Verbindung von synthetischer Datenerzeugung und Machine Learning zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren

Zur Erlangung des akademischen Grades eines

Dr.-Ing.

Von der Fakultät Maschinenbau

der Technischen Universität Dortmund

genehmigte Dissertation von

Matthias Brüggenolte, M.Sc.

aus

Lippstadt

Tag der mündlichen Prüfung: 22.06.2026

1. Gutachter: Prof. Dr. Dr. h.c. Michael Henke

2. Gutachterin: Prof. Dr. Anne Meyer

Dortmund, 2026

Vorwort

Diese Arbeit entstand während meiner Zeit als wissenschaftlicher Mitarbeiter am Lehrstuhl für Unternehmenslogistik (LFO) an der Technischen Universität Dortmund. Die Zeit am Lehrstuhl hat diese Arbeit besonders geprägt.

Mein besonderer Dank gilt meinem Doktorvater, Herrn Prof. Dr. Dr. h. c. Michael Henke, sowie meiner Zweitprüferin, Frau Prof. Dr. Anne Meyer, für die wissenschaftliche Betreuung, das entgegengebrachte Vertrauen und die wertvollen Impulse während der Entstehung dieser Arbeit. Die Möglichkeit, eigene Fragestellungen zu entwickeln und diese in einem anwendungsorientierten Forschungskontext zu bearbeiten, war für mich von großem Wert. Ebenso hat mir die gemeinsame Zusammenarbeit und das motivierende Arbeitsumfeld viel Freude bereitet.

Darüber hinaus danke ich den Kolleginnen und Kollegen des LFO und des Fraunhofer IML, die mich auf meinem Weg begleitet haben. Die zahlreichen fachlichen Diskussionen, kritischen Rückfragen und gemeinsamen Reflexionen haben wesentlich dazu beigetragen, die Arbeit weiterzuentwickeln und zu schärfen. Mein großer Dank gilt dabei dem Forschungsbereich Supply Chain Management und Einkauf: Die kollegiale Zusammenarbeit und die vielen schönen gemeinsamen Momente waren für mich fachlich wie persönlich von großer Bedeutung. Insbesondere danke ich Lara Kuhlman für ihre große Hilfsbereitschaft und ihr wertvolles wissenschaftliches Feedback.

Ein herzlicher Dank gilt auch meinen Eltern, die mich während der Promotionszeit begleitet und unterstützt haben. Ihr Zuspruch, ihr Verständnis und die gemeinsamen Momente abseits der Arbeit haben mir den nötigen Rückhalt gegeben.

Mein größter und persönlichster Dank gilt meiner Frau Anna Brüggenolte. Ohne ihre Geduld, ihr Verständnis, ihre Unterstützung und ihren Rückhalt wäre diese Arbeit in dieser Form nicht möglich gewesen. Und meinen Töchtern, Marcia und Paulina: Ihr seid die Besten!

Zusammenfassung

Die vorliegende Arbeit untersucht die Verbindung zwischen synthetischer Datenerzeugung und Machine Learning zur kostenorientierten Auswahl dynamischer Bestellmengenverfahren. In der Unternehmenspraxis werden Heuristiken zur Bestellmengenplanung wie Silver Meal, Groff, Least Unit Cost und Part Period zwar in ERP-Systemen bereitgestellt, jedoch häufig ohne systematische Auswahl eingesetzt. Dabei beeinflusst die Wahl des Bestellmengenverfahrens die Gesamtkosten der Planung maßgeblich. Vor diesem Hintergrund birgt die Kombination aus synthetischer Datenerzeugung und Machine Learning ein großes Potenzial, um unter gegebenen System- und Umweltbedingungen das jeweils kostenoptimale Verfahren zu identifizieren.

Zu diesem Zweck wird die dynamische Bestellmengenplanung zunächst systemtheoretisch modelliert. Anschließend werden durch Simulationsexperimente umfangreiche synthetische Daten generiert. Diese Datengrundlage ermöglicht eine umfassende Analyse der Leistungsfähigkeit der vier dynamischen Bestellmengenverfahren in Abhängigkeit von den zugrundeliegenden Systemfaktoren.

Auf Basis der synthetischen Daten werden zunächst die Wirkzusammenhänge zwischen den Systemfaktoren und der kostenoptimalen Heuristik untersucht. Die Ergebnisse zeigen, dass insbesondere der Prognosefehler sowie die Kostenstruktur (Time Between Order) den größten Einfluss auf die Wahl der Heuristik ausüben. Aufbauend auf diesen Erkenntnissen sowie weiteren statistischen Analysen werden niedrigschwellige, interpretierbare Entscheidungsregeln abgeleitet, die eine kostenorientierte Auswahl der Heuristik ermöglichen.

Parallel dazu wird ein Machine Learning-basiertes Klassifikationsmodell zur Heuristikwahl entwickelt. Nach der Identifikation geeigneter baumbasierter Verfahren werden diese auf Grundlage der synthetischen Daten trainiert und anhand etablierter Gütemaße verglichen. Das Verfahren LightGBM erweist sich dabei als das leistungsfähigste Modell und ermöglicht im Vergleich zu einer zufälligen Heuristikwahl signifikante Kosteneinsparungen.

Die Übertragbarkeit der Ergebnisse wird durch Simulationsexperimente mit realen Unternehmensdaten aus zwei Anwendungsfällen validiert. Sowohl die entwickelten Entscheidungsregeln als auch das Machine Learning Modell führen zu deutlichen Kosteneinsparungen. Insgesamt leistet die vorliegende Arbeit somit einen inhaltlichen und methodischen Beitrag zur kostenorientierten Auswahl dynamischer Bestellmengenverfahren, indem systemtheoretische Modellierung, simulative Datenerzeugung, statistische Analyse und Machine Learning miteinander verknüpft werden. Darüber hinaus bietet die Arbeit Unternehmen eine fundierte Grundlage für eine kostenorientierte Heuristikwahl.

Abstract

This dissertation addresses the combination of synthetic data generation and machine learning for the cost-oriented selection of dynamic lot sizing methods. In many companies, heuristics such as Silver Meal, Groff, Least Unit Cost and Part Period are provided in ERP systems, but are often used without systematic selection. In this context, the combination of synthetic data generation and machine learning offers great potential for identifying the most cost-effective method for given system and environmental conditions.

The dynamic lot sizing is first modelled using systems theory. Subsequently, extensive synthetic data is generated using simulation experiments. The resulting data basis allows a comprehensive analysis of the performance of the four dynamic lot sizing models as a function of the system factors.

Based on the synthetic data, the interdependencies between system factors and cost-optimal heuristics are analysed. The results show that the forecast error and the time between order have the strongest influence on the choice of heuristics. Based on this and further statistical analyses, simple, interpretable decision rules are derived that enable a cost-oriented choice of heuristics.

Additionally, a machine learning-based classification model for heuristic selection has been developed. After identifying suitable tree-based methods, these methods are trained using synthetic data and compared using established evaluation metrics. The LightGBM method is shown to provide the best performance and enables significant cost savings compared to random heuristic selection.

The applicability of the results is validated using simulation experiments with real company data from two use cases. Both the decision rules and the machine learning model enable significant cost savings. Overall, this work makes a substantive and methodological contribution to the cost-oriented selection of dynamic lot sizing methods by combining system theory modelling, simulative data generation, statistical analysis and machine learning. Additionally, it enables companies to make informed, cost-oriented heuristic choices.

