 18,19	(3,11 / 1,97) 2,33	(2,78 / 1,53) 1,33	(6,59 / 6,44) 5,48	(2,43 / 3,89) 4,38
Verfahren	S7	S8	S9	S10	mwMAE	Rank mwMAE
NF1	(7,03 / 11,92) 9,00	(6,84 / 3,11) 5,14	(6,59 / 3,94) 2,29	(8,82 / 8,47) 11,43	(12,22 / 8,18) 8,62	(11 / 11) 10
NF2	(6,63 / 10,97) 8,81	(5,93 / 3,08) 5,95	(6,26 / 3,94) 3,29	(8,53 / 7,53) 10,71	(11,13 / 7,25) 8,91	(10 / 10) 11
MA	(5,49 / 9,97) 6,86	(5,18 / 2,44) 5,05	(5,38 / 3,03) 1,76	(5,99 / 6,17) 7,67	(9,03 / 6,41) 6,09	(9 / 8) 5
S.ES	(5,29 / 10,00) 6,14	(5,22 / 3,03) 4,90	(4,97 / 3,06) 2,14	(5,75 / 5,81) 7,19	(8,44 / 6,19) 5,83	(8 / 5) 4
DT.ES	(4,93 / 10,06) 5,95	(5,01 / 2,58) 6,29	(4,91 / 3,11) 5,43	(5,00 / 5,14) 6,38	(7,77 / 5,74) 7,03	(4 / 3) 9
EXP.ES	(4,93 / 10,08) 5,95	(5,22 / 3,14) 6,86	(4,96 / 3,17) 4,52	(5,14 / 5,22) 5,95	(7,84 / 6,50) 6,75	(5 / 9) 8
EXP.ARIMA	(5,16 / 9,67) 5,95	(4,62 / 2,17) 5,71	(5,00 / 3,17) 4,52	(4,55 / 4,42) 6,24	(7,54 / 6,08) 5,60	(3 / 4) 2
EXP.ES&ARIMA	(4,96 / 10,14) 5,95	(5,22 / 3,14) 6,86	(5,20 / 3,36) 5,43	(5,06 / 5,06) 6,38	(7,93 / 6,39) 6,60	(7 / 7) 7
mv.EXP.ES	(4,96 / 10,14) 5,95	(4,41 / 2,56) 6,00	(5,20 / 3,36) 5,43	(5,06 / 5,06) 6,38	(7,85 / 6,33) 6,51	(6 / 6) 6
MLP	(6,34 / 9,67) 5,95	(4,15 / 3,17) 3,81	(3,39 / 3,14) 3,86	(4,36 / 4,44) 5,90	(7,31 / 5,39) 5,81	(2 / 2) 3
mv.MLP	(5,41 / 7,58) 4,71	(3,64 / 3,92) 3,71	(2,41 / 2,58) 4,52	(3,41 / 3,33) 4,76	(5,81 / 4,58) 5,04	(1 / 1) 1
Rang des MAE	S1	S2	S3	S4	S5	S6
NF1	(11 / 11) 9	(11 / 11) 10	(11 / 10) 9	(11 / 7) 10	(11 / 11) 11	(11 / 9) 10
NF2	(1 / 8) 10	(10 / 9) 11	(10 / 9) 11	(10 / 11) 11	(10 / 10) 10	(10 / 11) 11
MA	(3 / 10) 8	(9 / 8) 5	(8 / 4) 6	(9 / 2) 2	(9 / 8) 3	(9 / 8) 5
S.ES	(9 / 8) 6	(7 / 4) 4	(7 / 3) 8	(8 / 4) 4	(8 / 5) 3	(8 / 5) 1
DT.ES	(5 / 3) 1	(6 / 3) 9	(2 / 5) 5	(7 / 2) 7	(3 / 2) 1	(3 / 10) 2
EXP.ES	(5 / 3) 1	(3 / 10) 8	(4 / 6) 2	(4 / 6) 2	(7 / 4) 9	(2 / 4) 5
EXP.ARIMA	(10 / 7) 5	(2 / 5) 1	(2 / 10) 9	(3 / 10) 5	(4 / 9) 5	(7 / 3) 8
EXP.ES&ARIMA	(5 / 3) 1	(4 / 6) 6	(5 / 7) 2	(5 / 7) 7	(5 / 6) 6	(4 / 6) 3
mv.EXP.ES	(5 / 3) 1	(4 / 6) 6	(5 / 7) 2	(5 / 7) 7	(5 / 6) 6	(4 / 6) 3
MLP	(4 / 2) 10	(8 / 2) 3	(9 / 2) 7	(2 / 4) 1	(2 / 2) 8	(6 / 2) 9
mv.MLP	(2 / 1) 7	(1 / 1) 2	(1 / 1) 1	(1 / 1) 5	(1 / 1) 1	(1 / 1) 7
Verfahren	S7	S8	S9	S10	Sum of Ranks MAE	Rank of SoR MAE
NF1	(11 / 11) 11	(11 / 7) 5	(11 / 10) 3	(11 / 11) 11	(110 / 98) 89	(11 / 11) 10
NF2	(10 / 10) 10	(10 / 6) 7	(10 / 10) 4	(10 / 10) 10	(91 / 94) 95	(10 / 10) 11
MA	(8 / 4) 9	(6 / 2) 4	(9 / 2) 1	(9 / 9) 9	(79 / 57) 52	(9 / 6) 9
S.ES	(6 / 5) 8	(9 / 5) 3	(5 / 3) 2	(8 / 8) 8	(75 / 50) 47	(8 / 4) 2
DT.ES	(1 / 6) 2	(5 / 4) 9	(3 / 4) 9	(4 / 6) 5	(39 / 45) 50	(2 / 3) 6
EXP.ES	(1 / 7) 2	(7 / 8) 10	(4 / 6) 6	(7 / 7) 3	(44 / 61) 48	(3 / 8) 3
EXP.ARIMA	(5 / 2) 2	(4 / 1) 6	(6 / 6) 6	(3 / 2) 4	(46 / 55) 51	(4 / 5) 7
EXP.ES&ARIMA	(3 / 8) 2	(7 / 8) 10	(7 / 8) 9	(5 / 4) 5	(50 / 63) 51	(7 / 9) 7
mv.EXP.ES	(3 / 8) 2	(3 / 3) 8	(7 / 8) 9	(5 / 4) 5	(46 / 58) 49	(4 / 7) 4
MLP	(9 / 2) 2	(2 / 10) 2	(2 / 5) 5	(2 / 3) 2	(46 / 34) 49	(4 / 2) 4
mv.MLP	(7 / 1) 1	(1 / 11) 1	(1 / 1) 6	(1 / 1) 1	(17 / 20) 32	(1 / 1) 1

Tabelle 8.12. MAPE und Rang nach MAPE je Zeitreihe und Verfahren an Absatzstelle IN-1

Verfahren	S1	S2	S3	S4	S5	S6
NF1	(82,6% / 104,1%) 112,7%	(60,6% / 54,9%) 46,8%	(107,0%/105,5%) 101,7%	(94,0% / 103,9%) 100,8%	(69,4% / 71,6%) 79,6%	(100,3% / 69,1%) 66,7%
NF2	(56,8% / 82,7%) 104,8%	(69,0% / 51,3%) 61,6%	(104,3%/105,4%) 104,2%	(102,7%/121,4%) 120,6%	(75,4% / 82,1%) 84,6%	(102,5% / 93,3%) 84,0%
MA	(95,8% / 121,9%) 121,4%	(52,2% / 46,8%) 35,8%	(82,2% / 97,4%) 118,5%	(79,4% / 103,8%) 86,1%	(57,6% / 69,2%) 59,9%	(81,4% / 75,2%) 51,0%
S.ES	(117,9%/108,9%) 111,9%	(47,7% / 43,9%) 33,4%	(83,9% / 96,5%) 126,8%	(77,2% / 111,0%) 87,3%	(53,6% / 68,3%) 63,8%	(82,1% / 74,6%) 50,3%
DT.ES	(134,6%/102,8%) 108,7%	(47,9% / 42,4%) 49,8%	(74,0% / 118,9%) 90,6%	(91,4% / 108,7%) 113,5%	(55,7% / 64,9%) 55,2%	(77,3% / 91,3%) 53,8%
EXP.ES	(134,6%/102,8%) 108,7%	(48,8% / 54,6%) 46,0%	(80,2% / 124,8%) 92,6%	(77,7% / 120,6%) 97,2%	(65,3% / 74,8%) 63,5%	(79,6% / 87,0%) 61,8%
EXP.ARIMA	(140,4%/108,7%) 108,7%	(52,1% / 54,2%) 32,7%	(71,5% / 151,6%) 118,8%	(75,3% / 125,0%) 91,3%	(63,1% / 85,0%) 74,1%	(84,2% / 83,1%) 53,8%
EXP.ES&ARIMA	(134,6%/102,8%) 108,7%	(47,4% / 50,0%) 41,5%	(82,1% / 131,3%) 92,6%	(82,7% / 126,8%) 113,5%	(59,3% / 77,7%) 77,2%	(79,9% / 86,7%) 57,2%
mv.EXP.ES	(134,6%/102,8%) 108,7%	(47,4% / 50,0%) 41,5%	(82,1% / 131,3%) 92,6%	(82,7% / 126,8%) 113,5%	(59,3% / 77,7%) 77,2%	(79,9% / 86,7%) 57,2%
MLP	(132,3% / 99,3%) 120,6%	(47,0% / 42,6%) 24,6%	(79,0% / 86,0%) 94,4%	(64,9% / 118,3%) 84,9%	(44,9% / 66,2%) 85,9%	(76,8% / 79,1%) 55,3%
mv.MLP	(109,6% / 94,8%) 92,9%	(55,5% / 48,0%) 38,5%	(71,9% / 114,8%) 98,3%	(43,0% / 100,3%) 101,2%	(27,1% / 45,6%) 60,9%	(74,4% / 94,2%) 64,7%
Verfahren	S7	S8	S9	S10	mwMAPE	Rank mwMAPE
NF1	(91,4% / 87,9%) 70,8%	(121,3% / 44,1%) 95,4%	(82,5% / 105,0%) 109,1%	(71,9% / 68,8%) 105,6%	(88,1% / 81,5%) 88,9%	(10, / 1) 4
NF2	(95,3% / 138,6%) 75,7%	(120,8% / 53,4%) 138,3%	(113,6%/113,3%) 119,8%	(78,8% / 75,7%) 125,3%	(91,9% / 91,7%) 101,9%	(11, / 8) 8
MA	(86,1% / 106,0%) 55,6%	(117,1% / 55,4%) 113,9%	(75,4% / 88,4%) 107,2%	(58,0% / 60,5%) 84,6%	(78,5% / 82,5%) 83,4%	(3, / 3) 1
S.ES	(82,5% / 108,9%) 52,4%	(114,5%/133,5%) 116,2%	(74,1% / 89,8%) 121,5%	(55,5% / 57,9%) 78,4%	(78,9% / 89,3%) 84,2%	(4, / 5) 2
DT.ES	(79,3% / 113,6%) 59,1%	(114,4% / 82,9%) 166,8%	(78,3% / 90,0%) 270,8%	(53,0% / 54,7%) 70,1%	(80,6% / 87,0%) 103,9%	(5, / 4) 9
EXP.ES	(79,3% / 113,8%) 59,1%	(118,8%/134,4%) 181,8%	(78,0% / 95,3%) 235,5%	(54,6% / 59,5%) 64,8%	(81,7% / 96,8%) 101,1%	(6, / 10) 7
EXP.ARIMA	(81,1% / 120,4%) 58,1%	(117,1% / 72,1%) 145,8%	(78,4% / 95,3%) 235,5%	(56,9% / 53,2%) 69,4%	(82,0% / 94,8%) 98,8%	(8, / 9) 6
EXP.ES&ARIMA	(80,5% / 114,5%) 63,4%	(118,8%/134,4%) 181,8%	(82,2% / 109,1%) 270,8%	(54,1% / 54,0%) 70,1%	(82,2% / 98,7%) 107,7%	(9, / 11) 10
mv.EXP.ES	(80,5% / 114,5%) 63,4%	(115,6% / 60,8%) 181,8%	(82,2% / 109,1%) 270,8%	(54,1% / 54,0%) 70,1%	(81,8% / 91,4%) 107,7%	(7, / 7) 10
MLP	(103,5% / 123,2%) 57,6%	(88,0% / 44,0%) 72,5%	(68,8% / 99,5%) 200,2%	(59,4% / 57,1%) 67,3%	(76,5% / 81,5%) 86,3%	(2, / 2) 3
mv.MLP	(112,6% / 135,1%) 60,1%	(107,1% / 134,9%) 84,0%	(41,5% / 102,0%) 259,4%	(30,4% / 34,1%) 44,6%	(67,3% / 90,4%) 90,5%	(1, / 6) 5
Rang des MAE	S1	S2	S3	S4	S5	S6
NF1	(2 / 8) 9	(10 / 11) 9	(11 / 5) 7	(10 / 3) 6	(10 / 6) 9	(10 / 1) 10
NF2	(1 / 1) 2	(11 / 8) 11	(10 / 4) 8	(11 / 8) 11	(11 / 10) 10	(11 / 10) 11
MA	(3 / 11) 11	(8 / 4) 4	(8 / 3) 9	(6 / 2) 2	(5 / 5) 2	(7 / 3) 2
S.ES	(5 / 10) 8	(4 / 3) 3	(9 / 2) 11	(4 / 5) 3	(3 / 4) 5	(8 / 2) 1
DT.ES	(7 / 4) 3	(5 / 1) 10	(3 / 7) 1	(9 / 4) 8	(4 / 2) 1	(3 / 9) 4
EXP.ES	(7 / 4) 3	(6 / 10) 8	(5 / 8) 2	(5 / 7) 5	(9 / 7) 4	(4 / 8) 8
EXP.ARIMA	(11 / 9) 3	(7 / 9) 2	(1 / 11) 10	(3 / 9) 4	(8 / 11) 6	(9 / 5) 3
EXP.ES&ARIMA	(7 / 4) 3	(2 / 6) 6	(6 / 9) 2	(7 / 10) 8	(6 / 8) 7	(5 / 6) 6
mv.EXP.ES	(7 / 4) 3	(2 / 6) 6	(6 / 9) 2	(7 / 10) 8	(6 / 8) 7	(5 / 6) 6
MLP	(6 / 3) 10	(1 / 2) 1	(4 / 1) 5	(2 / 6) 1	(2 / 3) 11	(2 / 4) 5
mv.MLP	(4 / 2) 1	(9 / 5) 5	(2 / 6) 6	(1 / 1) 7	(1 / 1) 3	(1 / 11) 9
Verfahren	S7	S8	S9	S10	Sum of Ranks MAPE	Rank of SoR MAPE
NF1	(8 / 1) 10	(11 / 2) 3	(10 / 8) 2	(10 / 10) 10	(92 / 55) 75	(10 / 5) 10
NF2	(9 / 11) 11	(10 / 3) 6	(11 / 11) 3	(11 / 11) 11	(96 / 77) 84	(11 / 11) 11
MA	(7 / 2) 2	(7 / 4) 4	(4 / 1) 1	(8 / 9) 9	(63 / 44) 46	(8 / 2) 2
S.ES	(6 / 3) 1	(4 / 8) 5	(3 / 2) 4	(6 / 7) 8	(52 / 46) 49	(4 / 3) 3
DT.ES	(1 / 4) 5	(3 / 7) 8	(6 / 3) 9	(2 / 5) 5	(43 / 46) 54	(3 / 3) 7
EXP.ES	(1 / 5) 5	(8 / 9) 9	(5 / 4) 6	(5 / 8) 2	(55 / 70) 52	(6 / 8) 6
EXP.ARIMA	(5 / 8) 4	(6 / 6) 7	(7 / 4) 6	(7 / 2) 4	(64 / 74) 49	(9 / 10) 3
EXP.ES&ARIMA	(3 / 6) 8	(8 / 9) 9	(8 / 9) 9	(3 / 3) 5	(55 / 70) 63	(6 / 8) 8
mv.EXP.ES	(3 / 6) 8	(5 / 5) 9	(8 / 9) 9	(3 / 3) 5	(52 / 66) 63	(4 / 7) 8
MLP	(10 / 9) 3	(1 / 1) 1	(2 / 6) 5	(9 / 6) 3	(39 / 41) 45	(2 / 1) 1
mv.MLP	(11 / 10) 7	(2 / 11) 2	(1 / 7) 8	(1 / 1) 1	(33 / 55) 49	(1 / 5) 3