Inhaltsverzeichnis

Verbindung von synthetischer Datenerzeugung und Machine Learning zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ADF	Augmented Dickey-Fuller-Test
ADI	Average Demand Interval
AR	Autoregressiver Prozess
ARIMA	Autoregressives integriertes Moving-Average-Modell
ARIMAX	ARIMA-Modell mit exogenen Variablen
ARMA	Autoregressives Moving-Average-Modell
ARMAX	Autoregressives Moving-Average-Modell mit exogenen Variablen
ARX	Autoregressives Modell mit exogenen Variablen
ATH	Anteil der Top-Heuristiken
CK	Cohens Kappa
CSV	Comma-Separated Values
DGP	datengenerierender Prozess
d. h.	das heißt
DIN	Deutsches Institut für Normung
EFB	Exclusive Feature Bundling
EOQ	Economic Order Quantity
ERP	Enterprise-Resource-Planning
GAPE	Generalized Absolute Percentage Error
GOSS	Gradient-based One-Side Sampling
GR	Verfahren nach Groff
GSPE	Generalized Squared Percentage Error
i. A. a.	in Anlehnung an

i. d. R.	in der Regel
IQR	Interquartilsabstand
IT	Informationstechnik
KI	Künstliche Intelligenz
LBH	Verfahren nach Leinz-Bossert-Habe
LSTM	Long Short-Term Memory
LUC	gleitende wirtschaftliche Losgröße
MAD	mittlere absolute Abweichung
MAPE	mittlere absolute prozentuale Abweichung
ML	Machine Learning
MSE	mittlerer quadratischer Fehler
NFL	No-Free-Lunch-Theorem
NNPerioden	Nicht Null Perioden
PP	Stück-Perioden-Ausgleich-Verfahren
RMSE	Wurzel des mittleren quadratischen Fehlers
RP	Recall (korrekt klassifizierten positiven Fälle)
SM	dynamische Planungsrechnung
TBO	Time Between Order
UML	Unified Modeling Language
VarK	Variationskoeffizient
VDI	Verein Deutscher Ingenieure
WBZ	Wiederbeschaffungszeit
WW	Optimierungsverfahren nach Wagner-Whitin
XGBOOST	Extreme Gradient Boosting
z. B.	zum Beispiel

Abbildungsverzeichnis

Tabellenverzeichnis

1 Einleitung

1.1 Ausgangslage und Problemstellung

Unternehmen agieren heutzutage in einem globalen und dynamischen Geschäftsumfeld (Rolf et al. 2023, S. 7151), das von stetigem Wandel und Unsicherheit hinsichtlich der Kundennachfrage, Wiederbeschaffungszeiten und einer zunehmenden Arbeitsteilung geprägt ist. Die Dynamik und Komplexität innerhalb von Unternehmen sowie unternehmensübergreifend im Wertschöpfungsnetz steigen weiter an, unter anderem aufgrund der Forderung nach immer kürzeren Lieferzeiten bei gleichzeitigen Kostenersparnissen z. B. durch die Reduzierung von Beständen (Inman und Green 2022, S. 239). Kurzfristige und unregelmäßige Kundenbedarfe können häufig kaum prognostiziert und oft nur unzureichend befriedigt werden (Chowdhury et al. 2025, S. 2). Durch die Nachfrageunsicherheiten kommt es zu Bestandserhöhungen sowie -schwankungen (Chang et al. 2024, S. 226 ff.; Hofstra et al. 2024, S. 303). Vor diesem Hintergrund ist eine bedarfsgerechte Bevorratung unabdingbar.

In diesem komplexen und volatilen Geschäftsumfeld nimmt die Bestandsplanung eine zentrale Rolle ein (Olaniyi und Pugal 2024, S. 49). Die Bestandsplanung bildet die Grundlage für die Sicherstellung der Lieferfähigkeit und beeinflusst gleichzeitig die Unternehmenskosten erheblich (Hernawaty et al. 2024, S. 254; Kersten et al. 2017, S. 38). Die Bestandskosten setzen sich aus unterschiedlichen Kostenarten zusammen, wie beispielsweise Bestell-, Kapitalbindungs-, Lagerhaltungskosten (Schönsleben 2024, S. 498). Nach Angaben des US Census Bureau (2025) liegt der Gegenwartswert der Bestände von US-amerikanischen Unternehmen bei über 2,66 Billionen Dollar (United States Census Bureau 2025). Infolgedessen wird deutlich, dass eine effiziente Planung von Beständen ein erhebliches Kosteneinsparungspotenzial birgt.

Innerhalb der Bestandsplanung ist ein zentraler Stellhebel für die Kosteneinsparung die dynamische Bestellmengenplanung. Die Aufgabe ist es kostenoptimal die notwendigen Bedarfe nach Mengen und Zeit zu beschaffen (Thommen et al. 2016, S. 159 f.). Die Herausforderung besteht darin, ein Gleichgewicht zwischen Bestell- und Lagerhaltungskosten zu erreichen (Thommen et al. 2023, S. 195 f.; Kaul und Khurana 2022, S. 60), wobei Unsicherheiten in der Nachfrage die Komplexität der Entscheidungen deutlich erhöhen (Yuniasih und A'yuni 2024, S. 83). Um diese Herausforderungen zu lösen wurde eine Vielzahl an Methoden und Verfahren zur Bestellmengenplanung entwickelt (Glock et al. 2014, S. 40; Ma et al. 2019, S. 1491 ff.). Diese Methodenvielfalt eröffnet einerseits große Gestaltungsspielräume, erfordert andererseits aber auch eine fundierte Auswahlentscheidung. Nicht jede Methode ist für jeden Kontext gleichermaßen geeignet und die Leistungsfähigkeit einzelner Verfahren hängt stark von den zugrunde liegenden Systemfaktoren ab (Stadtler 2023, S. 11; Kian et al. 2021, S. 10; Beck et al. 2015, S. 28 ff.). Vielmehr

muss die Auswahl des richtigen Bestellmengenverfahrens in Abhängigkeit der spezifischen Systemfaktoren erfolgen (Dunke und Nickel 2021, S. 2489 ff.).

Neben klassischen Entscheidungsregeln, welche bei der Auswahl unterstützen können, bieten Machine Learning (ML) Methoden zur Klassifizierung und Auswahl von Verfahren große Potenziale (Tirkolaee et al. 2021, S. 3; Black et al. 2023, S. 200 ff.). Machine Learning ermöglicht es, Muster und Wirkzusammenhänge in großen Datenmengen zu erkennen. Insbesondere bei komplexen Zusammenhängen zwischen Einflussfaktoren kann ML einen erheblichen Mehrwert bieten, da klassische analytische Verfahren häufig an ihre Grenzen stoßen.

Eine zentrale Voraussetzung für die fundierte Auswahl der richtigen Verfahren ist die Verfügbarkeit geeigneter Daten (Hosen et al. 2024, S. 688 ff.; Rashedi 2022, S. 7 f.). In der Praxis zeigt sich jedoch häufig, dass die benötigten Daten für die Bewertung und Auswahl der richtigen Verfahren nicht in ausreichender Menge oder Qualität vorliegen. Das unterstreicht auch die Studie von BÜCHEL und ENGELS, welche aufzeigt, dass nur 29 % der Unternehmen effizient Daten bewirtschaften (Büchel und Engels 2022, S. 76). Fehlende oder unvollständige Datenbestände, inkonsistente Informationsflüsse oder die Fragmentierung von Daten über verschiedene IT-Systeme hinweg erschweren die Anwendung datengetriebener Entscheidungsmodelle erheblich (Sun et al. 2023, S. 1). Dies führt dazu, dass viele Unternehmen ohne belastbare Entscheidungsgrundlage Verfahren auswählen.

Vor diesem Hintergrund gewinnen synthetische Daten zunehmend an Bedeutung (Hecker et al. 2023, S. 162). Synthetische Daten sind künstlich generierte Datensätze, die reale Prozesse und Zusammenhänge nachbilden, ohne oder nur im geringen Umfang auf tatsächliche Unternehmensdaten angewiesen zu sein. Sie ermöglichen es, komplexe Entscheidungssituationen zu simulieren, alternative Szenarien zu untersuchen und die Leistungsfähigkeit verschiedener Verfahren unter kontrollierten Bedingungen zu bewerten (Wuttke et al. 2022, S. 229; Goyal und Mahmoud 2024, S. 1; Hollmann et al. 2025, S.319 ff.). Der Einsatz synthetischer Daten bietet damit das Potenzial, die Abhängigkeit von realen, oft schwer zugänglichen Daten zu reduzieren und gleichzeitig belastbare Entscheidungsgrundlagen zu schaffen. Das unterstreicht ebenfalls eine Studie von GARTNER, in der prognostiziert wird, dass zukünftig ein Großteil der KI-Anwendungen auf Basis von synthetischen Daten trainiert wird (Gartner 2023). Durch die Generierung synthetischer Daten kann eine umfangreiche und realitätsnahe Datengrundlage geschaffen werden, die es ermöglichen Entscheidungsmodelle unabhängig von der Datenverfügbarkeit in einem Unternehmen zu trainieren.