Tabelle 8.13. MAE und Rang nach MAE je Zeitreihe und Verfahren an Absatzstelle IN-2

MAE	S1	S2	S3	S4	S5	S6
NF1	(1,91 / 2,67) 2,76	(16,96 / 19,47) 9,86	(3,01 / 3,33) 2,29	(2,93 / 1,86) 2,07	(12,78 / 9,53) 8,29	(11,12 / 10,22) 11,86
NF2	(2,40 / 3,06) 2,43	(17,85 / 18,42) 11,86	(3,44 / 2,69) 1,86	(2,76 / 2,22) 1,67	(12,63 / 11,00) 8,52	(11,40 / 12,22) 8,52
MA	(2,00 / 2,33) 2,38	(13,82 / 15,47) 11,10	(2,69 / 2,56) 2,19	(2,31 / 1,64) 1,67	(10,38 / 8,39) 7,38	(8,75 / 8,81) 9,86
S.ES	(1,91 / 2,19) 2,19	(13,28 / 14,81) 9,57	(2,62 / 2,33) 2,05	(2,18 / 1,58) 1,53	(9,60 / 8,17) 6,81	(8,79 / 8,31) 9,00
DT.ES	(1,94 / 2,17) 1,81	(13,03 / 15,00) 9,24	(2,48 / 2,06) 2,29	(2,12 / 1,64) 1,87	(8,32 / 8,28) 6,57	(8,33 / 8,19) 10,57
EXP.ES	(1,97 / 2,14) 1,81	(13,00 / 15,08) 9,24	(2,48 / 2,06) 2,29	(2,12 / 1,64) 1,87	(8,36 / 8,47) 5,95	(8,23 / 8,11) 11,14
EXP.ARIMA	(1,80 / 2,08) 1,76	(12,80 / 14,17) 9,38	(2,43 / 2,06) 2,29	(1,96 / 1,86) 1,87	(8,83 / 8,89) 6,48	(9,39 / 8,06) 11,57
EXP.ES&ARIMA	(1,97 / 2,14) 1,81	(13,00 / 15,08) 9,24	(2,48 / 2,06) 2,29	(1,96 / 1,86) 1,87	(8,83 / 8,89) 6,48	(9,39 / 8,06) 11,57
mv.EXP.ES	(1,97 / 2,14) 1,81	(9,59 / 14,11) 9,71	(2,48 / 2,06) 2,29	(1,75 / 1,75) 1,87	(6,35 / 7,92) 6,33	(6,71 / 7,50) 10,86
MLP	(1,79 / 1,89) 2,24	(14,64 / 13,64) 9,57	(2,72 / 2,00) 2,24	(1,67 / 1,42) 1,47	(10,15 / 7,97) 6,24	(9,82 / 7,67) 9,76
mv.MLP	(1,79 / 1,94) 2,10	(9,79 / 13,39) 8,19	(2,77 / 2,03) 2,24	(1,48 / 1,67) 1,67	(6,43 / 7,75) 6,05	(6,20 / 7,75) 8,43
Verfahren	S7	S8	S9	mwMAE		Rank mwMAE
NF1	(13,22 / 7,92) 7,48	(1,90 / 1,39) 1,24	(3,43 / 8,00) 5,14	(7,47 / 7,15) 5,66		(10 / 10) 9
NF2	(14,47 / 10,56) 8,90	(2,29 / 1,19) 0,94	(3,56 / 3,53) 4,90	(7,87 / 7,21) 5,51		(11 / 11) 7
MA	(11,74 / 7,97) 7,76	(1,72 / 1,11) 0,76	(2,67 / 2,83) 3,86	(6,23 / 5,68) 5,22		(9 / 9) 5
S.ES	(10,96 / 6,86) 6,38	(1,75 / 1,14) 0,82	(2,67 / 2,86) 3,67	(5,97 / 5,36) 4,67		(8 / 7) 2
DT.ES	(10,26 / 6,14) 14,52	(1,68 / 0,92) 1,12	(2,70 / 2,75) 3,48	(5,65 / 5,24) 5,72		(4 / 4) 10
EXP.ES	(10,23 / 6,14) 14,48	(1,80 / 1,03) 0,94	(2,59 / 2,89) 3,24	(5,64 / 5,28) 5,66		(3 / 5) 8
EXP.ARIMA	(10,91 / 7,89) 7,48	(1,67 / 1,33) 1,35	(2,60 / 2,89) 3,24	(5,82 / 5,47) 5,05		(6 / 8) 3
EXP.ES&ARIMA	(10,23 / 6,14) 14,48	(1,80 / 1,03) 0,94	(2,59 / 2,89) 3,24	(5,80 / 5,35) 5,77		(5 / 6) 11
mv.EXP.ES	(7,99 / 6,28) 12,10	(1,80 / 1,03) 0,94	(2,49 / 2,78) 3,29	(4,57 / 5,06) 5,47		(2 / 3) 6
MLP	(8,30 / 5,75) 8,52	(1,16 / 1,03) 1,35	(2,25 / 2,81) 4,29	(5,83 / 4,91) 5,08		(7 / 1) 4
mv.MLP	(6,97 / 5,83) 8,10	(1,03 / 1,03) 1,00	(2,47 / 2,78) 3,76	(4,32 / 4,91) 4,61		(1 / 1) 1
Rang des MAE	S1	S2	S3	S4	S5	S6
NF1	(4 / 10) 11	(10 / 11) 9	(10 / 11) 6	(11 / 8) 11	(11 / 10) 10	(10 / 10) 11
NF2	(11 / 11) 10	(11 / 10) 11	(11 / 10) 1	(10 / 11) 3	(10 / 11) 11	(11 / 11) 2
MA	(10 / 9) 9	(8 / 9) 10	(7 / 9) 3	(9 / 3) 3	(9 / 6) 9	(5 / 9) 5
S.ES	(4 / 8) 7	(7 / 5) 6	(6 / 8) 2	(8 / 2) 2	(7 / 4) 8	(6 / 8) 3
DT.ES	(6 / 7) 2	(6 / 6) 2	(2 / 3) 6	(6 / 3) 6	(3 / 5) 7	(4 / 7) 6
EXP.ES	(7 / 4) 2	(4 / 7) 2	(2 / 3) 6	(6 / 3) 6	(4 / 7) 1	(3 / 6) 8
EXP.ARIMA	(3 / 3) 1	(3 / 4) 5	(1 / 3) 6	(4 / 8) 6	(5 / 8) 5	(7 / 4) 9
EXP.ES&ARIMA	(7 / 4) 2	(4 / 7) 2	(2 / 3) 6	(4 / 8) 6	(5 / 8) 5	(7 / 4) 9
mv.EXP.ES	(7 / 4) 2	(1 / 3) 8	(2 / 3) 6	(3 / 7) 6	(1 / 2) 4	(2 / 1) 7
MLP	(1 / 1) 8	(9 / 2) 6	(8 / 1) 4	(2 / 1) 1	(8 / 3) 3	(9 / 2) 4
mv.MLP	(1 / 2) 6	(2 / 1) 1	(9 / 2) 4	(1 / 6) 3	(2 / 1) 2	(1 / 3) 1
Verfahren	S7	S8	S9	Sum of Ranks MAE		Rank of SoR MAE
NF1	(10 / 9) 2	(10 / 11) 9	(10 / 11) 11	(86 / 91) 80		(10 / 10) 11
NF2	(11 / 11) 7	(11 / 9) 3	(11 / 10) 10	(97 / 94) 58		(11 / 11) 10
MA	(9 / 10) 4	(5 / 7) 1	(7 / 5) 8	(69 / 67) 52		(9 / 9) 8
S.ES	(8 / 7) 1	(6 / 8) 2	(7 / 6) 6	(59 / 56) 37		(8 / 8) 2
DT.ES	(6 / 3) 11	(4 / 1) 8	(9 / 1) 5	(46 / 36) 53		(7 / 4) 9
EXP.ES	(4 / 3) 9	(7 / 2) 3	(4 / 7) 1	(41 / 42) 38		(4 / 5) 3
EXP.ARIMA	(7 / 8) 2	(3 / 10) 10	(6 / 7) 1	(39 / 55) 45		(3 / 7) 5
EXP.ES&ARIMA	(4 / 3) 9	(7 / 2) 3	(4 / 7) 1	(44 / 46) 43		(6 / 6) 4
mv.EXP.ES	(2 / 6) 8	(7 / 2) 3	(3 / 2) 4	(28 / 30) 48		(2 / 3) 6
MLP	(3 / 1) 6	(2 / 2) 10	(1 / 4) 9	(43 / 17) 51		(5 / 1) 7
mv.MLP	(1 / 2) 5	(1 / 2) 7	(2 / 2) 7	(20 / 21) 36		(1 / 2) 1

Tabelle 8.14. MAPE und Rang nach MAPE je Zeitreihe und Verfahren an Absatzstelle IN-2

MAPE	S1	S2	S3	S4	S5	S6
NF1	(75,1% / 116,8%) 86,6%	(34,7% / 33,4%) 13,9%	(73,0% / 98,2%) 62,0%	(106,2% / 91,3%) 77,4%	(36,9% / 33,3%) 27,7%	(40,1% / 42,5%) 31,2%
NF2	(109,5% / 134,1%) 65,5%	(37,2% / 33,6%) 17,4%	(88,5% / 62,8%) 52,0%	(90,9% / 103,4%) 73,8%	(37,4% / 40,9%) 28,7%	(45,4% / 61,9%) 21,3%
MA	(88,0% / 102,3%) 76,2%	(30,3% / 26,7%) 16,0%	(76,3% / 82,2%) 66,8%	(89,2% / 80,3%) 73,9%	(32,3% / 29,9%) 25,3%	(34,4% / 39,4%) 26,4%
S.ES	(90,3% / 105,9%) 72,2%	(29,4% / 26,5%) 14,0%	(76,1% / 76,3%) 64,0%	(88,5% / 83,6%) 66,4%	(30,1% / 30,2%) 23,7%	(34,5% / 39,9%) 23,9%
DT.ES	(91,0% / 103,0%) 55,2%	(28,8% / 27,2%) 13,7%	(76,7% / 69,0%) 81,2%	(95,6% / 91,4%) 90,0%	(26,7% / 32,0%) 20,5%	(32,7% / 42,3%) 24,2%
EXP.ES	(91,4% / 102,0%) 55,2%	(28,7% / 27,3%) 13,7%	(76,7% / 69,0%) 81,2%	(95,6% / 91,4%) 90,0%	(27,2% / 33,7%) 20,8%	(33,9% / 40,9%) 26,5%
EXP.ARIMA	(81,7% / 98,2%) 53,6%	(28,9% / 24,7%) 12,6%	(76,0% / 69,0%) 81,2%	(85,3% / 106,5%) 90,0%	(27,9% / 35,8%) 24,3%	(39,1% / 35,7%) 26,3%
EXP.ES&ARIMA	(91,4% / 102,0%) 55,2%	(28,7% / 27,3%) 13,7%	(76,7% / 69,0%) 81,2%	(85,3% / 106,5%) 90,0%	(27,9% / 35,8%) 24,3%	(39,1% / 35,7%) 26,3%
mv.EXP.ES	(91,4% / 102,0%) 55,2%	(26,4% / 23,4%) 14,8%	(76,7% / 69,0%) 81,2%	(79,5% / 100,9%) 90,0%	(25,4% / 30,0%) 20,7%	(28,6% / 27,3%) 24,9%
MLP	(79,4% / 86,9%) 73,2%	(34,1% / 22,7%) 12,7%	(83,6% / 60,4%) 76,4%	(71,1% / 70,8%) 60,8%	(32,1% / 29,7%) 21,5%	(45,6% / 37,8%) 23,1%
mv.MLP	(78,8% / 87,8%) 65,2%	(32,0% / 23,4%) 11,6%	(84,9% / 64,5%) 77,4%	(68,3% / 83,6%) 70,6%	(15,5% / 29,2%) 21,2%	(37,7% / 33,5%) 21,9%
Verfahren	S7	S8	S9	mwMAPE		Rank mwMAPE
NF1	(37,0% / 32,9%) 19,5%	(92,6% / 80,8%) 86,3%	(97,2% / 94,6%) 77,3%	(65,9% / 69,3%) 53,5%		(10, / 10) 10
NF2	(40,6% / 46,8%) 22,3%	(110,4% / 74,0%) 65,7%	(96,4% / 92,6%) 79,1%	(72,9% / 72,2%) 47,3%		(11, / 11) 3
MA	(35,0% / 35,2%) 19,5%	(85,8% / 79,0%) 61,3%	(79,1% / 68,2%) 75,3%	(61,1% / 60,4%) 49,0%		(5, / 5) 4
S.ES	(32,3% / 31,7%) 15,8%	(93,6% / 78,1%) 70,1%	(78,3% / 71,7%) 72,8%	(61,5% / 60,4%) 47,0%		(6, / 6) 2
DT.ES	(29,4% / 27,6%) 34,6%	(98,6% / 69,3%) 58,3%	(91,4% / 69,0%) 78,9%	(63,4% / 59,0%) 50,7%		(7, / 4) 7
EXP.ES	(29,3% / 27,6%) 34,4%	(103,7% / 78,1%) 69,6%	(88,7% / 74,8%) 69,6%	(63,9% / 60,5%) 51,2%		(9, / 7) 8
EXP.ARIMA	(32,1% / 36,8%) 19,3%	(90,4% / 86,1%) 107,4%	(86,9% / 74,8%) 69,6%	(60,9% / 63,1%) 53,8%		(4, / 9) 11
EXP.ES&ARIMA	(29,3% / 27,6%) 34,4%	(103,7% / 78,1%) 69,6%	(88,7% / 74,8%) 69,6%	(63,4% / 61,9%) 51,6%		(8, / 8) 9
mv.EXP.ES	(28,1% / 25,2%) 28,2%	(103,7% / 78,1%) 69,6%	(69,2% / 63,4%) 63,1%	(58,8% / 57,7%) 49,7%		(3, / 3) 6
MLP	(24,9% / 26,0%) 19,8%	(78,4% / 73,2%) 96,1%	(71,7% / 63,4%) 60,1%	(57,9% / 52,3%) 49,3%		(2, / 1) 5
mv.MLP	(15,9% / 22,2%) 19,1%	(72,1% / 78,5%) 75,5%	(76,4% / 63,6%) 59,1%	(53,5% / 54,0%) 46,8%		(1, / 2) 1
Rang des MAE	S1	S2	S3	S4	S5	S6
NF1	(1 / 10) 11	(10 / 10) 7	(1 / 11) 2	(11 / 5) 6	(10 / 7) 10	(9 / 10) 11
NF2	(11 / 11) 7	(11 / 11) 11	(11 / 2) 1	(8 / 9) 4	(11 / 11) 11	(10 / 11) 1
MA	(5 / 7) 10	(7 / 6) 10	(4 / 10) 4	(7 / 2) 5	(9 / 3) 9	(4 / 6) 9
S.ES	(6 / 9) 8	(6 / 5) 8	(3 / 9) 3	(6 / 4) 2	(7 / 5) 6	(5 / 7) 4
DT.ES	(7 / 8) 2	(4 / 7) 4	(5 / 4) 7	(9 / 6) 7	(3 / 6) 1	(2 / 9) 5
EXP.ES	(8 / 4) 2	(2 / 8) 4	(5 / 4) 7	(9 / 6) 7	(4 / 8) 3	(3 / 8) 10
EXP.ARIMA	(4 / 3) 1	(5 / 4) 2	(2 / 4) 7	(4 / 10) 7	(5 / 9) 7	(7 / 3) 7
EXP.ES&ARIMA	(8 / 4) 2	(2 / 8) 4	(5 / 4) 7	(4 / 10) 7	(5 / 9) 7	(7 / 3) 7
mv.EXP.ES	(8 / 4) 2	(1 / 3) 9	(5 / 4) 7	(3 / 8) 7	(2 / 4) 2	(1 / 1) 6
MLP	(3 / 1) 9	(9 / 1) 3	(9 / 1) 5	(2 / 1) 1	(8 / 2) 5	(11 / 5) 3
mv.MLP	(2 / 2) 5	(8 / 2) 1	(10 / 3) 6	(1 / 3) 3	(1 / 1) 4	(6 / 2) 2
Verfahren	S7	S8	S9	Sum of Ranks MAPE		Rank of SoR MAPE
NF1	(10 / 8) 4	(5 / 10) 9	(11 / 11) 9	(68 / 82) 69		(10 / 11) 11
NF2	(11 / 11) 7	(11 / 3) 3	(10 / 10) 11	(94 / 79) 56		(11 / 10) 9
MA	(9 / 9) 5	(3 / 9) 2	(5 / 4) 8	(53 / 56) 62		(9 / 7) 10
S.ES	(8 / 7) 1	(6 / 7) 7	(4 / 6) 7	(51 / 59) 46		(7 / 8) 3
DT.ES	(6 / 4) 11	(7 / 1) 1	(9 / 5) 10	(52 / 50) 48		(8 / 4) 4
EXP.ES	(4 / 4) 9	(8 / 4) 4	(7 / 7) 4	(50 / 53) 50		(5 / 5) 7
EXP.ARIMA	(7 / 10) 3	(4 / 11) 11	(6 / 7) 4	(44 / 61) 49		(3 / 9) 6
EXP.ES&ARIMA	(4 / 4) 9	(8 / 4) 4	(7 / 7) 4	(50 / 53) 51		(5 / 5) 8
mv.EXP.ES	(3 / 2) 8	(8 / 4) 4	(1 / 1) 3	(32 / 31) 48		(1 / 3) 4
MLP	(2 / 3) 6	(2 / 2) 10	(2 / 2) 2	(48 / 18) 44		(4 / 1) 2
mv.MLP	(1 / 1) 2	(1 / 8) 8	(3 / 3) 1	(33 / 25) 32		(2 / 2) 1

8.5.2 Ergebnisse der Bedarfsprognose für Zeitreihe IN-3

Tabelle 8.15. Mittlere Fehlerwert und mittlerer Rang der Absatzstelle IN-3

Verfahren	Mittelwert MAE	Rang MAE	Mittelwert MAPE	Rang MAPE
NF1	(7,7 / 8,9) 5,4	(10 / 10) 8	(47,9%/57,9%) 61,1%	(10/3) 1
NF2	(7,9 / 9,9) 5,8	(11 / 11) 9	(51,0%/72,9%) 82,0%	(11/11) 4
MA	(5,9 / 7,7) 4,7	(8 / 4) 3	(40,4%/63,3%) 70,1%	(8/7) 3
S.ES	(6,0 / 7,7) 4,3	(9 / 7) 1	(41,1%/62,7%) 68,7%	(9/5) 2
DT.ES	(5,5 / 7,8) 5,2	(3 / 9) 7	(38,4%/66,3%) 110,0%	(6/9) 9
EXP.ES	(5,6 / 7,7) 5,0	(4 / 6) 5	(37,7%/63,7%) 98,0%	(4/8) 7
EXP.ARIMA	(5,6 / 7,8) 7,3	(5 / 8) 11	(37,9%/68,7%) 147,4%	(5/10) 11
EXP.ES&ARIMA	(5,6 / 7,7) 5,1	(6 / 5) 6	(39,0%/62,8%) 105,5%	(7/6) 8
mv.EXP.ES	(5,8 / 7,4) 4,7	(7 / 3) 4	(34,9%/58,9%) 89,3%	(3/4) 5
MLP	(4,9 / 5,8) 6,5	(2 / 2) 10	(34,2%/53,8%) 113,4%	(2/2) 10
mv.MLP	(4,6 / 5,6) 4,6	(1 / 1) 2	(29,8%/49,3%) 89,9%	(1/1) 6

* Ergebnisse werden in der Form ,(Fehler Trainingsmenge/ Fehler Validierungsmenge) Fehler Testmenge' präsentiert.

Tabelle 8.16. Mittlerer Absoluter Fehler je Zeitreihe und Verfahren an Absatzstelle IN-3 einschließlich Rang der Verfahren

MAE je Zeitreihe	S1	S2	S3	S4	S5	S6	S7	mwMAE	Rang mwMAE
NF1	(9,2 / 11,8)8,7	(9,5 / 6,9)5,6	(5,3 / 6,0)3,0	(5,5 / 4,8)2,9	(5,9 / 6,9)2,4	(15,5 / 20,6)13,5	(3,2 / 5,1)1,7	(7,7 / 8,9)5,4	(10 / 10)8
NF2	(10,4 / 12,7)8,8	(9,6 / 9,0)5,8	(5,2 / 5,5)3,2	(5,1 / 4,9)2,6	(6,5 / 7,7)3,0	(15,2 / 23,9)14,3	(3,6 / 5,6)2,9	(7,9 / 9,9)5,8	(11 / 11)9
MA	(7,6 / 10,1)7,4	(7,3 / 6,6)5,0	(3,9 / 4,7)2,6	(4,0 / 3,8)2,1	(4,9 / 5,9)2,2	(11,1 / 18,0)11,3	(2,6 / 4,4)2,3	(5,9 / 7,7)4,7	(8 / 4)3
S.ES	(7,6 / 10,6)6,6	(7,2 / 7,0)4,8	(3,9 / 4,9)2,6	(4,1 / 3,9)2,1	(4,6 / 5,9)2,0	(11,9 / 17,4)9,9	(2,6 / 4,2)2,2	(6,0 / 7,7)4,3	(9 / 7)1
DT.ES	(7,5 / 10,5)8,0	(6,9 / 7,2)4,3	(3,7 / 5,1)4,0	(3,7 / 4,5)2,3	(4,3 / 6,1)3,1	(10,0 / 17,1)10,2	(2,5 / 4,0)4,2	(5,5 / 7,8)5,2	(3 / 9)7
EXP.ES	(7,5 / 10,5)8,0	(7,0 / 6,9)4,3	(3,7 / 5,0)4,0	(3,7 / 4,3)2,2	(4,4 / 6,1)2,2	(10,5 / 17,1)11,0	(2,2 / 3,8)3,5	(5,6 / 7,7)5,0	(4 / 6)5
EXP.ARIMA	(8,0 / 9,8)11,0	(6,9 / 7,0)4,3	(3,7 / 5,4)6,8	(3,6 / 4,9)4,4	(4,4 / 6,6)7,4	(10,1 / 17,0)13,0	(2,5 / 3,6)4,2	(5,6 / 7,8)7,3	(5 / 8)11
EXP.ES&ARIMA	(7,5 / 10,5)8,0	(6,9 / 7,0)4,3	(3,7 / 5,0)4,0	(3,7 / 4,3)2,2	(4,4 / 6,1)2,2	(10,5 / 17,1)11,0	(2,5 / 3,6)4,2	(5,6 / 7,7)5,1	(6 / 5)6
mv.EXP.ES	(8,0 / 9,1)6,5	(7,9 / 7,3)5,9	(4,1 / 4,7)2,6	(3,9 / 4,2)2,4	(4,8 / 5,7)2,0	(9,6 / 16,8)10,0	(2,2 / 3,8)3,5	(5,8 / 7,4)4,7	(7 / 3)4
MLP	(6,2 / 6,9)12,6	(6,1 / 5,2)4,3	(3,3 / 3,9)4,9	(3,3 / 3,5)2,9	(4,0 / 5,3)2,8	(9,3 / 12,4)14,4	(2,0 / 3,1)3,5	(4,9 / 5,8)6,5	(2 / 2)10
mv.MLP	(6,4 / 6,8)6,3	(6,2 / 4,3)4,9	(2,8 / 3,7)2,6	(2,8 / 3,5)3,3	(3,8 / 4,6)2,0	(8,1 / 13,2)9,1	(2,0 / 3,5)3,7	(4,6 / 5,6)4,6	(1 / 1)2

Tabelle 8.17. Rang des Mittlerer Absoluten Fehler je Zeitreihe und Verfahren an Absatzstelle IN-3 einschließlich Summe der Ränge je Verfahren

Rang nach MAE	S1	S2	S3	S4	S5	S6	S7	Summe Rang	Rang
NF1	(10 / 10) 8	(10 / 4) 9	(11 / 11) 5	(11 / 9) 9	(10 / 10) 7	(11 / 10) 9	(10 / 10) 1	(73 / 64) 48	(10 / 10) 8
NF2	(11 / 11) 9	(11 / 11) 10	(10 / 10) 6	(10 / 11) 7	(11 / 11) 9	(10 / 11) 10	(11 / 11) 4	(74 / 76) 55	(11 / 11) 9
MA	(7 / 5) 4	(8 / 3) 8	(8 / 3) 2	(8 / 3) 1	(9 / 5) 4	(8 / 9) 7	(9 / 9) 3	(57 / 37) 29	(9 / 4) 2
S.ES	(6 / 9) 3	(7 / 6) 6	(7 / 5) 2	(9 / 4) 1	(7 / 4) 1	(9 / 8) 2	(8 / 8) 2	(53 / 44) 17	(8 / 7) 1
DT.ES	(3 / 6) 5	(3 / 9) 1	(4 / 8) 7	(4 / 8) 5	(3 / 6) 10	(4 / 7) 4	(5 / 7) 11	(26 / 51) 43	(3 / 9) 7
EXP.ES	(3 / 7) 5	(6 / 5) 1	(5 / 6) 7	(5 / 6) 3	(4 / 6) 4	(6 / 5) 5	(3 / 5) 5	(32 / 40) 30	(4 / 6) 3
EXP.ARIMA	(8 / 4) 10	(4 / 6) 1	(3 / 9) 11	(3 / 10) 11	(6 / 9) 11	(5 / 4) 8	(6 / 3) 9	(35 / 45) 61	(6 / 8) 11
EXP.ES&ARIMA	(3 / 7) 5	(4 / 6) 1	(5 / 6) 7	(5 / 6) 3	(4 / 6) 4	(6 / 5) 5	(6 / 3) 9	(33 / 39) 34	(5 / 5) 6
mv.EXP.ES	(9 / 3) 2	(9 / 10) 11	(9 / 3) 2	(7 / 5) 6	(8 / 3) 2	(3 / 3) 3	(4 / 6) 5	(49 / 33) 31	(7 / 3) 5
MLP	(1 / 2) 11	(1 / 2) 5	(2 / 2) 10	(2 / 2) 8	(2 / 2) 8	(2 / 1) 11	(1 / 1) 5	(11 / 12) 58	(2 / 2) 10
mv.MLP	(2 / 1) 1	(2 / 1) 7	(1 / 1) 1	(1 / 1) 10	(1 / 1) 2	(1 / 2) 1	(2 / 2) 8	(10 / 9) 30	(1 / 1) 3

Tabelle 8.18. Rang des MAPE je Zeitreihe und Verfahren an Absatzstelle IN-3 einschließlich Summe der Ränge je Verfahren

Rang nach MAE	S1	S2	S3	S4	S5	S6	S7	Summe Rang	Rang
NF1	(10 / 2) 6	(10 / 3) 7	(10 / 3) 1	(11 / 4) 4	(9 / 4) 1	(11 / 3) 9	(9 / 7) 1	(70 / 26) 29	(10 / 3) 2
NF2	(11 / 11) 4	(11 / 11) 10	(11 / 10) 6	(10 / 3) 3	(11 / 11) 6	(10 / 11) 8	(11 / 11) 4	(75 / 68) 41	(11 / 11) 7
MA	(2 / 5) 5	(3 / 10) 6	(4 / 6) 2	(9 / 5) 2	(10 / 5) 5	(8 / 10) 7	(10 / 10) 2	(46 / 51) 29	(8 / 8) 2
S.ES	(5 / 7) 3	(9 / 8) 5	(9 / 5) 4	(8 / 2) 1	(8 / 3) 3	(9 / 9) 3	(8 / 9) 3	(56 / 43) 22	(9 / 5) 1
DT.ES	(4 / 8) 7	(4 / 7) 1	(6 / 9) 7	(5 / 10) 5	(5 / 8) 10	(4 / 5) 4	(7 / 8) 11	(35 / 55) 45	(4 / 10) 9
EXP.ES	(6 / 9) 7	(5 / 4) 2	(7 / 7) 7	(6 / 8) 6	(6 / 6) 7	(6 / 6) 5	(4 / 5) 5	(40 / 45) 39	(6 / 7) 6
EXP.ARIMA	(9 / 6) 10	(6 / 5) 2	(5 / 11) 11	(3 / 11) 11	(4 / 10) 11	(5 / 8) 10	(5 / 2) 9	(37 / 53) 64	(5 / 9) 11
EXP.ES&ARIMA	(6 / 9) 7	(6 / 5) 2	(7 / 7) 7	(6 / 8) 6	(6 / 6) 7	(6 / 6) 5	(5 / 2) 9	(42 / 43) 43	(7 / 5) 8
mv.EXP.ES	(1 / 1) 11	(2 / 2) 8	(2 / 4) 10	(4 / 1) 9	(3 / 9) 9	(3 / 2) 11	(1 / 1) 5	(16 / 20) 63	(2 / 2) 10
MLP	(8 / 4) 2	(8 / 9) 11	(3 / 2) 5	(2 / 7) 8	(1 / 1) 2	(2 / 4) 2	(3 / 6) 5	(27 / 33) 35	(3 / 4) 4
mv.MLP	(3 / 3) 1	(1 / 1) 9	(1 / 1) 3	(1 / 6) 10	(2 / 2) 4	(1 / 1) 1	(2 / 4) 8	(11 / 18) 36	(1 / 1) 5

8.5.3 Ergebnisse des MAE nach Absatzverlauf

Tabelle 8.19. MAE über 4 Absatzstellen mit sporadischem Absatzverlauf*

Verfahren	Aggregierte Fehlermaße des MAE		
	sumMAE	mwMAE	mdMAE
NF1	(24,87 / 18,78) 21,14	(6,22 / 4,69) 5,29	(5,12 / 4,82) 4,38
NF2	(26,00 / 20,39) 22,14	(6,50 / 5,10) 5,54	(5,02 / 4,51) 5,14
MA	(20,13 / 16,36) 17,90	(5,03 / 4,09) 4,48	(3,63 / 3,76) 4,57
S.ES	(19,15 / 14,78) 15,86	(4,79 / 3,69) 3,96	(3,50 / 3,56) 3,86
DT.ES	(17,58 / 13,19) 23,38	(4,39 / 3,30) 5,85	(3,09 / 3,19) 3,83
EXP.ES	(17,64 / 13,28) 23,14	(4,41 / 3,32) 5,79	(3,16 / 3,24) 3,74
EXP.ARIMA	(17,81 / 14,08) 16,19	(4,45 / 3,52) 4,05	(2,91 / 2,74) 3,76
EXP.ES&ARIMA	(17,78 / 13,17) 23,33	(4,45 / 3,29) 5,83	(3,19 / 3,15) 3,83
mv.EXP.ES	(15,54 / 13,31) 20,95	(3,88 / 3,33) 5,24	(3,19 / 3,15) 3,83
MLP	(14,33 / 11,92) 16,76	(3,58 / 2,98) 4,19	(2,64 / 2,72) 3,62
mv.MLP	(12,20 / 10,86) 15,10	(3,05 / 2,72) 3,77	(2,25 / 2,15) 3,00
Verfahren	Rang der Fehlermaße des MAE		
	Rang sumMAE	Rang mwMAE	Rang mdMAE
NF1	(10 / 10) 7	(10 / 10) 7	(11 / 11) 9
NF2	(11 / 11) 8	(11 / 11) 8	(10 / 10) 11
MA	(9 / 9) 5	(9 / 9) 5	(9 / 9) 10
S.ES	(8 / 8) 2	(8 / 8) 2	(8 / 8) 8
DT.ES	(4 / 4) 11	(4 / 4) 11	(4 / 6) 5
EXP.ES	(5 / 5) 9	(5 / 5) 9	(5 / 7) 3
EXP.ARIMA	(7 / 7) 3	(7 / 7) 3	(3 / 3) 4
EXP.ES&ARIMA	(6 / 3) 10	(6 / 3) 10	(6 / 4) 5
mv.EXP.ES	(3 / 6) 6	(3 / 6) 6	(6 / 4) 5
MLP	(2 / 2) 4	(2 / 2) 4	(2 / 2) 2
mv.MLP	(1 / 1) 1	(1 / 1) 1	(1 / 1) 1

* Ergebnisse werden in der Form ‚(Fehler Trainingsmenge/ Fehler Validierungsmenge) Fehler Testmenge‘ präsentiert.