1.2 Zielsetzung

Angesichts der zuvor aufgezeigten Herausforderungen wird deutlich, dass die dynamische Bestellmengenplanung einen wesentlichen Einfluss auf die Gesamtkosten eines Unternehmens besitzt (Vania und Yolina 2021, S. 22). Gleichzeitig steht eine Vielzahl unterschiedlicher Verfahren zur kostenorientierten Auswahl zur Verfügung, sodass die Auswahl des geeigneten Bestellmengenverfahrens in Abhängigkeit von den jeweiligen Systemfaktoren erfolgen muss. Zwar liegen hierzu bereits grundlegende wissenschaftliche Arbeiten vor, diese sind jedoch aufgrund ihrer spezifischen Annahmen und eingeschränkten Allgemeingültigkeit nur bedingt übertragbar (Kian et al. 2021, S. 10; Fildes und Kingsman 2011, S. 483 ff.). Vor diesem Hintergrund birgt die Kombination aus synthetischer Datenerzeugung und Machine Learning ein großes Potenzial, um unter gegebenen System- und Umweltbedingungen das jeweils kostenoptimale Verfahren zu identifizieren.

Im Mittelpunkt steht dabei einerseits die Unternehmenspraxis zu befähigen, das kostenoptimale Bestellmengenverfahren auszuwählen, das unter gegebenen System- und Umweltbedingungen zu minimalen Gesamtkosten führt. Andererseits gilt es, wissenschaftlich fundiert aufzuzeigen, wie sich eine Kombination aus synthetischer Datenerzeugung, statistischer Datenanalyse und Machine Learning zur Entwicklung und Anwendung solcher Auswahlentscheidungen konzipieren und einsetzen lässt. Zur Erreichung der übergeordneten Zielsetzung wird das Forschungsvorhaben mittels vier Forschungsfragen konkretisiert und strukturiert.

Forschungsfrage 1: Welche Anforderungen an die dynamische Bestellmengenplanung charakterisieren das unternehmerische Umfeld?

Die erste Forschungsfrage beschäftigt sich im Kern mit der systematischen Erfassung, Klassifikation und Abgrenzung des Untersuchungsbereichs. Dabei werden zunächst die praxisrelevanten dynamischen Bestellmengenverfahren identifiziert. Aufbauend werden die zentralen Systemfaktoren bestimmt, welche die Auswahl und Leistungsfähigkeit der einzelnen dynamischen Bestellmengenverfahren beeinflussen. Die Ergebnisse bilden die systemischen Grundlagen und den Rahmen für die nachfolgenden Forschungsfragen.

Forschungsfrage 2: Wie lassen sich synthetische Daten zur dynamischen Bestellmengenplanung erzeugen?

Zur Gewährleistung von Reproduzierbarkeit, Generalisierbarkeit und Abbildungsgüte müssen umfangreiche und vielfältige Daten erhoben werden. Eine derart große und übertragbare Datenbasis lässt sich aus realen Systemen nur schwer erheben. Um diesen Anforderungen gerecht zu werden, werden synthetische Daten mittels Simulationsexperimenten erzeugt. Daher fokussiert die zweite Forschungsfrage die Erzeugung syntheti-

scher Daten als Grundlage für die kostenorientierte Auswahl dynamischer Bestellmengenverfahren. Dabei wird die dynamische Bestellmengenplanung systemtheoretisch beschrieben und modelliert. Die Simulationsexperimente ermöglichen es, eine Vielzahl von Faktorkombinationen unter kontrollierten Bedingungen zu untersuchen und damit eine belastbare Datengrundlage für die spätere Analyse und Modellbildung zu schaffen.

Auf Basis der synthetischen Daten beschäftigt sich die dritte Forschungsfrage mit der Entwicklung von Entscheidungsregeln zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren.

Forschungsfrage 3: Welche Entscheidungsregeln können zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren definiert werden?

Ziel dieser Forschungsfrage ist es, der Unternehmenspraxis Entscheidungsregeln bereitzustellen, welche es ermöglichen, in Abhängigkeit der Systemfaktoren das kostenoptimale dynamische Bestellmengenverfahren auszuwählen. Dabei müssen die Entscheidungsregeln vor allem niederschwellig und einfach anwendbar sein. Denn mit steigender Komplexität steigt das Risiko, dass die Entscheidungsregeln in der Praxis keine Anwendung finden, aufgrund von beispielsweise fehlender Interpretierbarkeit.

Parallel zu den Entscheidungsregeln, adressiert die vierte Forschungsfrage die Entwicklung eines ML-Modells zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren.

Forschungsfrage 4: Wie kann ein Machine Learning Modell zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren gestaltet werden?

Auf Basis der umfassenden synthetischen Daten soll untersucht werden ob und wie ein ML-Modell entwickelt werden kann. Dabei soll ein lernfähiges Klassifizierungsmodell trainiert werden, das in der Lage ist, Muster und Wirkzusammenhänge zwischen Systemfaktoren und Leistung der Bestellmengenverfahren zu erkennen und daraus Entscheidungen zur kostenorientierten Auswahl abzuleiten.

Mit der Beantwortung der vier Forschungsfrage leistet die vorliegende Arbeit einen wissenschaftlichen und anwendungsorientierten Beitrag zur **integrierten Betrachtung von synthetischen Daten und datengetriebener Entscheidungsfindung** in der Bestellmengenplanung. Durch die Verknüpfung von simulativer Datenerzeugung, statistischer Analyse und Machine Learning wird eine Entscheidungsfindung ermöglicht, welche nur im geringen Umfang auf reale Daten angewiesen ist. Ebenso wird damit der wissenschaftliche Diskurs erweitert, indem ein Vorgehen zur methodischen Kopplung von Simulation und ML vorgestellt und validiert wird.

1.3 Forschungsmethodik und Aufbau der Arbeit

Die vorliegende Dissertation lässt sich einer anwendungsorientierten Forschung zuordnen (Ulrich 1982, S. 3 ff.). Ziel ist die Entwicklung und Validierung eines Modells zur kostenorientierten Auswahl dynamischer Bestellmengenverfahren. Damit soll den Unternehmen eine kostenorientierte Auswahl, unter gegebenen System- und Umweltbedingungen, die zu minimalen Gesamtkosten führt, ermöglicht werden. Der Forschungsprozess folgt einer systematischen Vorgehensweise, die sich an der Design Science Research Methodik von PEFFERS et al. orientiert (Peffers et al. 2007, S. 54). PEFFERS et al. unterteilen den Forschungsprozess in sechs Phasen: Problemidentifizierung, Zieldefinition, Design und Entwicklung, Demonstration, Evaluation sowie Kommunikation (Peffers et al. 2007, S. 54). Dabei werden die Phasen Demonstration und Evaluation zusammengefasst, da diese eng miteinander verknüpft sind (Meyer 2024, S. 6). Die Design Science Research Methodik sieht regelmäßige Möglichkeiten für Iterationen und Anpassungen vor. Während des Forschungsprozesses wurden diese Anpassungen und Iterationen von Teilergebnissen durchgeführt. Die Arbeit gliedert sich in sieben Kapitel (vgl. Abbildung 1-1).

Im **ersten Kapitel** werden zunächst die Relevanz und Herausforderungen innerhalb der Bestellmengenplanung aufgezeigt (vgl. Kapitel 1.1). Auf dieser Basis wird die Zielsetzung beschrieben und mittels der vier Forschungsfragen strukturiert und konkretisiert (vgl. Kapitel 1.2).

Im **zweiten Kapitel** werden die konzeptionellen Grundlagen und der Stand der Forschung beschrieben. Zunächst werden die theoretischen Grundlagen der Bestandsplanung erläutert, bevor die relevanten Teilbereiche Bedarfsplanung (vgl. Kapitel 2.1), Sicherheitsbestandsplanung (vgl. Kapitel 2.2) und Bestellplanung (vgl. Kapitel 2.3) systematisch aufgearbeitet werden, um die Anforderungen an den Stand der Forschung zu erheben. Im Anschluss erfolgt eine kritische Analyse der bestehenden dynamischen Bestellmengenverfahren sowie eine Einordnung des Forschungsstands zur kostenorientierten Auswahl der Verfahren (vgl. Kapitel 2.4).

Das **dritte Kapitel** behandelt die Generierung von synthetischen Daten für die dynamische Bestellmengenplanung. Im ersten Schritt werden die theoretischen Grundlagen zur Simulation und Modellierung logistischer Systeme dargestellt (vgl. Kapitel 3.1). Anschließend erfolgt die Modellierung des definierten Systems sowie die Generierung der synthetischen Daten zur dynamischen Bestellmengenplanung durch die Simulationsexperimente (vgl. Kapitel 3.2). Die synthetischen Daten bilden die Grundlage für die nachfolgenden Analysen und Entwicklungen.