Tabelle 8.20. MAE über 21 Absatzstellen mit unregelmäßigem Absatzverlauf*

Verfahren	Aggregierte Fehlermaße des MAE		
	sumMAE	mwMAE	mdMAE
NF1	(94,45 / 91,89) 74,02	(4,50 / 4,38) 3,52	(2,57 / 2,33) 2,07
NF2	(94,29 / 93,36) 72,99	(4,49 / 4,45) 3,48	(2,38 / 2,69) 1,71
MA	(77,46 / 76,44) 65,24	(3,69 / 3,64) 3,11	(2,12 / 2,11) 1,67
S.ES	(74,96 / 74,94) 61,17	(3,57 / 3,57) 2,91	(2,03 / 2,19) 1,53
DT.ES	(70,31 / 72,86) 67,13	(3,35 / 3,47) 3,20	(1,80 / 2,06) 1,62
EXP.ES	(71,49 / 74,06) 67,05	(3,40 / 3,53) 3,19	(1,80 / 2,06) 1,62
EXP.ARIMA	(72,28 / 73,19) 69,41	(3,44 / 3,49) 3,31	(1,80 / 2,00) 1,90
EXP.ES&ARIMA	(73,35 / 74,92) 69,52	(3,49 / 3,57) 3,31	(1,80 / 2,06) 1,62
mv.EXP.ES	(63,31 / 71,47) 66,71	(3,01 / 3,40) 3,18	(1,80 / 1,97) 1,62
MLP	(71,19 / 68,86) 61,58	(3,39 / 3,28) 2,93	(1,67 / 1,86) 1,48
mv.MLP	(55,82 / 66,31) 57,57	(2,66 / 3,16) 2,74	(1,62 / 1,75) 1,57
Verfahren	Rang der Fehlermaße des MAE		
	Rang sumMAE	Rang mwMAE	Rang mdMAE
NF1	(11 / 10) 11	(11 / 10) 11	(11 / 10) 11
NF2	(10 / 11) 10	(10 / 11) 10	(10 / 11) 9
MA	(9 / 9) 4	(9 / 9) 4	(9 / 8) 8
S.ES	(8 / 8) 2	(8 / 8) 2	(8 / 9) 2
DT.ES	(3 / 4) 7	(3 / 4) 7	(3 / 5) 4
EXP.ES	(5 / 6) 6	(5 / 6) 6	(3 / 5) 4
EXP.ARIMA	(6 / 5) 8	(6 / 5) 8	(3 / 4) 10
EXP.ES&ARIMA	(7 / 7) 9	(7 / 7) 9	(3 / 5) 4
mv.EXP.ES	(2 / 3) 5	(2 / 3) 5	(3 / 3) 4
MLP	(4 / 2) 3	(4 / 2) 3	(2 / 2) 1
mv.MLP	(1 / 1) 1	(1 / 1) 1	(1 / 1) 3

* Ergebnisse werden in der Form ‚(Fehler Trainingsmenge/ Fehler Validierungsmenge) Fehler Testmenge‘ präsentiert.

Tabelle 8.21. MAE über 19 Absatzstellen mit regelmäßigem Absatzverlauf*

Verfahren	Aggregierte Fehlermaße des MAE		
	sumMAE	mwMAE	mdMAE
NF1	(130,40 / 104,28) 100,28	(6,86 / 5,49) 5,28	(3,01 / 3,33) 3,13
NF2	(122,33 / 92,42) 100,78	(6,44 / 4,86) 5,30	(3,04 / 3,53) 3,19
MA	(100,12 / 80,11) 74,46	(5,27 / 4,22) 3,92	(2,50 / 2,86) 2,48
S.ES	(93,48 / 75,94) 70,79	(4,92 / 4,00) 3,73	(2,38 / 2,78) 2,43
DT.ES	(87,18 / 71,14) 81,82	(4,59 / 3,74) 4,31	(2,28 / 2,75) 2,53
EXP.ES	(87,25 / 78,11) 78,70	(4,59 / 4,11) 4,14	(2,32 / 2,78) 2,47
EXP.ARIMA	(85,12 / 75,78) 67,42	(4,48 / 3,99) 3,55	(2,38 / 2,83) 2,71
EXP.ES&ARIMA	(88,02 / 76,83) 74,09	(4,63 / 4,04) 3,90	(2,35 / 2,78) 2,38
mv.EXP.ES	(86,23 / 75,08) 72,85	(4,54 / 3,95) 3,83	(2,09 / 2,56) 2,38
MLP	(84,16 / 65,58) 70,96	(4,43 / 3,45) 3,73	(2,33 / 2,50) 2,24
mv.MLP	(71,69 / 60,11) 64,95	(3,77 / 3,16) 3,42	(2,28 / 2,44) 2,33

Verfahren	Rang der Fehlermaßes des MAE		
	Rang sumMAE	Rang mwMAE	Rang mdMAE
NF1	(11 / 11) 10	(11 / 11) 10	(10 / 10) 10
NF2	(10 / 10) 11	(10 / 10) 11	(11 / 11) 11
MA	(9 / 9) 7	(9 / 9) 7	(9 / 9) 7
S.ES	(8 / 6) 3	(8 / 6) 3	(8 / 5) 5
DT.ES	(5 / 3) 9	(5 / 3) 9	(2 / 4) 8
EXP.ES	(6 / 8) 8	(6 / 8) 8	(4 / 5) 6
EXP.ARIMA	(3 / 5) 2	(3 / 5) 2	(7 / 8) 9
EXP.ES&ARIMA	(7 / 7) 6	(7 / 7) 6	(6 / 5) 3
mv.EXP.ES	(4 / 4) 5	(4 / 4) 5	(1 / 3) 3
MLP	(2 / 2) 4	(2 / 2) 4	(5 / 2) 1
mv.MLP	(1 / 1) 1	(1 / 1) 1	(3 / 1) 2

* Ergebnisse werden in der Form ‚(Fehler Trainingsmenge/ Fehler Validierungsmenge) Fehler Testmenge‘ präsentiert.

Tabelle 8.22. MAE über 36 Absatzstellen mit langsamer Umschlagsgeschwindigkeit *

Verfahren	Aggregierte Fehlermaße des MAE		
	sumMAE	mwMAE	mdMAE
NF1	(143,58 / 142,56) 119,46	(3,99 / 3,96) 3,32	(2,72 / 3,01) 2,38
NF2	(143,80 / 142,53) 122,91	(3,99 / 3,96) 3,41	(2,76 / 3,01) 2,64
MA	(118,34 / 118,67) 103,94	(3,29 / 3,30) 2,89	(2,32 / 2,49) 2,21
S.ES	(113,59 / 114,44) 97,38	(3,16 / 3,18) 2,70	(2,20 / 2,35) 2,10
DT.ES	(106,30 / 109,39) 111,70	(2,95 / 3,04) 3,10	(2,09 / 2,28) 2,07
EXP.ES	(107,67 / 110,97) 109,88	(2,99 / 3,08) 3,05	(2,08 / 2,26) 2,24
EXP.ARIMA	(107,68 / 110,94) 104,96	(2,99 / 3,08) 2,92	(2,04 / 2,26) 2,14
EXP.ES&ARIMA	(108,77 / 111,47) 110,36	(3,02 / 3,10) 3,07	(2,08 / 2,26) 2,12
mv.EXP.ES	(97,66 / 107,19) 104,60	(2,71 / 2,98) 2,91	(2,04 / 2,26) 1,95
MLP	(104,10 / 102,39) 96,98	(2,89 / 2,84) 2,69	(2,07 / 2,18) 1,95
mv.MLP	(88,48 / 97,86) 92,29	(2,46 / 2,72) 2,56	(1,98 / 2,00) 1,93

Verfahren	Rang der Fehlermaßes des MAE		
	Rang sumMAE	Rang mwMAE	Rang mdMAE
NF1	(10 / 11) 10	(10 / 11) 10	(10 / 10) 10
NF2	(11 / 10) 11	(11 / 10) 11	(11 / 11) 11
MA	(9 / 9) 4	(9 / 9) 4	(9 / 9) 8
S.ES	(8 / 8) 3	(8 / 8) 3	(8 / 8) 5
DT.ES	(4 / 4) 9	(4 / 4) 9	(7 / 7) 4
EXP.ES	(5 / 6) 7	(5 / 6) 7	(5 / 3) 9
EXP.ARIMA	(6 / 5) 6	(6 / 5) 6	(3 / 3) 7
EXP.ES&ARIMA	(7 / 7) 8	(7 / 7) 8	(5 / 3) 6
mv.EXP.ES	(2 / 3) 5	(2 / 3) 5	(2 / 3) 2
MLP	(3 / 2) 2	(3 / 2) 2	(4 / 2) 2
mv.MLP	(1 / 1) 1	(1 / 1) 1	(1 / 1) 1

* Ergebnisse werden in der Form ‚(Fehler Trainingsmenge/ Fehler Validierungsmenge) Fehler Testmenge‘ präsentiert.

Tabelle 8.23. MAE über 8 Absatzstellen mit schneller Umschlagsgeschwindigkeit *

Verfahren	Aggregierte Fehlermaße des MAE		
	sumMAE	mwMAE	mdMAE
NF1	(106,14 / 72,39) 75,97	(13,27 / 9,05) 9,50	(6,15 / 6,58) 5,90
NF2	(98,81 / 63,64) 73,00	(12,35 / 7,95) 9,13	(5,33 / 4,86) 6,02
MA	(79,37 / 54,25) 53,67	(9,92 / 6,78) 6,71	(4,39 / 3,92) 4,07
S.ES	(74,01 / 51,22) 50,44	(9,25 / 6,40) 6,30	(4,28 / 3,88) 3,86
DT.ES	(68,76 / 47,81) 60,63	(8,59 / 5,98) 7,58	(4,11 / 3,97) 3,81
EXP.ES	(68,71 / 54,47) 59,01	(8,59 / 6,81) 7,38	(3,74 / 3,88) 3,76
EXP.ARIMA	(67,54 / 52,11) 48,06	(8,44 / 6,51) 6,01	(3,68 / 3,85) 3,93
EXP.ES&ARIMA	(70,39 / 53,44) 56,58	(8,80 / 6,68) 7,07	(4,03 / 3,93) 3,71
mv.EXP.ES	(67,41 / 52,67) 55,91	(8,43 / 6,58) 6,99	(4,03 / 3,88) 3,74
MLP	(65,57 / 43,97) 52,32	(8,20 / 5,50) 6,54	(3,26 / 3,69) 4,52
mv.MLP	(51,22 / 39,42) 45,33	(6,40 / 4,93) 5,67	(2,63 / 3,33) 4,07
Verfahren	Rang der Fehlermaße des MAE		
	Rang sumMAE	Rang mwMAE	Rang mdMAE
NF1	(11 / 11) 11	(11 / 11) 11	(11 / 11) 10
NF2	(10 / 10) 10	(10 / 10) 10	(10 / 10) 11
MA	(9 / 8) 5	(9 / 8) 5	(9 / 7) 7
S.ES	(8 / 4) 3	(8 / 4) 3	(8 / 4) 5
DT.ES	(6 / 3) 9	(6 / 3) 9	(7 / 9) 4
EXP.ES	(5 / 9) 8	(5 / 9) 8	(4 / 4) 3
EXP.ARIMA	(4 / 5) 2	(4 / 5) 2	(3 / 3) 6
EXP.ES&ARIMA	(7 / 7) 7	(7 / 7) 7	(5 / 8) 1
mv.EXP.ES	(3 / 6) 6	(3 / 6) 6	(5 / 4) 2
MLP	(2 / 2) 4	(2 / 2) 4	(2 / 2) 9
mv.MLP	(1 / 1) 1	(1 / 1) 1	(1 / 1) 7

* Ergebnisse werden in der Form ‚(Fehler Trainingsmenge/ Fehler Validierungsmenge) Fehler Testmenge‘ präsentiert.

8.6 Ergebnistabellen der Bestandsrechnung

8.6.1 Ergebnisse der Summe der Ränge über alle Zeitreihen

Tabelle 8.24. mdRang der Sum of Ranks nach MLLC Kosten der Warenbestände und Rang je Absatzstelle* - identisch mit mdRang des SumofRanks nach SLLC

Verfahren	mdMLLC			
	AU-1	AU-2	IN-1	IN-2
NF1Bestand	(11,5 / 12,0) 11,0	(11,0 / 11,0) 11,0	(11,5 / 12,0) 11,5	(12,0 / 11,0) 11,0
NF2 Bestand	(11,0 / 10,5) 10,0	(12,0 / 11,0) 12,0	(8,5 / 7,0) 10,0	(11,0 / 12,0) 11,0
MA Bestand	(8,0 / 9,0) 9,0	(9,0 / 9,0) 8,0	(8,5 / 10,0) 7,0	(5,0 / 8,0) 6,0
S.ES Bestand	(8,5 / 8,5) 5,5	(8,0 / 4,0) 6,0	(7,0 / 6,5) 6,0	(6,0 / 8,0) 5,0
DT.ES Bestand	(4,0 / 6,0) 5,5	(5,0 / 4,0) 6,0	(7,0 / 6,0) 6,5	(7,0 / 4,0) 5,0
EXP.ES Bestand	(6,0 / 4,5) 6,0	(5,0 / 6,0) 6,0	(7,0 / 6,0) 3,5	(7,0 / 4,0) 5,0
EXP.ARIMA B.	(6,0 / 4,5) 5,5	(6,0 / 6,0) 3,0	(5,0 / 5,0) 4,5	(8,0 / 8,0) 6,0
EXP.ES&ARIMA	(7,0 / 4,5) 5,0	(8,0 / 5,0) 4,0	(8,0 / 9,0) 7,5	(7,0 / 4,0) 6,0
mv.EXP.ES Bestand	(4,0 / 5,0) 3,0	(3,0 / 3,0) 4,0	(3,5 / 7,0) 6,5	(2,0 / 4,0) 4,0
MLP Bestand	(2,0 / 3,5) 2,0	(2,0 / 8,0) 2,0	(3,5 / 6,0) 5,5	(4,0 / 5,0) 8,0
mv.MLP Bestand	(2,0 / 3,5) 3,0	(3,0 / 10,0) 2,0	(3,0 / 4,5) 4,5	(5,0 / 6,0) 7,0
mv.MLP.ACF B.	(2,5 / 1,0) 4,0	(2,0 / 2,0) 9,0	(1,0 / 1,0) 5,0	(1,0 / 2,0) 3,0
Verfahren	Rang mdMLLC			
	AU-1	AU-2	IN-1	IN-2
NF1Bestand	(12 / 12) 12	(11 / 11) 11	(12 / 12) 12	(12 / 11) 11
NF2 Bestand	(11 / 11) 11	(12 / 11) 12	(10 / 8) 11	(11 / 12) 11
MA Bestand	(9 / 10) 10	(10 / 9) 9	(10 / 11) 9	(4 / 8) 6
S.ES Bestand	(10 / 9) 6	(8 / 3) 6	(6 / 7) 6	(6 / 8) 3
DT.ES Bestand	(4 / 8) 6	(5 / 3) 6	(6 / 4) 7	(7 / 2) 3
EXP.ES Bestand	(6 / 4) 9	(5 / 6) 6	(6 / 4) 1	(7 / 2) 3
EXP.ARIMA B.	(6 / 4) 6	(7 / 6) 3	(5 / 3) 2	(10 / 8) 6
EXP.ES&ARIMA	(8 / 4) 5	(8 / 5) 4	(9 / 10) 10	(7 / 2) 6
mv.EXP.ES Bestand	(4 / 7) 2	(3 / 2) 4	(3 / 8) 7	(2 / 2) 2
MLP Bestand	(1 / 2) 1	(1 / 8) 1	(3 / 4) 5	(3 / 6) 10
mv.MLP Bestand	(1 / 2) 2	(3 / 10) 1	(2 / 2) 2	(4 / 7) 9
mv.MLP.ACF B.	(3 / 1) 4	(1 / 1) 10	(1 / 1) 4	(1 / 1) 1

8.6.2 Ergebnistabellen der Disposition je Zeitreihe

Tabelle 8.25. MLLC je Zeitreihe und Verfahren an Absatzstelle AU-1

MLLC	S1	S2	S3	S4	S5	S6	S7
NF1Bestand	(0,017 / 0,014) 0,015	(0,029 / 0,050) 0,029	(0,032 / 0,035) 0,024	(0,052 / 0,046) 0,038	(0,036 / 0,055) 0,038	(0,018 / 0,017) 0,015	(0,028 / 0,026) 0,019
NF2 Bestand	(0,016 / 0,014) 0,015	(0,036 / 0,030) 0,032	(0,037 / 0,030) 0,023	(0,060 / 0,039) 0,042	(0,032 / 0,037) 0,031	(0,021 / 0,014) 0,015	(0,029 / 0,019) 0,018
MA Bestand	(0,013 / 0,014) 0,012	(0,028 / 0,031) 0,026	(0,026 / 0,031) 0,021	(0,040 / 0,035) 0,032	(0,026 / 0,030) 0,035	(0,014 / 0,014) 0,018	(0,018 / 0,018) 0,016
S.ES Bestand	(0,013 / 0,014) 0,012	(0,026 / 0,033) 0,025	(0,027 / 0,031) 0,019	(0,042 / 0,035) 0,031	(0,026 / 0,030) 0,030	(0,015 / 0,013) 0,017	(0,018 / 0,019) 0,016
DT.ES Bestand	(0,015 / 0,012) 0,010	(0,026 / 0,033) 0,033	(0,029 / 0,026) 0,019	(0,033 / 0,031) 0,024	(0,023 / 0,027) 0,029	(0,015 / 0,012) 0,014	(0,015 / 0,020) 0,019
EXP.ES Bestand	(0,013 / 0,012) 0,010	(0,027 / 0,035) 0,033	(0,029 / 0,026) 0,019	(0,038 / 0,032) 0,030	(0,024 / 0,030) 0,030	(0,015 / 0,010) 0,011	(0,016 / 0,018) 0,018
EXP.ARIMA B.	(0,014 / 0,012) 0,010	(0,027 / 0,029) 0,029	(0,030 / 0,026) 0,019	(0,041 / 0,033) 0,034	(0,025 / 0,026) 0,028	(0,017 / 0,010) 0,011	(0,015 / 0,018) 0,019
EXP.ES&ARIMA	(0,013 / 0,012) 0,010	(0,028 / 0,025) 0,029	(0,029 / 0,026) 0,019	(0,042 / 0,033) 0,034	(0,025 / 0,026) 0,028	(0,013 / 0,010) 0,011	(0,017 / 0,018) 0,018
mv.EXP.ES Bestand	(0,012 / 0,012) 0,010	(0,024 / 0,033) 0,025	(0,023 / 0,025) 0,019	(0,034 / 0,030) 0,026	(0,019 / 0,032) 0,023	(0,013 / 0,010) 0,011	(0,017 / 0,018) 0,018
MLP Bestand	(0,011 / 0,011) 0,010	(0,020 / 0,035) 0,024	(0,018 / 0,031) 0,016	(0,022 / 0,034) 0,023	(0,026 / 0,020) 0,027	(0,014 / 0,010) 0,011	(0,011 / 0,018) 0,015
mv.MLP Bestand	(0,011 / 0,011) 0,010	(0,019 / 0,032) 0,024	(0,021 / 0,034) 0,013	(0,026 / 0,033) 0,022	(0,022 / 0,018) 0,030	(0,008 / 0,008) 0,014	(0,015 / 0,016) 0,014
mv.MLP.ACF B.	(0,011 / 0,011) 0,011	(0,027 / 0,023) 0,027	(0,022 / 0,022) 0,023	(0,026 / 0,029) 0,027	(0,026 / 0,022) 0,028	(0,012 / 0,010) 0,011	(0,013 / 0,016) 0,015
Verfahren	S8	S9	S10	S11	S12	mwMLLC	Rank mwMLLC
NF1Bestand	(0,048 / 0,033) 0,030	(0,028 / 0,024) 0,019	(0,022 / 0,019) 0,018	(0,028 / 0,025) 0,056	(0,021 / 0,026) 0,020		
NF2 Bestand	(0,033 / 0,042) 0,029	(0,023 / 0,019) 0,017	(0,021 / 0,019) 0,018	(0,021 / 0,023) 0,062	(0,022 / 0,019) 0,022		
MA Bestand	(0,042 / 0,035) 0,026	(0,020 / 0,022) 0,015	(0,018 / 0,016) 0,015	(0,023 / 0,019) 0,023	(0,017 / 0,015) 0,028		
S.ES Bestand	(0,038 / 0,036) 0,025	(0,021 / 0,021) 0,015	(0,018 / 0,015) 0,015	(0,021 / 0,018) 0,034	(0,021 / 0,017) 0,023		
DT.ES Bestand	(0,029 / 0,032) 0,021	(0,019 / 0,017) 0,015	(0,014 / 0,013) 0,014	(0,021 / 0,017) 0,045	(0,015 / 0,018) 0,022		
EXP.ES Bestand	(0,033 / 0,034) 0,025	(0,018 / 0,016) 0,015	(0,017 / 0,014) 0,014	(0,021 / 0,017) 0,041	(0,016 / 0,018) 0,021		
EXP.ARIMA B.	(0,041 / 0,036) 0,031	(0,018 / 0,016) 0,015	(0,015 / 0,016) 0,014	(0,020 / 0,016) 0,037	(0,016 / 0,018) 0,021		
EXP.ES&ARIMA	(0,033 / 0,034) 0,026	(0,018 / 0,016) 0,015	(0,017 / 0,014) 0,014	(0,021 / 0,016) 0,041	(0,014 / 0,018) 0,021		
mv.EXP.ES Bestand	(0,026 / 0,035) 0,020	(0,017 / 0,016) 0,015	(0,017 / 0,014) 0,014	(0,020 / 0,017) 0,036	(0,015 / 0,018) 0,021		
MLP Bestand	(0,034 / 0,034) 0,028	(0,016 / 0,016) 0,015	(0,014 / 0,014) 0,013	(0,018 / 0,020) 0,057	(0,022 / 0,016) 0,016		
mv.MLP Bestand	(0,029 / 0,038) 0,025	(0,015 / 0,016) 0,015	(0,014 / 0,014) 0,013	(0,017 / 0,020) 0,060	(0,023 / 0,016) 0,018		
mv.MLP.ACF B.	(0,020 / 0,036) 0,022	(0,014 / 0,013) 0,015	(0,011 / 0,011) 0,011	(0,024 / 0,014) 0,057	(0,026 / 0,015) 0,016		

Tabelle 8.26. Summe der LLC Kosten (SLLC) je Zeitreihe und Verfahren an Absatzstelle AU-1

MLLC	S1	S2	S3	S4	S5	S6	S7
NF1 Bestand	(1,094 / 0,511) 0,310	(1,926 / 1,801) 0,619	(2,122 / 1,275) 0,500	(3,452 / 1,657) 0,792	(2,373 / 1,973) 0,802	(1,163 / 0,611) 0,317	(1,868 / 0,952) 0,400
NF2 Bestand	(1,051 / 0,511) 0,317	(2,356 / 1,079) 0,666	(2,461 / 1,085) 0,475	(3,967 / 1,418) 0,875	(2,116 / 1,330) 0,644	(1,385 / 0,508) 0,313	(1,928 / 0,673) 0,385
MA Bestand	(0,839 / 0,514) 0,259	(1,829 / 1,121) 0,544	(1,709 / 1,131) 0,432	(2,664 / 1,269) 0,670	(1,698 / 1,096) 0,740	(0,904 / 0,514) 0,381	(1,184 / 0,665) 0,342
S.ES Bestand	(0,846 / 0,496) 0,259	(1,742 / 1,181) 0,522	(1,801 / 1,102) 0,396	(2,772 / 1,276) 0,655	(1,701 / 1,088) 0,626	(0,975 / 0,485) 0,348	(1,170 / 0,672) 0,328
DT.ES Bestand	(0,974 / 0,438) 0,220	(1,727 / 1,170) 0,684	(1,943 / 0,919) 0,403	(2,214 / 1,122) 0,504	(1,554 / 0,988) 0,605	(0,993 / 0,414) 0,299	(0,983 / 0,722) 0,392
EXP.ES Bestand	(0,874 / 0,438) 0,220	(1,778 / 1,245) 0,684	(1,975 / 0,919) 0,403	(2,527 / 1,150) 0,634	(1,608 / 1,084) 0,634	(0,986 / 0,371) 0,223	(1,098 / 0,640) 0,385
EXP.ARIMA B.	(0,934 / 0,424) 0,220	(1,807 / 1,027) 0,612	(1,996 / 0,919) 0,403	(2,780 / 1,176) 0,709	(1,683 / 0,926) 0,583	(1,168 / 0,364) 0,223	(1,008 / 0,658) 0,392
EXP.ES&ARIMA	(0,881 / 0,438) 0,220	(1,843 / 0,913) 0,612	(1,975 / 0,919) 0,403	(2,823 / 1,176) 0,709	(1,687 / 0,923) 0,583	(0,889 / 0,371) 0,223	(1,109 / 0,644) 0,385
mv.EXP.ES Bestand	(0,781 / 0,438) 0,220	(1,583 / 1,183) 0,515	(1,528 / 0,908) 0,389	(2,246 / 1,096) 0,540	(1,274 / 1,158) 0,490	(0,896 / 0,371) 0,223	(1,120 / 0,633) 0,382
MLP Bestand	(0,637 / 0,399) 0,202	(1,163 / 1,244) 0,500	(1,072 / 1,118) 0,331	(1,304 / 1,242) 0,493	(1,526 / 0,713) 0,572	(0,828 / 0,364) 0,223	(0,666 / 0,650) 0,317
mv.MLP Bestand	(0,644 / 0,399) 0,212	(1,141 / 1,137) 0,500	(1,218 / 1,207) 0,277	(1,523 / 1,177) 0,464	(1,274 / 0,641) 0,629	(0,493 / 0,277) 0,284	(0,877 / 0,564) 0,292
mv.MLP.ACF B.	(0,684 / 0,378) 0,238	(1,649 / 0,816) 0,569	(1,346 / 0,788) 0,479	(1,548 / 1,039) 0,572	(1,570 / 0,806) 0,590	(0,734 / 0,374) 0,223	(0,756 / 0,579) 0,320
Verfahren	S8	S9	S10	S11	S12	mwSLLC	Rank mwSLLC
NF1 Bestand	(3,182 / 1,194) 0,634	(1,828 / 0,848) 0,389	(1,475 / 0,666) 0,378	(1,843 / 0,896) 1,183	(1,112 / 0,948) 0,414		
NF2 Bestand	(2,191 / 1,508) 0,605	(1,503 / 0,688) 0,356	(1,371 / 0,670) 0,371	(1,390 / 0,845) 1,297	(1,112 / 0,701) 0,464		
MA Bestand	(2,776 / 1,276) 0,547	(1,335 / 0,794) 0,324	(1,181 / 0,583) 0,306	(1,498 / 0,688) 0,475	(0,924 / 0,526) 0,589		
S.ES Bestand	(2,496 / 1,300) 0,515	(1,381 / 0,772) 0,310	(1,166 / 0,544) 0,310	(1,400 / 0,644) 0,721	(1,114 / 0,604) 0,489		
DT.ES Bestand	(1,926 / 1,160) 0,432	(1,284 / 0,608) 0,317	(0,947 / 0,482) 0,284	(1,392 / 0,619) 0,953	(0,846 / 0,664) 0,468		
EXP.ES Bestand	(2,209 / 1,232) 0,533	(1,223 / 0,586) 0,317	(1,145 / 0,497) 0,284	(1,410 / 0,598) 0,860	(0,914 / 0,650) 0,436		
EXP.ARIMA B.	(2,744 / 1,286) 0,652	(1,184 / 0,582) 0,317	(0,997 / 0,558) 0,295	(1,342 / 0,576) 0,782	(0,870 / 0,650) 0,436		
EXP.ES&ARIMA	(2,206 / 1,232) 0,536	(1,238 / 0,586) 0,317	(1,159 / 0,497) 0,284	(1,407 / 0,590) 0,856	(0,796 / 0,636) 0,432		
mv.EXP.ES Bestand	(1,754 / 1,270) 0,428	(1,138 / 0,590) 0,317	(1,152 / 0,497) 0,284	(1,314 / 0,601) 0,760	(0,832 / 0,654) 0,439		
MLP Bestand	(2,032 / 1,210) 0,587	(0,950 / 0,575) 0,317	(0,810 / 0,490) 0,277	(1,076 / 0,729) 1,196	(1,068 / 0,575) 0,335		
mv.MLP Bestand	(1,684 / 1,356) 0,522	(0,868 / 0,564) 0,317	(0,821 / 0,493) 0,281	(1,015 / 0,704) 1,264	(1,096 / 0,564) 0,374		
mv.MLP.ACF B.	(1,209 / 1,295) 0,461	(0,817 / 0,461) 0,313	(0,680 / 0,403) 0,227	(1,422 / 0,500) 1,192	(1,285 / 0,554) 0,346		

Tabelle 8.27. Rang nach MLLC je Zeitreihe und Verfahren an Absatzstelle AU-1 (identisch mit Rang nach SLLC)

Verfahren	MLLC						
	S1	S2	S3	S4	S5	S6	S7
NF1Bestand	(12 / 11) 11	(11 / 12) 9	(11 / 12) 12	(11 / 12) 11	(12 / 12) 12	(11 / 12) 10	(11 / 12) 12
NF2 Bestand	(11 / 10) 12	(12 / 4) 10	(12 / 7) 10	(12 / 11) 12	(11 / 11) 10	(12 / 10) 9	(12 / 10) 7
MA Bestand	(5 / 12) 10	(10 / 5) 5	(5 / 10) 9	(7 / 9) 8	(7 / 9) 11	(5 / 11) 12	(10 / 8) 5
S.ES Bestand	(6 / 9) 9	(5 / 8) 4	(6 / 8) 4	(9 / 10) 7	(8 / 8) 7	(8 / 9) 11	(9 / 9) 4
DT.ES Bestand	(10 / 5) 3	(4 / 7) 11	(7 / 3) 5	(4 / 3) 3	(3 / 6) 6	(9 / 8) 8	(3 / 11) 10
EXP.ES Bestand	(7 / 5) 3	(6 / 11) 11	(8 / 3) 5	(6 / 4) 6	(4 / 7) 9	(7 / 4) 1	(6 / 4) 7
EXP.ARIMA B.	(9 / 4) 3	(7 / 3) 7	(10 / 3) 5	(8 / 5) 9	(5 / 5) 3	(10 / 2) 1	(5 / 7) 10
EXP.ES&ARIMA	(8 / 5) 3	(9 / 2) 7	(8 / 3) 5	(10 / 5) 9	(6 / 4) 3	(3 / 4) 1	(7 / 5) 7
mv.EXP.ES Bestand	(4 / 5) 3	(3 / 9) 3	(4 / 2) 3	(5 / 2) 4	(1 / 10) 1	(4 / 4) 1	(8 / 3) 6
MLP Bestand	(1 / 2) 1	(2 / 10) 2	(1 / 9) 2	(1 / 8) 2	(9 / 2) 2	(6 / 2) 1	(1 / 6) 2
mv.MLP Bestand	(2 / 2) 2	(1 / 6) 1	(2 / 11) 1	(3 / 7) 1	(2 / 1) 8	(1 / 1) 7	(4 / 1) 1
mv.MLP.ACF B.	(3 / 1) 8	(8 / 1) 6	(3 / 1) 11	(2 / 1) 5	(10 / 3) 5	(2 / 7) 1	(2 / 2) 3
Verfahren	S8	S9	S10	S11	S12	mwRang MLLC	Rank mwRang MLLC
NF1Bestand	(12 / 2) 11	(12 / 12) 12	(12 / 11) 12	(12 / 12) 8	(7 / 12) 4		
NF2 Bestand	(7 / 12) 10	(11 / 9) 11	(11 / 12) 11	(8 / 11) 12	(9 / 11) 9		
MA Bestand	(11 / 7) 8	(9 / 11) 10	(10 / 10) 9	(10 / 8) 1	(6 / 1) 12		
S.ES Bestand	(9 / 10) 4	(10 / 10) 1	(9 / 8) 10	(9 / 7) 2	(8 / 5) 11		
DT.ES Bestand	(4 / 1) 2	(8 / 8) 3	(4 / 2) 4	(5 / 6) 7	(3 / 10) 10		
EXP.ES Bestand	(6 / 4) 6	(6 / 5) 3	(6 / 5) 4	(7 / 4) 6	(5 / 7) 6		
EXP.ARIMA B.	(10 / 8) 12	(5 / 4) 3	(5 / 9) 8	(4 / 2) 4	(4 / 7) 6		
EXP.ES&ARIMA	(5 / 4) 7	(7 / 5) 3	(8 / 5) 4	(6 / 3) 5	(1 / 6) 5		
mv.EXP.ES Bestand	(2 / 6) 1	(4 / 7) 3	(7 / 5) 4	(3 / 5) 3	(2 / 9) 8		
MLP Bestand	(8 / 3) 9	(3 / 3) 3	(2 / 3) 2	(2 / 10) 10	(10 / 4) 1		
mv.MLP Bestand	(3 / 11) 5	(2 / 2) 3	(3 / 4) 3	(1 / 9) 11	(11 / 3) 3		
mv.MLP.ACF B.	(1 / 9) 3	(1 / 1) 2	(1 / 1) 1	(11 / 1) 9	(12 / 2) 2		