Im **vierten Kapitel** werden Entscheidungsregeln zur kostenorientierten Auswahl dynamischer Bestellmengenverfahren abgeleitet. Aufbauend auf den im dritten Kapitel gene-

rierten Daten erfolgt eine Untersuchung der relevanten Systemfaktoren und deren Einfluss auf die Auswahl geeigneter Verfahren (vgl. Kapitel 4.1 und Kapitel 4.2). Anschließend erfolgt die differenzierte Analyse nach den systematischen Fehlertypen *Rauschen*, *Über*- und *Unterschätzung* (vgl. Kapitel 4.3). Auf Basis der Analysen werden Entscheidungsregeln entwickelt, die eine situative und kostenorientierte Auswahl der Verfahren ermöglichen (vgl. Kapitel 4.4).

Das **fünfte Kapitel** widmet sich der Entwicklung eines Machine Learning Modells zur kostenorientierten Auswahl dynamischer Bestellmengenverfahren. Zunächst werden die Grundlagen zum ML erläutert (vgl. Kapitel 5.1) und die relevanten Klassifizierungsmodelle identifiziert (vgl. Kapitel 5.2). Anschließend werden ausgewählte Klassifikationsmodelle implementiert, trainiert (vgl. Kapitel 5.3) und hinsichtlich ihrer Leistungsfähigkeit verglichen (vgl. Kapitel 5.4).

Im **sechsten Kapitel** erfolgt die Validierung der Ergebnisse. Dabei werden die Entscheidungsregeln sowie das ML-Modell anhand von zwei Anwendungsfällen (vgl. Kapitel 6.1.1 und Kapitel 6.2.1), basierend auf realen Unternehmensdaten, überprüft und deren Anwendbarkeit in praktischen Szenarien bewertet (vgl. Kapitel 6.1.2 und Kapitel 6.2.2).

Das **siebte Kapitel** fasst die zentralen Ergebnisse der Arbeit zusammen (vgl. Kapitel 7.1) und diskutiert kritisch die Limitationen (vgl. Kapitel 7.2). Abschließend werden die wissenschaftlichen und praktischen Implikationen der entwickelten Ergebnisse reflektiert und ein Ausblick gegeben (vgl. Kapitel 7.3).

Die Arbeit folgt damit einem stringenten Aufbau, der von den theoretischen Grundlagen über die konzeptionelle und methodische Entwicklung bis zur Validierung reicht und so eine nachvollziehbare Brücke zwischen Forschung und Praxis im Kontext der dynamischen Bestellmengenplanung schlägt.

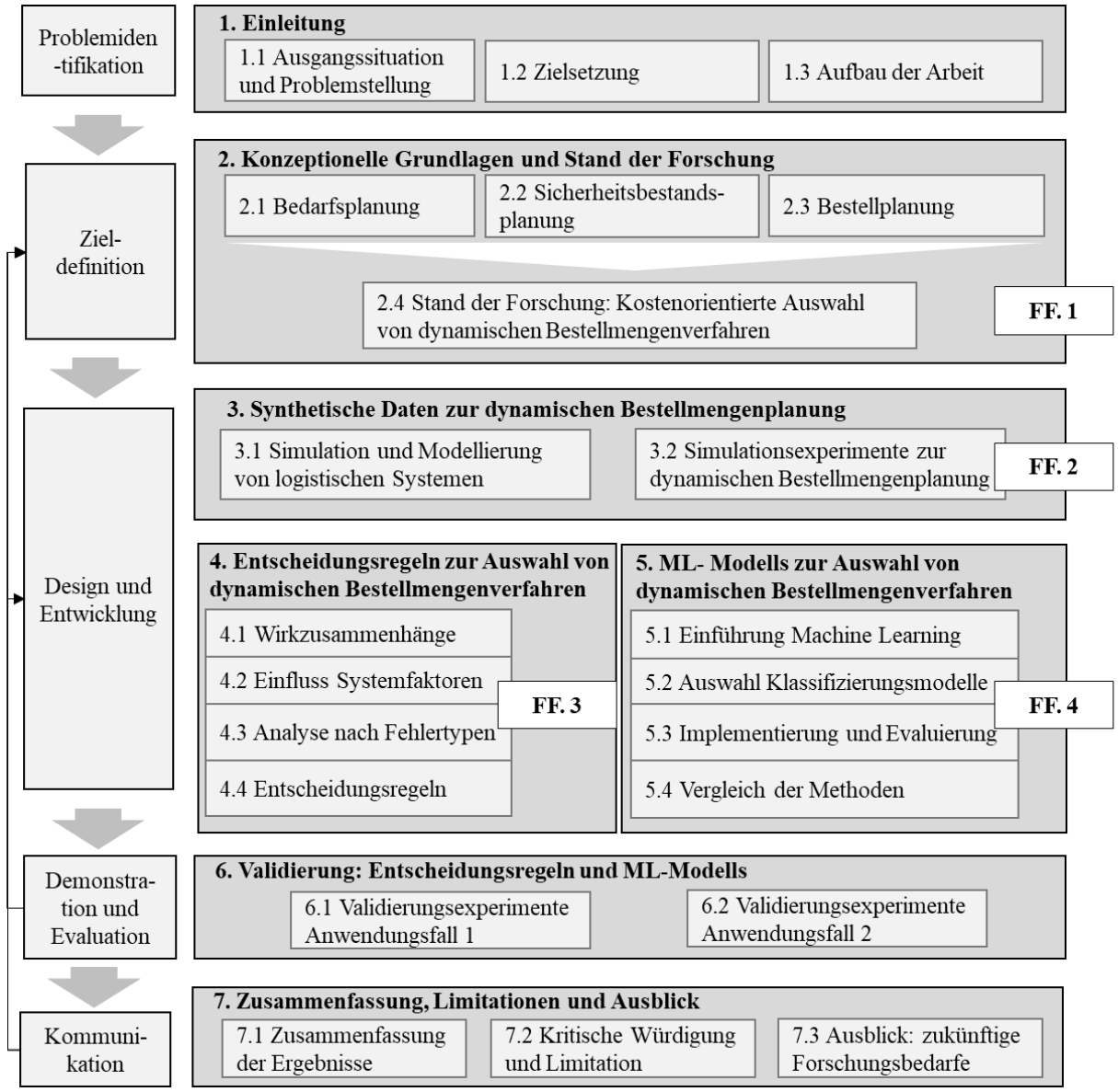

Abbildung 1-1: Aufbau der Arbeit i.A.a. (Peffers et al. 2007, S. 54)

2 Konzeptionelle Grundlagen und Stand der Forschung

In diesem Kapitel werden die konzeptionellen Grundlagen für die vorliegende Arbeit beschrieben sowie der Stand der Forschung zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren dargelegt. Zunächst erfolgt die Definition der Bestandsplanung und Abgrenzung des zugrunde liegenden Untersuchungsbereichs. Anschließend werden die Grundlagen der Bestandsplanung anhand der drei Teilbereiche *Bedarfsplanung* (vgl. Kapitel 2.1), *Sicherheitsbestandsplanung* (vgl. Kapitel 2.2) sowie *Bestellplanung* (vgl. Kapitel 2.3) erläutert. Auf dieser Basis werden die relevanten Verfahren zur dynamischen Bestellmengenplanung identifiziert und die entscheidungsrelevanten Anforderungen zur kostenorientierten Auswahl abgeleitet. Anhand der Anforderungen wird der Forschungsstand in diesem Bereich kritisch reflektiert, um den Beitrag der vorliegenden Arbeit im wissenschaftlichen Diskurs einzuordnen (vgl. Kapitel 2.4). Insgesamt adressiert damit das Kapitel 2 die erste Forschungsfrage: „*Welche Anforderungen an die Bestellmengenplanung charakterisieren das unternehmerische Umfeld?*“

Abgrenzung des Untersuchungsbereichs: Bestandsplanung

Die Bestandsplanung stellt in der betrieblichen Praxis eine Querschnittsfunktion über unterschiedliche Unternehmensbereiche hinweg dar, sodass beispielsweise in der Produktion Losgrößen- und Pufferbestände (Schmidt und Nyhuis 2021, S. 121 f.), in der Beschaffung Rohwaren- und Sicherheitsbestände (Schönsleben 2016, S. 459) oder in der Instandhaltung Ersatzteilbestände (Pawellek 2013, S. 42) geplant werden. Die vorliegende Arbeit zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren lässt sich an der Schnittstelle zwischen der Bestandsplanung und der operativ-taktischen Beschaffung verorten (Alard 2002, S. 46).

Die Bestandsplanung wird in der Literatur sowie in der Unternehmenspraxis unterschiedlich definiert. Dies bezieht sich sowohl auf die Begrifflichkeit selbst wie auch auf die jeweiligen Aufgabenbereiche. Aus diesem Grund werden nachfolgend zunächst einige zentrale Definitionen der Bestandsplanung sowie die entsprechenden Aufgaben aufgezeigt und diskutiert, um anschließend die Bestandsplanung als Grundlage und Rahmen dieser Arbeit zu definieren.

Nach STÖLZLE wird die Bestandsplanung als Teilbereich des übergeordneten Bestandsmanagements verstanden und stellt vor allem die kurz- und mittelfristige Planung und Steuerung von Beständen in den Mittelpunkt (Stölzle et al. 2004, S. 43 ff.). Dabei unterteilt sich die Bestandsplanung in die drei Aufgabenbereiche *Bedarfsplanung*, *Planung der Sicherheitsbestände* sowie *Bestellplanung*. Die Aufgabe der Bedarfsplanung ist die Identifikation von Materialbedarfen für die Beschaffung (Stölzle et al. 2004, S. 61). Die Bestellplanung oder auch Bestandssteuerung genannt, beschäftigt sich mit der Disposition

der Bestände. Damit die identifizierten Bedarfe zum richtigen Zeitpunkt, in der richtigen Menge, am richtigen Ort zur Verfügung stehen. Dazu werden unterschiedliche Dispositionsverfahren herangezogen, um die Bestellzeitpunkte sowie die Bestellmengen festzulegen (Stölzle et al. 2004, S. 90 f.). Als Ausgleichsfunktion gegen Schwankungen in der Kundennachfrage oder im Materialfluss gehört die Planung der Sicherheitsbestände ebenso zu den Aufgaben (Stölzle et al. 2004, S. 15).

SCHUH definiert ebenfalls die Bestandsplanung als einen Teilbereich des Bestandsmanagements. Das übergeordnete Ziel ist die Gewährleistung einer hohen Lieferfähigkeit, bei möglichst geringen Lagerkosten. Die Teilaufgaben der Bestandsplanung werden nach SCHUH durch die *Bedarfsermittlung*, die *Festlegung der Bevorratungsebenen* sowie die *Festlegung der Dispositionsparameter* definiert. Die Aufgaben der Bedarfsermittlung sind weitestgehend analog zu STÖLZLE zu verstehen. Die Wahl der Bevorratungsebene legt den Fertigstellungsgrad eines Artikels fest. Die Festlegung der Dispositionsparameter umfasst die Auslegung des Sicherheitsbestands, des Höchstbestands sowie des Lieferservicegrads. (Schuh 2006, S. 64 ff.)

GLEIßNER & FEMERLING erweitern die Begrifflichkeit Bestandsplanung, um die Steuerung (Bestandsplanung und -steuerung) (Gleißner und Femerling 2008, S. 146). Die Aufgaben unterteilen sich dabei in *Dispositionsverfahren*, *Lagerhaltungsstrategien* sowie *Planung der Sicherheitsbestände*. Die Dispositionsverfahren legen die mengenmäßige Einteilung sowie die terminierte Zuweisung der internen Aufträge zu Ressourcen fest. Dabei wird zwischen *verbrauchsgesteuerter*, *stochastischer* und *programmgesteuerter* Disposition unterschieden. Die verbrauchsgesteuerte Disposition ermittelt den Materialbedarf basierend auf den vergangenen Verbrauchswerten. Darüber hinaus stellt die stochastische Disposition ebenfalls eine verbrauchsgesteuerte Disposition dar, wobei auf Basis vergangener Verbrauchswerte Bedarfsprognosen erstellt werden. Werden Bestellungen aufgrund des Produktionsprogramms ausgeführt, wird von der programmgesteuerten Disposition gesprochen. Demzufolge sind die Aufgaben der Dispositionsverfahren ähnlich zu den Aufgaben der Bedarfsermittlung nach STÖLZLE und SCHUH zu verstehen. Anschließend kommen die Lagerhaltungsstrategien (Bestellpolitiken) zum Einsatz. Dabei werden die Parameter für die Bestellmengen und -zeitpunkte definiert. (Gleißner und Femerling 2008, S. 146 ff.)

Demgegenüber definieren MEYER & SANDER, neben der Bedarfs- und der Beschaffungsplanung, die Bestandsplanung als ein Teilbereich der Materialdisposition. Dabei konzentriert sich die Bestandsplanung auf die kurzfristige und wirtschaftliche Planung der Versorgungssicherheit des Unternehmens. Die Aufgaben der Bedarfsplanung und Bestellplanung werden äquivalent zu den vorherigen Autoren verstanden. (Meyer und Sander 2009, S. 19 f.) Wie auch MEYER & SANDER definieren mehrere Autoren die Materi-

aldisposition als den übergeordneten Bereich der Bestandsplanung. Dabei werden im Wesentlichen alle notwendigen Aufgaben, von der Bedarfsidentifikation bis hin zur Planung der Bestellmengen und -zeitpunkte sowie Sicherheitsbestände verstanden. Diese Definitionen unterscheiden sich jedoch häufig in den verwendeten Begrifflichkeiten und in der Anzahl sowie im Umfang der jeweiligen Teilaufgaben. (Tiedtke 2007, S. 290; Melzer-Ridinger 2009, S. 179 ff.; Härdler 2012, S. 218 f.)

Aus den zuvor aufgezeigten Definitionen geht hervor, dass häufig gleiche Aufgaben sowie ein ähnliches Vorgehen innerhalb der Bestandsplanung beschrieben wird. Im ersten Schritt erfolgt die **Bedarfsplanung**, wobei die relevanten Bedarfe identifiziert werden. Aufbauend werden die Planung der Sicherheitsbestände und die Bestellplanung durchgeführt. Die Planung der **Sicherheitsbestände** erfolgt, um Fehlmengen (-kosten), durch beispielsweise Störungen oder Schwankungen in den betrieblichen Abläufen zu vermeiden. Innerhalb der **Bestellplanung** werden die erforderlichen Bedarfe hinsichtlich Menge und Zeit wirtschaftlich beschafft. Dabei kommen unterschiedliche Dispositionsregeln (Bestellpolitiken) zum Einsatz. Demzufolge wird als Grundlage dieser Arbeit die Bestandsplanung durch die drei Aufgabenbereiche *Bedarfsplanung*, *Sicherheitsbestandsplanung* und *Bestellplanung* definiert. Diese aufgabenorientierte Definition der Bestandsplanung zeigt ebenfalls die nachfolgende Abbildung 2-1.

Abbildung 2-1: Aufgabenbereiche der Bestandsplanung

Zur Beschreibung der Wirkzusammenhänge sowie Schnittstellen der drei Aufgabenbereiche eignet sich die sogenannte Sägezahnkurve. Die Sägezahnkurve stellt ein grundlegendes Modell der Bestandsplanung dar, wobei die Kurvenform einem Sägeblatt ähnelt. In diesem Modell wird die Menge (Bestandshöhe) über die Zeit idealisiert abgebildet. Eine beispielhafte Sägezahnkurve zeigt die nachfolgende Abbildung 2-2 auf.