Tabelle 8.28. MLLC je Zeitreihe und Verfahren an Absatzstelle AU-2

Verfahren	MLLC							Rank mwMLLC
	S1	S2	S3	S4	S5	S6	S7	
NF1Bestand	(0,030 / 0,053) 0,029	(0,028 / 0,023) 0,025	(0,026 / 0,038) 0,029	(0,032 / 0,037) 0,029	(0,019 / 0,014) 0,014	(0,025 / 0,027) 0,019	(0,026 / 0,025) 0,025	
NF2 Bestand	(0,028 / 0,031) 0,038	(0,027 / 0,022) 0,026	(0,031 / 0,038) 0,028	(0,034 / 0,039) 0,031	(0,020 / 0,017) 0,015	(0,025 / 0,024) 0,019	(0,029 / 0,032) 0,029	
MA Bestand	(0,024 / 0,044) 0,027	(0,023 / 0,018) 0,020	(0,025 / 0,034) 0,023	(0,028 / 0,023) 0,025	(0,016 / 0,013) 0,013	(0,025 / 0,025) 0,016	(0,026 / 0,023) 0,022	
S.ES Bestand	(0,023 / 0,036) 0,045	(0,023 / 0,018) 0,019	(0,023 / 0,036) 0,022	(0,026 / 0,021) 0,024	(0,015 / 0,012) 0,011	(0,024 / 0,024) 0,015	(0,027 / 0,021) 0,021	
DT.ES Bestand	(0,020 / 0,040) 0,032	(0,020 / 0,018) 0,019	(0,020 / 0,031) 0,021	(0,022 / 0,021) 0,024	(0,017 / 0,012) 0,014	(0,019 / 0,020) 0,013	(0,024 / 0,022) 0,024	
EXP.ES Bestand	(0,020 / 0,040) 0,032	(0,021 / 0,021) 0,019	(0,020 / 0,033) 0,021	(0,022 / 0,021) 0,026	(0,016 / 0,012) 0,011	(0,020 / 0,020) 0,013	(0,027 / 0,023) 0,020	
EXP.ARIMA B.	(0,022 / 0,039) 0,032	(0,019 / 0,020) 0,016	(0,022 / 0,031) 0,021	(0,024 / 0,022) 0,021	(0,018 / 0,013) 0,014	(0,019 / 0,020) 0,013	(0,025 / 0,021) 0,020	
EXP.ES&ARIMA	(0,020 / 0,037) 0,032	(0,019 / 0,020) 0,019	(0,020 / 0,031) 0,021	(0,023 / 0,021) 0,025	(0,017 / 0,012) 0,011	(0,020 / 0,020) 0,013	(0,028 / 0,023) 0,020	
mv.EXP.ES Bestand	(0,020 / 0,038) 0,030	(0,018 / 0,018) 0,019	(0,017 / 0,028) 0,021	(0,020 / 0,021) 0,023	(0,013 / 0,012) 0,011	(0,020 / 0,020) 0,013	(0,024 / 0,020) 0,020	
MLP Bestand	(0,017 / 0,042) 0,032	(0,017 / 0,020) 0,022	(0,033 / 0,035) 0,017	(0,023 / 0,020) 0,022	(0,014 / 0,012) 0,011	(0,016 / 0,026) 0,011	(0,022 / 0,031) 0,019	
mv.MLP Bestand	(0,016 / 0,047) 0,031	(0,018 / 0,025) 0,017	(0,029 / 0,037) 0,018	(0,021 / 0,022) 0,021	(0,014 / 0,012) 0,011	(0,020 / 0,027) 0,010	(0,020 / 0,034) 0,021	
mv.MLP.ACF B.	(0,036 / 0,030) 0,028	(0,023 / 0,020) 0,022	(0,018 / 0,025) 0,023	(0,019 / 0,017) 0,015	(0,009 / 0,012) 0,011	(0,017 / 0,018) 0,017	(0,024 / 0,020) 0,023	
Verfahren	S8	S9	S10	S11	S12	S13		
NF1Bestand	(0,028 / 0,024) 0,023	(0,033 / 0,059) 0,038	(0,019 / 0,014) 0,015	(0,029 / 0,022) 0,027	(0,025 / 0,021) 0,022	(0,034 / 0,034) 0,030		
NF2 Bestand	(0,028 / 0,025) 0,026	(0,035 / 0,040) 0,032	(0,023 / 0,015) 0,015	(0,028 / 0,017) 0,020	(0,029 / 0,022) 0,022	(0,039 / 0,032) 0,043		
MA Bestand	(0,023 / 0,029) 0,020	(0,029 / 0,022) 0,025	(0,017 / 0,012) 0,012	(0,025 / 0,025) 0,023	(0,021 / 0,019) 0,019	(0,029 / 0,026) 0,025		
S.ES Bestand	(0,023 / 0,025) 0,019	(0,026 / 0,023) 0,025	(0,016 / 0,011) 0,011	(0,027 / 0,018) 0,022	(0,019 / 0,017) 0,017	(0,027 / 0,022) 0,023		
DT.ES Bestand	(0,021 / 0,028) 0,021	(0,027 / 0,021) 0,031	(0,014 / 0,014) 0,013	(0,024 / 0,019) 0,022	(0,019 / 0,016) 0,016	(0,028 / 0,022) 0,025		
EXP.ES Bestand	(0,021 / 0,028) 0,021	(0,023 / 0,022) 0,021	(0,014 / 0,013) 0,014	(0,023 / 0,015) 0,022	(0,020 / 0,017) 0,019	(0,028 / 0,022) 0,025		
EXP.ARIMA B.	(0,021 / 0,032) 0,017	(0,023 / 0,024) 0,022	(0,016 / 0,011) 0,013	(0,025 / 0,013) 0,020	(0,019 / 0,018) 0,020	(0,028 / 0,023) 0,024		
EXP.ES&ARIMA	(0,022 / 0,028) 0,021	(0,025 / 0,022) 0,021	(0,017 / 0,012) 0,013	(0,026 / 0,020) 0,020	(0,020 / 0,017) 0,019	(0,029 / 0,022) 0,025		
mv.EXP.ES Bestand	(0,020 / 0,028) 0,021	(0,024 / 0,019) 0,021	(0,016 / 0,012) 0,013	(0,020 / 0,018) 0,020	(0,019 / 0,017) 0,019	(0,018 / 0,022) 0,027		
MLP Bestand	(0,026 / 0,031) 0,017	(0,022 / 0,029) 0,025	(0,007 / 0,011) 0,044	(0,026 / 0,015) 0,016	(0,013 / 0,019) 0,014	(0,024 / 0,023) 0,027		
mv.MLP Bestand	(0,029 / 0,033) 0,016	(0,021 / 0,030) 0,028	(0,009 / 0,012) 0,040	(0,024 / 0,017) 0,015	(0,014 / 0,013) 0,014	(0,023 / 0,019) 0,024		
mv.MLP.ACF B.	(0,025 / 0,026) 0,025	(0,023 / 0,020) 0,021	(0,009 / 0,011) 0,041	(0,023 / 0,019) 0,019	(0,016 / 0,018) 0,020	(0,021 / 0,020) 0,029		

* Ergebnisse werden in der Form ‚(Fehler Trainingsmenge/ Fehler Validierungsmenge) Fehler Tesmenge‘ präsentiert.

Tabelle 8.29. Summe der LLC Kosten (SLLC) je Zeitreihe und Verfahren an Absatzstelle AU-2

Verfahren	MLLC							Rank mwSLLC
	S1	S2	S3	S4	S5	S6	S7	
NF1Bestand	(2,002 / 1,897) 0,601	(1,818 / 0,828) 0,522	(1,706 / 1,376) 0,608	(2,131 / 1,322) 0,619	(1,274 / 0,515) 0,299	(1,655 / 0,970) 0,392	(1,713 / 0,900) 0,526	
NF2 Bestand	(1,854 / 1,125) 0,794	(1,750 / 0,785) 0,547	(2,054 / 1,379) 0,590	(2,213 / 1,390) 0,659	(1,289 / 0,608) 0,306	(1,619 / 0,852) 0,389	(1,919 / 1,136) 0,601	
MA Bestand	(1,613 / 1,578) 0,564	(1,488 / 0,648) 0,418	(1,644 / 1,235) 0,493	(1,840 / 0,828) 0,529	(1,026 / 0,461) 0,263	(1,650 / 0,894) 0,342	(1,748 / 0,810) 0,464	
S.ES Bestand	(1,580 / 0,434) 0,538	(1,494 / 0,630) 0,396	(1,493 / 1,292) 0,472	(1,692 / 0,767) 0,504	(0,983 / 0,418) 0,234	(1,556 / 0,865) 0,324	(1,790 / 0,752) 0,439	
DT.ES Bestand	(1,336 / 1,448) 0,664	(1,340 / 0,654) 0,407	(1,314 / 1,103) 0,446	(1,480 / 0,742) 0,511	(1,153 / 0,428) 0,295	(1,305 / 0,718) 0,281	(1,633 / 0,791) 0,504	
EXP.ES Bestand	(1,332 / 1,455) 0,664	(1,404 / 0,740) 0,407	(1,343 / 1,203) 0,446	(1,501 / 0,756) 0,536	(1,096 / 0,428) 0,234	(1,334 / 0,718) 0,281	(1,793 / 0,837) 0,428	
EXP.ARIMA B.	(1,472 / 1,408) 0,664	(1,258 / 0,704) 0,331	(1,467 / 1,103) 0,446	(1,598 / 0,795) 0,436	(1,239 / 0,454) 0,299	(1,284 / 0,718) 0,281	(1,664 / 0,744) 0,428	
EXP.ES&ARIMA	(1,343 / 1,348) 0,664	(1,301 / 0,736) 0,407	(1,346 / 1,099) 0,446	(1,508 / 0,752) 0,533	(1,111 / 0,428) 0,234	(1,348 / 0,718) 0,281	(1,900 / 0,834) 0,428	
mv.EXP.ES Bestand	(1,332 / 1,351) 0,628	(1,212 / 0,640) 0,407	(1,136 / 1,002) 0,446	(1,364 / 0,744) 0,493	(0,893 / 0,428) 0,234	(1,341 / 0,718) 0,281	(1,593 / 0,734) 0,428	
MLP Bestand	(1,015 / 1,501) 0,664	(0,986 / 0,729) 0,467	(1,952 / 1,248) 0,364	(1,361 / 0,709) 0,468	(0,826 / 0,428) 0,234	(0,952 / 0,925) 0,230	(1,294 / 1,122) 0,392	
mv.MLP Bestand	(0,954 / 1,694) 0,660	(1,035 / 0,911) 0,353	(1,719 / 1,348) 0,371	(1,253 / 0,777) 0,450	(0,826 / 0,428) 0,234	(1,175 / 0,989) 0,205	(1,169 / 1,229) 0,432	
mv.MLP.ACF B.	(2,153 / 1,084) 0,587	(1,357 / 0,702) 0,454	(1,098 / 0,896) 0,479	(1,161 / 0,604) 0,324	(0,565 / 0,435) 0,234	(1,004 / 0,630) 0,353	(1,444 / 0,731) 0,490	
Verfahren	S8	S9	S10	S11	S12	S13		
NF1Bestand	(1,880 / 0,863) 0,475	(2,192 / 2,108) 0,802	(1,271 / 0,518) 0,306	(1,597 / 0,802) 0,564	(1,637 / 0,760) 0,461	(2,275 / 1,219) 0,623		
NF2 Bestand	(1,817 / 0,888) 0,554	(2,342 / 1,458) 0,670	(1,519 / 0,529) 0,306	(1,482 / 0,619) 0,418	(1,909 / 0,774) 0,468	(2,574 / 1,137) 0,912		
MA Bestand	(1,493 / 1,037) 0,425	(1,939 / 0,788) 0,529	(1,141 / 0,443) 0,259	(1,424 / 0,887) 0,492	(1,364 / 0,691) 0,396	(1,940 / 0,953) 0,522		
S.ES Bestand	(1,503 / 0,908) 0,403	(1,734 / 0,834) 0,529	(1,076 / 0,403) 0,234	(1,535 / 0,654) 0,471	(1,264 / 0,626) 0,367	(1,804 / 0,806) 0,482		
DT.ES Bestand	(1,438 / 1,018) 0,432	(1,811 / 0,762) 0,659	(0,940 / 0,503) 0,276	(1,396 / 0,694) 0,464	(1,295 / 0,558) 0,342	(1,900 / 0,799) 0,532		
EXP.ES Bestand	(1,435 / 1,018) 0,436	(1,515 / 0,787) 0,443	(0,918 / 0,482) 0,284	(1,331 / 0,536) 0,454	(1,364 / 0,623) 0,392	(1,893 / 0,803) 0,532		
EXP.ARIMA B.	(1,410 / 1,165) 0,360	(1,559 / 0,876) 0,467	(1,040 / 0,403) 0,276	(1,438 / 0,475) 0,414	(1,305 / 0,652) 0,418	(1,845 / 0,839) 0,508		
EXP.ES&ARIMA	(1,449 / 1,015) 0,432	(1,644 / 0,776) 0,443	(1,127 / 0,418) 0,276	(1,492 / 0,736) 0,410	(1,364 / 0,626) 0,392	(1,922 / 0,806) 0,528		
mv.EXP.ES Bestand	(1,342 / 1,018) 0,432	(1,633 / 0,672) 0,443	(1,105 / 0,414) 0,276	(1,132 / 0,644) 0,410	(1,282 / 0,626) 0,389	(1,195 / 0,794) 0,560		
MLP Bestand	(1,519 / 1,129) 0,360	(1,300 / 1,026) 0,518	(0,438 / 0,412) 0,933	(1,275 / 0,529) 0,338	(0,756 / 0,672) 0,292	(1,412 / 0,837) 0,567		
mv.MLP Bestand	(1,701 / 1,179) 0,335	(1,245 / 1,076) 0,578	(0,560 / 0,420) 0,830	(1,189 / 0,604) 0,324	(0,827 / 0,457) 0,302	(1,376 / 0,673) 0,496		
mv.MLP.ACF B.	(1,472 / 0,924) 0,515	(1,350 / 0,716) 0,446	(0,511 / 0,402) 0,869	(1,170 / 0,677) 0,407	(0,972 / 0,630) 0,410	(1,252 / 0,724) 0,614		

* Ergebnisse werden in der Form '(Fehler Trainingsmenge/ Fehler Validierungsmenge) Fehler Tesmenge' präsentiert.