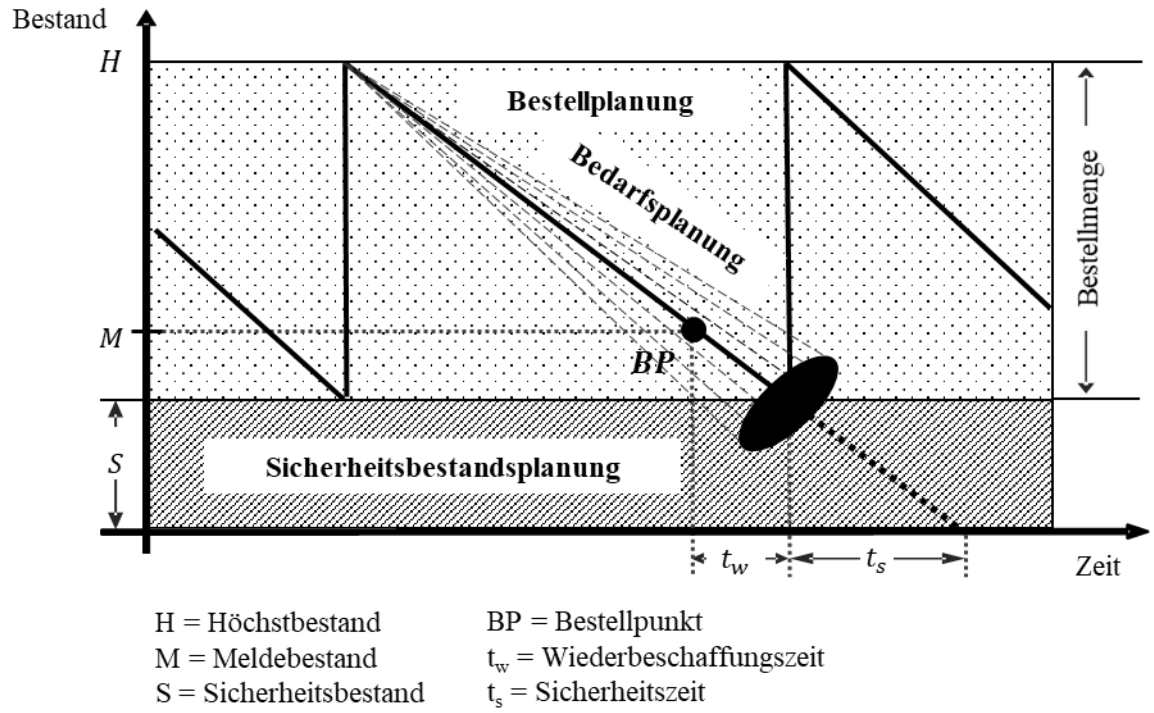

Abbildung 2-2: Sägezahnkurve i. A. a. (REFA 1985, S. 131)

Zunächst erfolgt ein Lagerzugang auf Basis einer Bestellung, wodurch das Lager beispielsweise bis zum Höchstbestand H aufgefüllt wird. Anschließend werden die Materialien oder Produkte über die Zeit verbraucht. Innerhalb des Modells wird der Verbrauch idealisiert und vereinfacht abgebildet, indem die Lagerabgänge kontinuierlich und linear über die Zeit erfolgen. Die Bedarfsunsicherheit ist vereinfacht durch die unterschiedlichen Lagerabgangsgeraden dargestellt (vgl. Abbildung 2-2). Erreicht der Bestand den Bestellpunkt BP, so wird eine neue Bestellung ausgelöst. Der Bestellpunkt ergibt sich durch die Zeitspanne für die noch ausreichend Materialien oder Produkte im Lager zur Verfügung stehen und aus der Wiederbeschaffungszeit t_w, sodass rechtzeitig ein Wareneingang erfolgt. Wobei die Wiederbeschaffungszeit, durch die Zeitdauer von der Bestellauslösung bis zum Wareneingang definiert ist (Schmidt et al. 1993, S. 239 f.). Die Bestandshöhe beim Bestellpunkt wird als Meldebestand M bezeichnet. In der Unternehmenspraxis sind nur in Ausnahmefällen idealisierte sowie kontinuierliche Lagerabgänge zu beobachten. Vielmehr sind die Lagerabgänge schwankend und mit Unsicherheiten behaftet. Aus diesem Grund werden die Sicherheitsbestände S benötigt, um Fehlmengen (-kosten) zu vermeiden. Sollten größere Lagerabgänge als geplant auftreten, dient der Sicherheitsbestand als Puffer. Die Zeitspanne, in der ein Sicherheitsbestand den Materialverbrauch deckt, wird Sicherheitszeit T_s genannt. (Ihme 2006, S. 237 f.; Wiendahl 2014, S. 311 ff.)

Zur Gewährleistung eines bedarfsgerechten Lieferservicegrads (vgl. Kapitel 2.2) mittels möglichst geringer Kosten (vgl. Kapitel 2.3) wurden eine Vielzahl unterschiedlicher Vorgehensweisen und Methoden zur Planung der drei Aufgabenbereiche entwickelt, welche in den nachfolgenden drei Unterkapitel näher erläutert werden.

2.1 Bedarfsplanung

Ein Bedarf beschreibt die Art und Menge eines Materials oder Produkts, das zur Herstellung von Produkten oder für einen Kundenbedarf erforderlich ist (Meyer und Sander 2009, S. 21) und stellt somit die Grundlage der Bedarfsplanung dar. Im Allgemeinen kann der Bedarf hinsichtlich *Ursprungs- und Erzeugnisebene* sowie der *Berücksichtigung von Lagerbeständen* klassifiziert werden. Nach Ursprung und Erzeugnisebene unterteilt sich der Bedarf in *Primär-, Sekundär-* und *Tertiärbedarf*. Auf der anderen Seite wird zwischen dem *Brutto-* und *Nettobedarf*, unter Berücksichtigung von Lagerbeständen unterschieden. Die Klassifizierung der Bedarfsarten zeigt die nachfolgende Abbildung 2-3 auf.

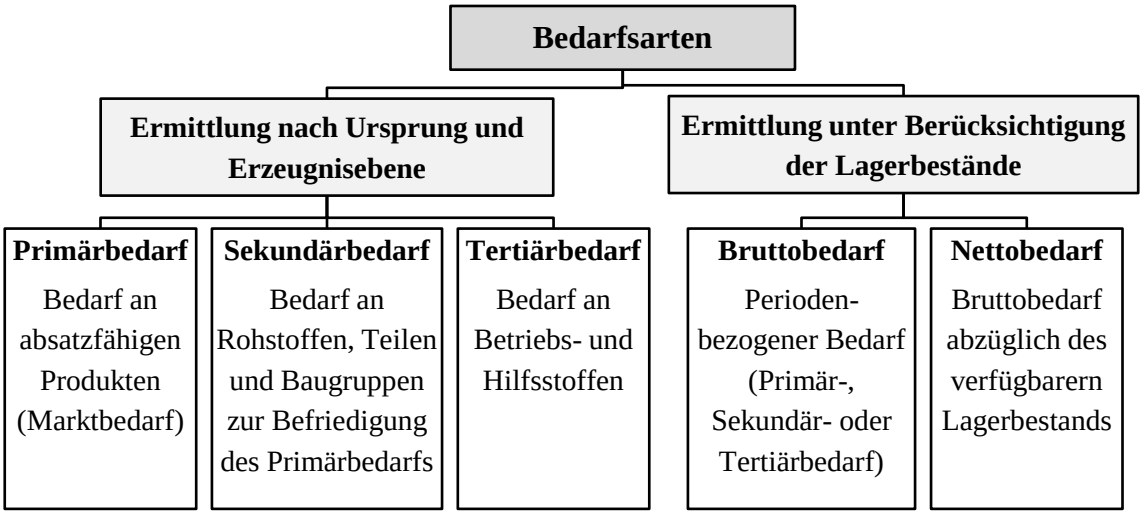

Abbildung 2-3: Unterteilung der Bedarfsarten i.A.a. (Hartmann 2005, S. 278)

Der **Primärbedarf** bezeichnet den Bedarf an verkaufsfähigen Produkten. Folglich entspricht das dem Marktbedarf, welcher sich beispielsweise aus dem Absatzplan ableiten lässt. Demgegenüber setzt sich der **Sekundärbedarf** aus den Materialien, die für die Produktion oder Montage des Primärbedarfs benötigt werden, zusammen. Der Bedarf an Hilfs- und Betriebsstoffen, welcher für die Herstellung des Primär- und Sekundärbedarfs benötigt wird sowie der Bedarf an Verschleißwerkzeugen wird als **Tertiärbedarf** bezeichnet. (Westkämper 2006, S. 184; Huber und Laverentz 2019, S. 43)

Darüber hinaus kann der Materialbedarf unter Berücksichtigung der Lagerbestände in Brutto- und Nettobedarf unterschieden werden. Der **Bruttobedarf** (Nachfrage) bezeichnet den periodenspezifischen Bedarf ohne Berücksichtigung von Lager- und Umlaufbeständen. Zur Ermittlung des Bruttobedarfs stehen verschiedene Verfahren der Bedarfsermittlung zur Verfügung, die im weiteren Verlauf des Kapitels näher erläutert werden. Der **Nettobedarf** stellt den tatsächlichen Bedarf einer Periode dar und leitet sich aus dem Bruttobedarf ab, indem vom Bruttobedarf bzw. der Nachfrage der verfügbare Lagerbestand abgezogen wird. (Schulte 2001, S. 114) In diesem Rahmen gilt es innerhalb der Bedarfsplanung, die zuvor aufgezeigten Bedarfe hinsichtlich Art, Menge und Zeit zu ermitteln.

2.1.1 Vorgehensweisen zur Bedarfsermittlung

Zur Bestimmung des Bedarfs nach Menge und Zeit können unterschiedliche Vorgehensweisen verwendet werden, welche *deterministische*, *heuristische* und *stochastische* Bedarfsermittlung genannt werden (Klaus et al. 2012, S. 46 f.).

Die **deterministische Bedarfsermittlung** wird auf Grundlage von fest eingeplanten Kunden-, Lager- oder Produktionsaufträgen durchgeführt. Folglich wird diese Vorgehensweise auch als bedarfsgesteuerte, auftragsorientierte oder programmgebundene Bedarfsermittlung bezeichnet. (Schulte 2001, S. 129) Charakteristisch für die deterministische Bedarfsplanung ist, dass die Planung auf fest definierten Informationen basiert und keine Unsicherheiten hinsichtlich der Bedarfe zulässt. Demzufolge eignet sich die deterministische Bedarfsermittlung nur bei Aufträgen, die sich in der Zukunft nicht ändern. Dieses Szenario ist in der Praxis, aufgrund von Schwankungen der Kundenbedarfe selten anzutreffen. (Seeck 2010, S. 48)

Eine weitere Möglichkeit zur Planung der Bedarfe ist die **heuristische Bedarfsermittlung**. Die heuristische Vorgehensweise erfolgt häufig auf Grundlage von subjektiven Schätzungen und eignet sich zur Planung von Bedarfen, bei denen weder konkrete Kunden-, Produktionsaufträge oder Vergangenheitsdaten vorliegen. Dies ist der Fall, wenn keine historischen Verbrauchsdaten (z. B. bei Neuprodukten) vorliegen, zukünftig starke Veränderung im Absatz erwartet werden oder wenn der Bedarfsverlauf völlig unregelmäßig und nicht vorhersehbar ist. (Seeck 2010, S. 62 f.; Wannenwetsch 2010, S. 57)

Im Rahmen der Schätzung wird zwischen der analogen und intuitiven Methode unterschieden. Die analoge Schätzung schätzt den Bedarf anhand vergleichbarer Produkte, z. B. anhand eines Vorgängerprodukts oder einer Produktvariante. Liegen keine Bedarfsdaten zu vergleichbaren Produkten vor, erfolgt eine intuitive Schätzung, häufig auf Basis von Kundenumfragen, Expertenschätzungen oder Delphi-Umfragen (Ross 2018, S. 201).

Auf Basis der hohen Fehleranfälligkeit sollte die heuristische Bedarfsermittlung nur verwendet werden, wenn keine oder nicht ausreichende Daten zur Verfügung stehen. (Hartmann 2005, S. 335 f.)

Die **stochastische Bedarfsermittlung** stellt die am häufigsten eingesetzte Vorgehensweise zur Bestimmung von Bedarfen dar. Hierbei wird ausgehend von historischen Lagerabgängen oder Kundenaufträgen der zukünftige Bedarf prognostiziert. (Seeck 2010, S. 49) Die Prognose wird nach TEMPELMEIER durch die Extrapolation der vergangenen Bedarfswerte in die Zukunft, mittels unterschiedlichen Verfahren definiert (Tempelmeier 2016, S. 53). SCHNEEWEIß erweitert diese Definition, indem sämtliche Daten berücksichtigt werden müssen, die einen Aufschluss über künftige Bedarfe geben (Schneeweiß 1981, S. 87). Demzufolge stellen die Datengrundlage hinsichtlich der vergangenen Bedarfe und die Auswahl des richtigen Prognoseverfahrens die zentralen Faktoren für die Prognosegüte dar.

2.1.2 Zeitreihenanalyse

In der Unternehmenspraxis sowie in den gängigen ERP-Systemen ist die Nachfrageprognose auf Basis von Zeitreihen das am häufigsten verwendete Vorgehen zur Bestimmung der zukünftigen Nachfrage (Cattani et al. 2011, S. 492 ff.). Die Zeitreihendaten bestehen aus Beobachtungen zu einer oder mehreren Variablen, die im Zeitverlauf gesammelt wurden. Folglich wird die historische Nachfrage als eine zeitlich geordnete Folge von diskreten Beobachtungen dargestellt (Das 2019, S. 16). Charakteristisch für Zeitreihen ist, dass zu jedem Zeitpunkt genau eine Beobachtung vorliegt. Die Zeitspanne zwischen zwei beobachteten Werten wird als *Lag* bezeichnet (Schlittgen und Streitberg 2001, S. 100). Die Zeitachse wird dabei in gleich große Zeitabschnitte (Perioden) unterteilt. Das zeitliche Aggregationsniveau einer Periode wird in der Praxis häufig auf Tagesbasis gewählt (Bantel 2014, S. 6). Die nachfolgende Abbildung zeigt eine beispielhafte monatliche Nachfrage als diskrete Zeitreihe, wobei die Nachfrage über die Perioden (Tage) abgebildet ist.

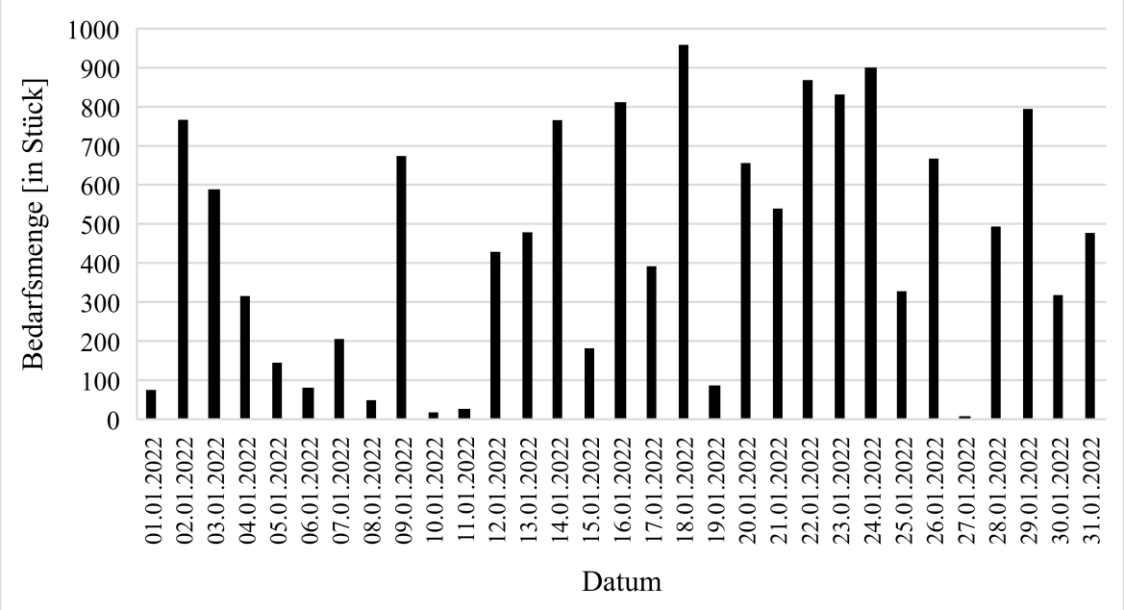

Abbildung 2-4: Diskrete Zeitreihe einer Nachfrage

Für die weitere Betrachtung werden nachfolgend zunächst die zentralen statistischen Kennzahlen zur Bedarfsplanung vorgestellt. Das **arithmetische Mittel** $\bar{x}$ einer Zeitreihe von n Werten $x_1, \dots, x_n$ ist definiert als (Eckstein 2016, S. 42):

$$\bar{x} = \frac{1}{n} * \sum_{i=1}^{n} x_i \tag{2-1}$$

Der **Erwartungswert** $E(X)$ einer diskreten Zufallsvariablen X mit den Werten $x_1, \dots, x_n$ und $P(X = x_i)$ als Eintrittswahrscheinlichkeit des Wertes x_i ergibt sich aus (Mittag 2017, S. 166):

$$E(X) = \sum_{i=1}^{n} x_i \, P(X = x_i) \tag{2-2}$$

Die **Varianz** σ^2 einer Zeitreihe von n Werten $x_1, \dots, x_n$ ist definiert als (Alicke 2005, S. 28):