Tabelle 8.30. Rang nach MLLC je Zeitreihe und Verfahren an Absatzstelle AU-2 (identisch mit Rang nach SLLC)

Verfahren	MLLC						
	S1	S2	S3	S4	S5	S6	S7
NF1Bestand	(11 / 12) 3	(12 / 11) 11	(9 / 11) 12	(11 / 11) 11	(11 / 11) 10	(12 / 11) 12	(7 / 9) 11
NF2 Bestand	(10 / 2) 11	(11 / 10) 12	(11 / 12) 11	(12 / 12) 12	(12 / 12) 12	(10 / 7) 11	(12 / 11) 12
MA Bestand	(9 / 10) 1	(8 / 3) 8	(8 / 7) 10	(10 / 10) 8	(6 / 10) 8	(11 / 9) 9	(8 / 6) 8
S.ES Bestand	(8 / 3) 12	(10 / 1) 3	(7 / 9) 8	(9 / 7) 6	(5 / 1) 1	(9 / 8) 8	(10 / 4) 7
DT.ES Bestand	(5 / 7) 6	(6 / 4) 4	(3 / 4) 3	(4 / 3) 7	(9 / 2) 9	(4 / 2) 3	(5 / 5) 10
EXP.ES Bestand	(4 / 8) 6	(7 / 9) 4	(4 / 6) 3	(5 / 6) 10	(7 / 2) 1	(5 / 2) 3	(9 / 8) 2
EXP.ARIMA B.	(7 / 6) 6	(4 / 6) 1	(6 / 4) 3	(8 / 9) 2	(10 / 9) 10	(3 / 2) 3	(6 / 3) 2
EXP.ES&ARIMA	(6 / 4) 6	(5 / 8) 4	(5 / 3) 3	(6 / 5) 9	(8 / 2) 1	(8 / 2) 3	(11 / 7) 2
mv.EXP.ES Bestand	(3 / 5) 4	(3 / 2) 4	(1 / 2) 3	(2 / 4) 5	(2 / 2) 1	(7 / 2) 3	(3 / 2) 2
MLP Bestand	(2 / 9) 6	(1 / 7) 10	(12 / 8) 1	(7 / 2) 4	(4 / 2) 1	(1 / 10) 2	(2 / 10) 1
mv.MLP Bestand	(1 / 11) 5	(2 / 12) 2	(10 / 10) 2	(3 / 8) 3	(3 / 2) 1	(6 / 12) 1	(1 / 12) 6
mv.MLP.ACF B.	(12 / 1) 2	(9 / 5) 9	(2 / 1) 9	(1 / 1) 1	(1 / 8) 1	(2 / 1) 10	(4 / 1) 9
Verfahren	S8	S9	S10	S11	S12	S13	Rank mwRang MLLC
NF1Bestand	(11 / 1) 10	(11 / 12) 12	(11 / 11) 8	(12 / 11) 12	(11 / 11) 11	(11 / 12) 11	
NF2 Bestand	(10 / 2) 12	(12 / 11) 11	(12 / 12) 9	(11 / 5) 7	(12 / 12) 12	(12 / 11) 12	
MA Bestand	(6 / 9) 5	(10 / 6) 8	(10 / 8) 2	(7 / 12) 11	(10 / 10) 8	(10 / 10) 4	
S.ES Bestand	(7 / 3) 4	(8 / 7) 7	(7 / 2) 1	(10 / 7) 10	(5 / 4) 4	(5 / 7) 1	
DT.ES Bestand	(4 / 6) 6	(9 / 3) 10	(5 / 10) 3	(4 / 9) 9	(6 / 2) 3	(8 / 4) 6	
EXP.ES Bestand	(3 / 6) 9	(4 / 5) 1	(4 / 9) 7	(3 / 3) 8	(8 / 3) 6	(7 / 5) 6	
EXP.ARIMA B.	(2 / 11) 2	(5 / 8) 5	(6 / 3) 3	(6 / 1) 6	(7 / 8) 10	(6 / 9) 3	
EXP.ES&ARIMA	(5 / 5) 6	(7 / 4) 1	(9 / 6) 3	(9 / 10) 4	(8 / 5) 6	(9 / 6) 5	
mv.EXP.ES Bestand	(1 / 6) 6	(6 / 1) 1	(8 / 5) 3	(1 / 6) 4	(4 / 5) 5	(1 / 3) 8	
MLP Bestand	(9 / 10) 2	(2 / 9) 6	(1 / 4) 12	(8 / 2) 2	(1 / 9) 1	(4 / 8) 9	
mv.MLP Bestand	(12 / 12) 1	(1 / 10) 9	(3 / 7) 10	(5 / 4) 1	(2 / 1) 2	(3 / 1) 2	
mv.MLP.ACF B.	(8 / 4) 11	(3 / 2) 4	(2 / 1) 11	(2 / 8) 3	(3 / 7) 9	(2 / 2) 10	

* Ergebnisse werden in der Form '(Fehler Trainingsmenge/ Fehler Validierungsmenge) Fehler Testmenge' präsentiert.

Tabelle 8.31. MLLC je Zeitreihe und Verfahren an Absatzstelle IN-1

Verfahren	MLLC							Rank mwMLLC
	S1	S2	S3	S4	S5	S6	S7	
NF1Bestand	(0,017 / 0,034) 0,014	(0,629 / 0,468) 0,421	(0,049 / 0,042) 0,045	(0,092 / 0,058) 0,050	(0,391 / 0,229) 0,200	(0,065 / 0,052) 0,059	(0,105 / 0,167) 0,087	
NF2 Bestand	(0,013 / 0,026) 0,014	(0,450 / 0,363) 0,326	(0,042 / 0,037) 0,039	(0,077 / 0,044) 0,037	(0,292 / 0,182) 0,161	(0,065 / 0,058) 0,077	(0,076 / 0,140) 0,079	
MA Bestand	(0,014 / 0,031) 0,013	(0,531 / 0,335) 0,304	(0,048 / 0,032) 0,030	(0,077 / 0,042) 0,036	(0,285 / 0,160) 0,142	(0,050 / 0,047) 0,080	(0,107 / 0,185) 0,067	
S.ES Bestand	(0,014 / 0,031) 0,012	(0,519 / 0,321) 0,286	(0,045 / 0,034) 0,030	(0,076 / 0,041) 0,035	(0,293 / 0,155) 0,137	(0,052 / 0,052) 0,069	(0,096 / 0,152) 0,062	
DT.ES Bestand	(0,014 / 0,026) 0,011	(0,522 / 0,280) 0,406	(0,048 / 0,041) 0,042	(0,068 / 0,035) 0,033	(0,323 / 0,129) 0,106	(0,057 / 0,082) 0,054	(0,099 / 0,148) 0,062	
EXP.ES Bestand	(0,014 / 0,026) 0,011	(0,494 / 0,305) 0,250	(0,045 / 0,040) 0,038	(0,066 / 0,036) 0,029	(0,289 / 0,123) 0,094	(0,050 / 0,066) 0,042	(0,099 / 0,151) 0,062	
EXP.ARIMA B.	(0,011 / 0,026) 0,011	(0,495 / 0,290) 0,270	(0,055 / 0,039) 0,033	(0,062 / 0,032) 0,027	(0,267 / 0,134) 0,123	(0,058 / 0,087) 0,063	(0,085 / 0,136) 0,059	
EXP.ES&ARIMA	(0,014 / 0,026) 0,011	(0,523 / 0,327) 0,356	(0,043 / 0,040) 0,038	(0,082 / 0,041) 0,032	(0,297 / 0,149) 0,138	(0,056 / 0,082) 0,052	(0,097 / 0,153) 0,070	
mv.EXP.ES Bestand	(0,010 / 0,026) 0,011	(0,478 / 0,323) 0,346	(0,039 / 0,040) 0,038	(0,064 / 0,041) 0,031	(0,240 / 0,147) 0,137	(0,050 / 0,074) 0,052	(0,090 / 0,134) 0,070	
MLP Bestand	(0,011 / 0,028) 0,010	(0,753 / 0,258) 0,355	(0,077 / 0,035) 0,044	(0,062 / 0,026) 0,023	(0,168 / 0,176) 0,091	(0,061 / 0,094) 0,066	(0,074 / 0,136) 0,061	
mv.MLP Bestand	(0,008 / 0,028) 0,010	(0,719 / 0,221) 0,354	(0,071 / 0,034) 0,038	(0,060 / 0,027) 0,025	(0,169 / 0,142) 0,073	(0,037 / 0,070) 0,080	(0,069 / 0,132) 0,053	
mv.MLP.ACF B.	(0,012 / 0,022) 0,012	(0,303 / 0,214) 0,181	(0,024 / 0,030) 0,062	(0,041 / 0,027) 0,028	(0,117 / 0,095) 0,081	(0,037 / 0,034) 0,063	(0,090 / 0,087) 0,082	
Verfahren	S8	S9	S10					
NF1Bestand	(0,085 / 0,060) 0,057	(0,069 / 0,062) 0,051	(0,089 / 0,096) 0,079					
NF2 Bestand	(0,092 / 0,053) 0,096	(0,062 / 0,059) 0,052	(0,070 / 0,080) 0,081					
MA Bestand	(0,084 / 0,052) 0,084	(0,058 / 0,063) 0,044	(0,063 / 0,086) 0,095					
S.ES Bestand	(0,091 / 0,050) 0,086	(0,054 / 0,060) 0,042	(0,064 / 0,071) 0,093					
DT.ES Bestand	(0,087 / 0,053) 0,074	(0,055 / 0,059) 0,058	(0,067 / 0,073) 0,111					
EXP.ES Bestand	(0,093 / 0,052) 0,068	(0,056 / 0,063) 0,054	(0,071 / 0,055) 0,079					
EXP.ARIMA B.	(0,089 / 0,051) 0,077	(0,055 / 0,063) 0,055	(0,047 / 0,080) 0,120					
EXP.ES&ARIMA	(0,098 / 0,048) 0,068	(0,055 / 0,065) 0,058	(0,065 / 0,070) 0,110					
mv.EXP.ES Bestand	(0,068 / 0,046) 0,069	(0,046 / 0,062) 0,058	(0,050 / 0,064) 0,105					
MLP Bestand	(0,083 / 0,067) 0,126	(0,044 / 0,057) 0,041	(0,046 / 0,073) 0,160					
mv.MLP Bestand	(0,071 / 0,054) 0,097	(0,046 / 0,056) 0,040	(0,048 / 0,072) 0,107					
mv.MLP.ACF B.	(0,040 / 0,039) 0,091	(0,028 / 0,047) 0,040	(0,038 / 0,046) 0,086					

* Ergebnisse werden in der Form '(Fehler Trainingsmenge/ Fehler Validierungsmenge) Fehler Tesmenge' präsentiert.

Tabelle 8.32. Summe der LLC Kosten (SLLC) je Zeitreihe und Verfahren an Absatzstelle IN-1

Verfahren	MLLC							Rank mwSLLC
	S1	S2	S3	S4	S5	S6	S7	
NF1Bestand	(0,758 / 1,222) 0,299	(41,521 / 16,852) 8,842	(3,216 / 1,516) 0,672	(5,535 / 2,092) 1,051	(25,779 / 8,233) 4,190	(4,299 / 1,879) 1,248	(6,957 / 6,010) 1,832	
NF2 Bestand	(0,528 / 0,943) 0,295	(29,702 / 13,079) 6,854	(2,772 / 1,321) 0,583	(4,391 / 1,598) 0,785	(19,241 / 6,534) 3,384	(4,318 / 2,074) 1,609	(5,040 / 5,030) 1,668	
MA Bestand	(0,636 / 1,132) 0,266	(35,059 / 12,042) 6,390	(3,180 / 1,148) 0,450	(4,637 / 1,494) 0,749	(18,777 / 5,742) 2,988	(3,272 / 1,700) 1,675	(7,030 / 6,652) 1,404	
S.ES Bestand	(0,636 / 1,121) 0,248	(34,260 / 11,538) 5,998	(2,951 / 1,216) 0,450	(4,583 / 1,472) 0,734	(19,350 / 5,569) 2,869	(3,440 / 1,861) 1,439	(6,337 / 5,472) 1,300	
DT.ES Bestand	(0,654 / 0,946) 0,238	(34,975 / 10,069) 8,532	(3,213 / 1,487) 0,624	(4,222 / 1,253) 0,695	(21,650 / 4,637) 2,228	(3,795 / 2,957) 1,124	(6,654 / 5,338) 1,310	
EXP.ES Bestand	(0,654 / 0,946) 0,238	(33,130 / 10,980) 5,254	(3,004 / 1,423) 0,567	(4,104 / 1,300) 0,605	(19,367 / 4,432) 1,984	(3,324 / 2,374) 0,892	(6,646 / 5,434) 1,310	
EXP.ARIMA B.	(0,485 / 0,931) 0,238	(33,164 / 10,444) 5,665	(3,687 / 1,400) 0,490	(3,789 / 1,156) 0,562	(17,897 / 4,827) 2,592	(3,858 / 3,146) 1,323	(5,679 / 4,894) 1,246	
EXP.ES&ARIMA	(0,643 / 0,942) 0,238	(35,054 / 11,768) 7,476	(2,889 / 1,452) 0,567	(5,032 / 1,490) 0,662	(19,922 / 5,371) 2,891	(3,727 / 2,957) 1,096	(6,528 / 5,520) 1,462	
mv.EXP.ES Bestand	(0,454 / 0,949) 0,238	(32,035 / 11,628) 7,261	(2,625 / 1,438) 0,567	(3,946 / 1,465) 0,655	(16,051 / 5,303) 2,873	(3,356 / 2,654) 1,096	(5,997 / 4,838) 1,462	
MLP Bestand	(0,417 / 1,010) 0,212	(44,442 / 9,280) 7,447	(4,530 / 1,271) 0,656	(3,333 / 0,947) 0,490	(9,916 / 6,351) 1,904	(3,626 / 3,366) 1,376	(4,344 / 4,905) 1,274	
mv.MLP Bestand	(0,313 / 0,995) 0,220	(42,404 / 7,955) 7,435	(4,186 / 1,218) 0,574	(3,264 / 0,965) 0,529	(9,996 / 5,105) 1,526	(2,166 / 2,515) 1,680	(4,044 / 4,757) 1,120	
mv.MLP.ACF B.	(0,472 / 0,792) 0,256	(18,173 / 7,711) 3,798	(1,422 / 1,065) 0,932	(2,269 / 0,972) 0,583	(7,042 / 3,433) 1,696	(2,201 / 1,235) 1,316	(5,407 / 3,120) 1,717	
Verfahren	S8	S9	S10					
NF1Bestand	(5,605 / 2,146) 1,202	(4,561 / 2,244) 1,062	(5,852 / 3,462) 1,652					
NF2 Bestand	(6,050 / 1,904) 2,006	(4,071 / 2,117) 1,087	(4,629 / 2,870) 1,708					
MA Bestand	(5,512 / 1,866) 1,765	(3,804 / 2,266) 0,914	(4,129 / 3,109) 1,992					
S.ES Bestand	(5,974 / 1,805) 1,801	(3,553 / 2,169) 0,882	(4,235 / 2,555) 1,945					
DT.ES Bestand	(5,833 / 1,905) 1,545	(3,670 / 2,122) 1,213	(4,488 / 2,638) 2,332					
EXP.ES Bestand	(6,230 / 1,854) 1,428	(3,766 / 2,262) 1,138	(4,776 / 1,962) 1,664					
EXP.ARIMA B.	(5,988 / 1,833) 1,626	(3,713 / 2,272) 1,145	(3,178 / 2,892) 2,528					
EXP.ES&ARIMA	(6,552 / 1,726) 1,435	(3,655 / 2,338) 1,224	(4,364 / 2,524) 2,311					
mv.EXP.ES Bestand	(4,532 / 1,642) 1,458	(3,088 / 2,231) 1,217	(3,364 / 2,316) 2,214					
MLP Bestand	(4,899 / 2,407) 2,644	(2,616 / 2,051) 0,868	(2,703 / 2,621) 3,352					
mv.MLP Bestand	(4,183 / 1,926) 2,041	(2,724 / 2,008) 0,850	(2,840 / 2,594) 2,248					
mv.MLP.ACF B.	(2,384 / 1,410) 1,917	(1,706 / 1,696) 0,844	(2,264 / 1,671) 1,813					

* Ergebnisse werden in der Form '(Fehler Trainingsmenge/ Fehler Validierungsmenge) Fehler Tesmenge' präsentiert.