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{2-3}$$

Die **Standardabweichung** σ einer Zeitreihe von n Werten $x_1, \dots, x_n$ ergibt sich aufbauend auf der Varianz zu (Alicke 2005, S. 28):

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2-4}$$

Der **Variationskoeffizient** $VarK$ einer Zeitreihe von n Werten $x_1, \dots, x_n$ ist definiert als der Quotient der Standardabweichung und des arithmetischen Mittelwerts (Alicke 2005, S. 29)

$$VarK = \frac{\sigma}{\bar{x}} \tag{2-5}$$

Das **durchschnittliche Nachfrageintervall** ADI (eng. Average inter-Demand Intervall) beschreibt das durchschnittliche Intervall zwischen zwei positiven Nachfragen einer Zeitreihe und berechnet sich wie folgend (Kalaya et al. 2019, S. 687):

$$ADI = \frac{Anzahl\ der\ Perioden}{Anzahl\ der\ Perioden\ mit\ Nachfrage} \tag{2-6}$$

Nach SYNTETOS et al. lassen sich die (Nachfrage-) Zeitreihen in vier Gruppen klassifizieren, *gleichmäßig* (smooth), *erratisch* (erratic), *intermittierend* oder *sporadisch* (intermittent) und *geklumpt* (lumpy). Die Unterteilung der vier Kategorien erfolgt anhand der Kriterien *Average inter-Demand Interval* (ADI) und *quadrierter Variationskoeffizient* ($VarK^2$). (Syntetos et al. 2005, S. 500; Boylan et al. 2008, S. 474 f.) Die Berechnung des Variationskoeffizienten $VarK$ und des Average inter-Demand Interval (ADI) zeigen die Formeln (2-5) und (2-6).

Die **gleichmäßige Nachfrage** zeichnet sich durch eine geringe Schwankung aus. Ebenso gibt es nur selten Perioden in denen keine Nachfrage auftritt. Hinsichtlich der beiden Kriterien ist die gleichmäßige Nachfrage durch einen ADI-Wert kleiner als 1,32 und einen

$Vark^2$-Wert kleiner als 0,49 definiert. Demgegenüber besitzt die **sporadische Nachfrage** eine Vielzahl an Perioden ohne Nachfrage ($ADI > 1{,}32$). Tritt jedoch eine Nachfrage auf, so sind die Mengenschwankungen gering ($Vark^2 < 0{,}49$). Die **erratische Nachfrage** folgt einer regelmäßigen Nachfrageverteilung ($ADI < 1{,}32$), jedoch mit deutlichen Schwankungen in der Menge ($Vark^2 > 0{,}49$). Abschließend ergibt sich die **geklumpte Nachfrage**, wobei die Nachfrage unregelmäßig ($ADI > 1{,}32$), mit stark schwankenden Nachfragemengen ($Vark^2 > 0{,}49$) ist. Resultierend sind die Grenzwerte für das Inter-Demand Intervall (ADI) mit 1,32 sowie für den quadrierten Variationskoeffizienten ($Vark^2$) mit 0,49 definiert, sodass sich die nachfolgenden vier Kategorien mit den entsprechenden Grenzwerten ergeben: (Ramaekers und Janssens 2014, S. 111):

- Glatte Nachfrage: ADI < 1,32; $VarK^2$ < 0,49
- Erratische Nachfrage: ADI < 1,32; $VarK^2$ > 0,49
- Sporadische Nachfrage: ADI > 1,32; $VarK^2$ < 0,49
- Geklumpte Nachfrage: ADI > 1,32; $VarK^2$ > 0,49

Die nachfolgende Abbildung 2-5 zeigt, anhand von vier beispielhaften Nachfrageverläufen, die unterschiedlichen Kategorien auf.

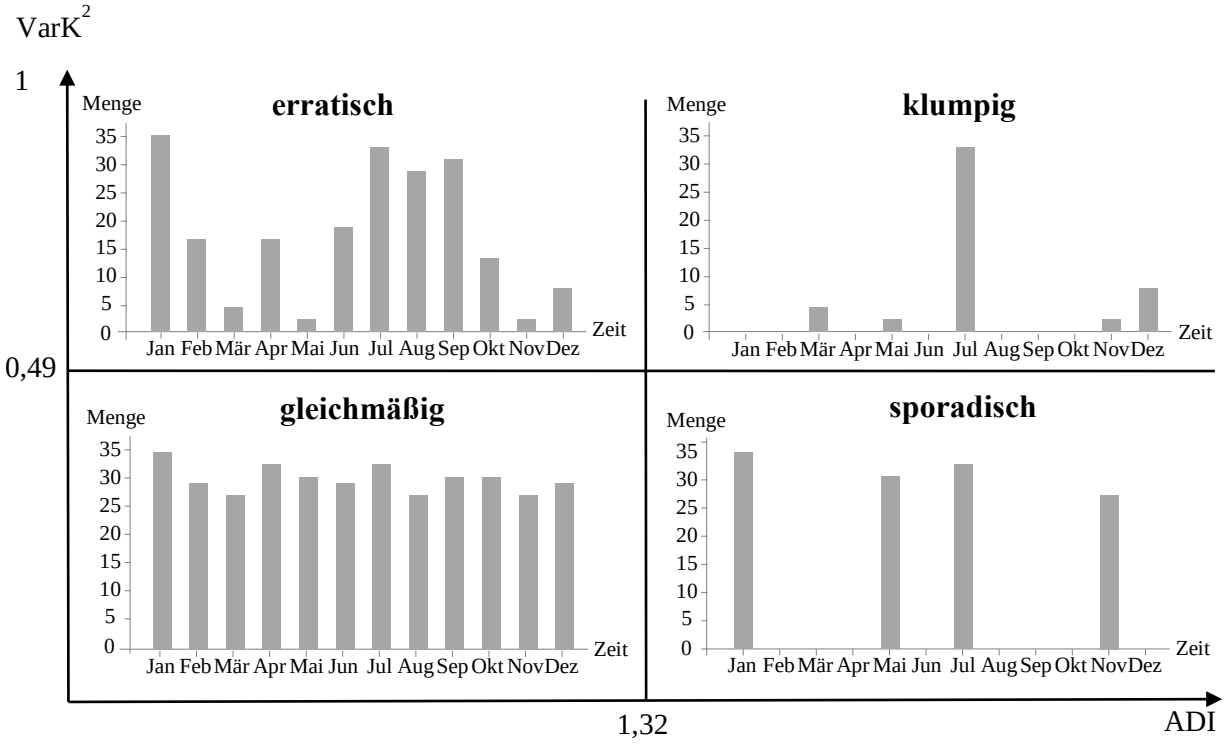

Abbildung 2-5: Klassifizierung von Nachfrageverläufen i.A.a. (Syntetos et al. 2005, S. 500)

Bevor die Nachfrage mittels konkreter Prognosemethoden prognostiziert werden kann, ist es wichtig die unterschiedlichen Zeitreihen tiefergehend, hinsichtlich der grundlegenden Merkmale und des Prozesses, welcher die Daten erzeugt, zu analysieren (Tu und Zhou 2004, S. 386). Denn die Daten einer Zeitreihe sind nicht nur zeitlich geordnete numerische Zahlen, sondern werden durch einen Prozess erzeugt, der als Datengenerierungsprozess (DGP) bezeichnet wird. Der Datenerzeugungsprozess von Zeitreihen ist stochastisch und die Realisierung von Zeitreihendaten ist durch eine gemeinsame Wahrscheinlichkeitsdichtefunktion gekennzeichnet. (Das 2019, S. 15 f.)

Im Allgemeinen wird bei der Nachfrageprognose versucht, ein stochastisches Modell an die vorliegende Zeitreihe anzupassen. Die Grundvoraussetzung für viele Prognoseverfahren ist, dass die zu prognostizierende Zeitreihe der Nachfrage einen stationären