Tabelle 8.33. Rang nach MLLC je Zeitreihe und Verfahren an Absatzstelle IN-1 (identisch mit Rang nach SLLC)

Verfahren	MLLC						
	S1	S2	S3	S4	S5	S6	S7
NF1Bestand	(12 / 12) 12	(10 / 12) 12	(9 / 12) 11	(12 / 12) 12	(12 / 12) 12	(11 / 4) 5	(11 / 11) 12
NF2 Bestand	(6 / 4) 11	(2 / 11) 6	(3 / 6) 8	(9 / 11) 11	(8 / 11) 11	(12 / 5) 10	(3 / 6) 10
MA Bestand	(10 / 11) 10	(9 / 10) 5	(8 / 2) 2	(10 / 10) 10	(6 / 9) 10	(3 / 2) 11	(12 / 12) 7
S.ES Bestand	(11 / 10) 8	(6 / 7) 4	(5 / 3) 1	(8 / 8) 9	(9 / 8) 7	(6 / 3) 9	(7 / 9) 4
DT.ES Bestand	(7 / 5) 3	(7 / 4) 11	(7 / 11) 9	(7 / 5) 8	(11 / 3) 5	(8 / 9) 4	(10 / 7) 5
EXP.ES Bestand	(7 / 5) 3	(4 / 6) 2	(6 / 8) 4	(6 / 6) 5	(7 / 2) 4	(4 / 6) 1	(9 / 8) 5
EXP.ARIMA B.	(3 / 2) 3	(5 / 5) 3	(10 / 7) 3	(4 / 4) 3	(5 / 4) 6	(9 / 11) 7	(4 / 4) 2
EXP.ES&ARIMA	(9 / 3) 3	(8 / 9) 10	(4 / 10) 4	(11 / 9) 7	(10 / 7) 9	(7 / 9) 2	(8 / 10) 8
mv.EXP.ES Bestand	(2 / 7) 3	(3 / 8) 7	(2 / 9) 4	(5 / 7) 6	(4 / 6) 8	(5 / 8) 2	(5 / 3) 8
MLP Bestand	(4 / 9) 1	(12 / 3) 9	(12 / 5) 10	(3 / 1) 1	(2 / 10) 3	(10 / 12) 8	(2 / 5) 3
mv.MLP Bestand	(1 / 8) 2	(11 / 2) 8	(11 / 4) 7	(2 / 2) 2	(3 / 5) 1	(2 / 7) 12	(1 / 2) 1
mv.MLP.ACF B.	(5 / 1) 9	(1 / 1) 1	(1 / 1) 12	(1 / 3) 4	(1 / 1) 2	(1 / 1) 6	(6 / 1) 11
Verfahren	S8	S9	S10	Rank mwRang MLLC			
NF1Bestand	(6 / 11) 1	(12 / 8) 6	(12 / 12) 1				
NF2 Bestand	(10 / 8) 10	(11 / 4) 7	(10 / 9) 3				
MA Bestand	(5 / 7) 7	(10 / 10) 5	(6 / 11) 6				
S.ES Bestand	(9 / 4) 8	(5 / 6) 4	(7 / 5) 5				
DT.ES Bestand	(7 / 9) 5	(7 / 5) 10	(9 / 8) 10				
EXP.ES Bestand	(11 / 6) 2	(9 / 9) 8	(11 / 2) 2				
EXP.ARIMA B.	(8 / 5) 6	(8 / 11) 9	(3 / 10) 11				
EXP.ES&ARIMA	(12 / 3) 3	(6 / 12) 12	(8 / 4) 9				
mv.EXP.ES Bestand	(2 / 2) 4	(3 / 7) 11	(5 / 3) 7				
MLP Bestand	(4 / 12) 12	(2 / 3) 3	(2 / 7) 12				
mv.MLP Bestand	(3 / 10) 11	(4 / 2) 2	(4 / 6) 8				
mv.MLP.ACF B.	(1 / 1) 9	(1 / 1) 1	(1 / 1) 4				

* Ergebnisse werden in der Form '(Fehler Trainingsmenge/ Fehler Validierungsmenge) Fehler Tesmenge' präsentiert.

Tabelle 8.34. MLLC je Zeitreihe und Verfahren an Absatzstelle IN-2

Verfahren	MLLC							Rank mwMLLC
	S1	S2	S3	S4	S5	S6	S7	
NF1Bestand	(0,018 / 0,032) 0,041	(0,309 / 0,181) 0,156	(0,036 / 0,028) 0,026	(0,031 / 0,025) 0,022	(0,176 / 0,115) 0,104	(0,151 / 0,115) 0,095	(0,195 / 0,113) 0,097	
NF2 Bestand	(0,031 / 0,027) 0,025	(0,208 / 0,282) 0,169	(0,036 / 0,029) 0,027	(0,031 / 0,026) 0,024	(0,159 / 0,120) 0,108	(0,137 / 0,151) 0,090	(0,177 / 0,173) 0,105	
MA Bestand	(0,022 / 0,030) 0,024	(0,189 / 0,157) 0,130	(0,023 / 0,023) 0,021	(0,029 / 0,019) 0,018	(0,121 / 0,095) 0,086	(0,105 / 0,087) 0,079	(0,149 / 0,105) 0,090	
S.ES Bestand	(0,024 / 0,029) 0,023	(0,166 / 0,158) 0,127	(0,023 / 0,022) 0,021	(0,030 / 0,019) 0,018	(0,113 / 0,092) 0,082	(0,101 / 0,084) 0,068	(0,126 / 0,101) 0,086	
DT.ES Bestand	(0,023 / 0,023) 0,022	(0,136 / 0,150) 0,126	(0,025 / 0,020) 0,022	(0,032 / 0,018) 0,019	(0,133 / 0,090) 0,063	(0,126 / 0,083) 0,132	(0,155 / 0,094) 0,185	
EXP.ES Bestand	(0,024 / 0,023) 0,022	(0,136 / 0,151) 0,126	(0,025 / 0,020) 0,022	(0,032 / 0,018) 0,019	(0,134 / 0,094) 0,074	(0,119 / 0,078) 0,109	(0,155 / 0,094) 0,185	
EXP.ARIMA B.	(0,020 / 0,026) 0,023	(0,146 / 0,155) 0,107	(0,023 / 0,020) 0,022	(0,034 / 0,020) 0,019	(0,140 / 0,100) 0,089	(0,150 / 0,079) 0,130	(0,146 / 0,108) 0,082	
EXP.ES&ARIMA	(0,028 / 0,023) 0,022	(0,132 / 0,145) 0,126	(0,023 / 0,020) 0,022	(0,033 / 0,020) 0,019	(0,136 / 0,100) 0,089	(0,147 / 0,079) 0,129	(0,152 / 0,092) 0,190	
mv.EXP.ES Bestand	(0,018 / 0,023) 0,022	(0,123 / 0,178) 0,107	(0,022 / 0,020) 0,022	(0,022 / 0,018) 0,019	(0,090 / 0,067) 0,055	(0,063 / 0,075) 0,173	(0,087 / 0,085) 0,202	
MLP Bestand	(0,021 / 0,024) 0,025	(0,229 / 0,158) 0,108	(0,026 / 0,022) 0,022	(0,024 / 0,015) 0,014	(0,143 / 0,088) 0,076	(0,136 / 0,081) 0,118	(0,126 / 0,090) 0,140	
mv.MLP Bestand	(0,025 / 0,022) 0,024	(0,149 / 0,209) 0,113	(0,030 / 0,022) 0,024	(0,028 / 0,017) 0,015	(0,145 / 0,063) 0,062	(0,084 / 0,082) 0,118	(0,144 / 0,108) 0,129	
mv.MLP.ACF B.	(0,022 / 0,021) 0,031	(0,178 / 0,150) 0,100	(0,016 / 0,017) 0,017	(0,017 / 0,018) 0,018	(0,078 / 0,066) 0,062	(0,053 / 0,084) 0,096	(0,084 / 0,078) 0,077	
Verfahren	S8	S9						
NF1Bestand	(0,022 / 0,018) 0,018	(0,043 / 0,031) 0,073						
NF2 Bestand	(0,032 / 0,020) 0,017	(0,044 / 0,031) 0,053						
MA Bestand	(0,021 / 0,016) 0,014	(0,041 / 0,029) 0,046						
S.ES Bestand	(0,022 / 0,016) 0,013	(0,039 / 0,022) 0,051						
DT.ES Bestand	(0,018 / 0,013) 0,010	(0,035 / 0,022) 0,040						
EXP.ES Bestand	(0,020 / 0,015) 0,012	(0,035 / 0,022) 0,039						
EXP.ARIMA B.	(0,020 / 0,016) 0,015	(0,038 / 0,022) 0,039						
EXP.ES&ARIMA	(0,020 / 0,015) 0,012	(0,035 / 0,022) 0,039						
mv.EXP.ES Bestand	(0,019 / 0,015) 0,012	(0,027 / 0,029) 0,038						
MLP Bestand	(0,014 / 0,012) 0,011	(0,026 / 0,046) 0,059						
mv.MLP Bestand	(0,011 / 0,012) 0,011	(0,029 / 0,031) 0,047						
mv.MLP.ACF B.	(0,012 / 0,013) 0,011	(0,024 / 0,023) 0,058						

* Ergebnisse werden in der Form ‚(Fehler Trainingsmenge/ Fehler Validierungsmenge) Fehler Tesmenge‘ präsentiert.

Tabelle 8.35. Summe der LLC Kosten (SLLC) je Zeitreihe und Verfahren an Absatzstelle IN-2

Verfahren	MLLC							Rank mwSLLC
	S1	S2	S3	S4	S5	S6	S7	
NF1Bestand	(1,179 / 1,167) 0,854	(20,382 / 6,533) 3,283	(2,374 / 1,012) 0,544	(2,022 / 0,904) 0,324	(11,610 / 4,147) 2,192	(9,966 / 4,122) 1,991	(12,882 / 4,068) 2,027	
NF2 Bestand	(2,046 / 0,968) 0,533	(13,709 / 10,159) 3,553	(2,368 / 1,033) 0,558	(2,073 / 0,936) 0,356	(10,497 / 4,327) 2,261	(9,070 / 5,432) 1,883	(11,662 / 6,240) 2,196	
MA Bestand	(1,433 / 1,088) 0,496	(12,446 / 5,639) 2,736	(1,504 / 0,824) 0,443	(1,888 / 0,698) 0,263	(7,961 / 3,402) 1,800	(6,898 / 3,118) 1,659	(9,816 / 3,766) 1,896	
S.ES Bestand	(1,568 / 1,034) 0,485	(10,928 / 5,699) 2,668	(1,526 / 0,796) 0,432	(2,012 / 0,691) 0,263	(7,476 / 3,298) 1,714	(6,646 / 3,010) 1,429	(8,342 / 3,632) 1,798	
DT.ES Bestand	(1,568 / 0,837) 0,467	(9,094 / 5,412) 2,639	(1,661 / 0,716) 0,457	(2,165 / 0,652) 0,281	(8,883 / 3,229) 1,325	(8,447 / 2,992) 2,776	(10,362 / 3,372) 3,885	
EXP.ES Bestand	(1,576 / 0,841) 0,467	(9,145 / 5,423) 2,650	(1,657 / 0,716) 0,457	(2,165 / 0,652) 0,281	(8,979 / 3,391) 1,548	(7,990 / 2,812) 2,286	(10,366 / 3,372) 3,885	
EXP.ARIMA B.	(1,350 / 0,923) 0,492	(9,808 / 5,574) 2,248	(1,568 / 0,716) 0,457	(2,286 / 0,713) 0,281	(9,363 / 3,586) 1,861	(10,018 / 2,852) 2,722	(9,754 / 3,902) 1,727	
EXP.ES&ARIMA	(1,876 / 0,837) 0,467	(8,837 / 5,234) 2,646	(1,543 / 0,716) 0,457	(2,179 / 0,716) 0,281	(9,127 / 3,593) 1,865	(9,836 / 2,856) 2,718	(10,151 / 3,305) 3,999	
mv.EXP.ES Bestand	(1,186 / 0,837) 0,467	(8,235 / 6,409) 2,253	(1,464 / 0,716) 0,457	(1,460 / 0,655) 0,281	(6,047 / 2,408) 1,155	(4,230 / 2,710) 3,630	(5,822 / 3,066) 4,244	
MLP Bestand	(1,253 / 0,869) 0,517	(13,509 / 5,698) 2,266	(1,554 / 0,788) 0,472	(1,424 / 0,544) 0,205	(8,423 / 3,182) 1,606	(8,042 / 2,925) 2,476	(7,421 / 3,234) 2,935	
mv.MLP Bestand	(1,464 / 0,808) 0,506	(8,792 / 7,521) 2,379	(1,798 / 0,806) 0,497	(1,670 / 0,622) 0,227	(8,563 / 2,260) 1,306	(4,929 / 2,935) 2,468	(8,510 / 3,874) 2,716	
mv.MLP.ACF B.	(1,306 / 0,749) 0,656	(10,665 / 5,394) 2,097	(0,968 / 0,601) 0,364	(0,996 / 0,630) 0,263	(4,656 / 2,365) 1,310	(3,208 / 3,020) 2,008	(5,023 / 2,808) 1,624	
Verfahren	S8	S9						
NF1Bestand	(1,454 / 0,648) 0,302	(2,498 / 1,126) 1,531						
NF2 Bestand	(2,081 / 0,731) 0,292	(2,408 / 1,098) 1,113						
MA Bestand	(1,396 / 0,572) 0,238	(2,377 / 1,046) 0,972						
S.ES Bestand	(1,453 / 0,562) 0,227	(2,241 / 0,806) 1,062						
DT.ES Bestand	(1,181 / 0,461) 0,176	(2,081 / 0,796) 0,834						
EXP.ES Bestand	(1,349 / 0,533) 0,198	(2,102 / 0,788) 0,822						
EXP.ARIMA B.	(1,373 / 0,583) 0,259	(2,202 / 0,788) 0,829						
EXP.ES&ARIMA	(1,356 / 0,536) 0,198	(2,051 / 0,785) 0,818						
mv.EXP.ES Bestand	(1,264 / 0,533) 0,198	(1,626 / 1,055) 0,790						
MLP Bestand	(0,837 / 0,425) 0,194	(1,342 / 1,655) 1,242						
mv.MLP Bestand	(0,644 / 0,432) 0,194	(1,511 / 1,109) 0,985						
mv.MLP.ACF B.	(0,749 / 0,457) 0,191	(1,289 / 0,842) 1,214						

* Ergebnisse werden in der Form '(Fehler Trainingsmenge/ Fehler Validierungsmenge) Fehler Teshmenge' präsentiert.

Tabelle 8.36. Rang nach MLLC je Zeitreihe und Verfahren an Absatzstelle IN-2 (identisch mit Rang nach SLLC)

Verfahren	MLLC						
	S1	S2	S3	S4	S5	S6	S7
NF1Bestand	(2 / 12) 12	(12 / 10) 11	(12 / 11) 11	(7 / 11) 11	(12 / 11) 11	(12 / 11) 4	(12 / 11) 5
NF2 Bestand	(12 / 9) 10	(10 / 12) 12	(11 / 12) 12	(8 / 12) 12	(11 / 12) 12	(9 / 12) 3	(11 / 12) 6
MA Bestand	(5 / 11) 7	(9 / 6) 10	(3 / 10) 3	(5 / 8) 3	(4 / 8) 8	(5 / 10) 2	(7 / 8) 4
S.ES Bestand	(9 / 10) 5	(7 / 8) 9	(5 / 8) 2	(6 / 7) 3	(3 / 6) 7	(4 / 8) 1	(4 / 7) 3
DT.ES Bestand	(7 / 4) 1	(3 / 3) 6	(8 / 2) 4	(9 / 4) 6	(5 / 5) 4	(7 / 7) 11	(9 / 5) 9
EXP.ES Bestand	(8 / 6) 1	(4 / 4) 8	(7 / 2) 4	(9 / 4) 6	(6 / 7) 5	(6 / 2) 6	(10 / 5) 9
EXP.ARIMA B.	(3 / 8) 6	(5 / 5) 2	(6 / 2) 4	(12 / 9) 6	(8 / 9) 9	(11 / 3) 10	(6 / 10) 2
EXP.ES&ARIMA	(11 / 3) 1	(2 / 1) 7	(4 / 2) 4	(11 / 10) 6	(7 / 10) 10	(10 / 4) 9	(8 / 4) 11
mv.EXP.ES Bestand	(1 / 4) 1	(1 / 9) 3	(2 / 2) 4	(2 / 6) 6	(2 / 3) 1	(2 / 1) 12	(2 / 2) 12
MLP Bestand	(4 / 7) 9	(11 / 7) 4	(9 / 7) 9	(3 / 1) 1	(9 / 4) 6	(8 / 5) 8	(3 / 3) 8
mv.MLP Bestand	(10 / 2) 8	(6 / 11) 5	(10 / 9) 10	(4 / 2) 2	(10 / 1) 2	(3 / 6) 7	(5 / 9) 7
mv.MLP.ACF B.	(6 / 1) 11	(8 / 2) 1	(1 / 1) 1	(1 / 3) 3	(1 / 2) 3	(1 / 9) 5	(1 / 1) 1
Verfahren	S8	S9	Rank mwRang MLLC				
NF1Bestand	(11 / 11) 12	(11 / 11) 12					
NF2 Bestand	(12 / 12) 11	(12 / 9) 9					
MA Bestand	(9 / 9) 9	(10 / 7) 6					
S.ES Bestand	(10 / 8) 8	(9 / 5) 8					
DT.ES Bestand	(4 / 4) 1	(5 / 4) 5					
EXP.ES Bestand	(6 / 5) 5	(7 / 2) 3					
EXP.ARIMA B.	(8 / 10) 10	(8 / 2) 4					
EXP.ES&ARIMA	(7 / 7) 5	(6 / 1) 2					
mv.EXP.ES Bestand	(5 / 5) 5	(3 / 8) 1					
MLP Bestand	(3 / 1) 3	(2 / 12) 11					
mv.MLP Bestand	(1 / 2) 4	(4 / 10) 7					
mv.MLP.ACF B.	(2 / 3) 2	(1 / 6) 10					

* Ergebnisse werden in der Form '(Fehler Trainingsmenge/ Fehler Validierungsmenge) Fehler Tesmenge' präsentiert.

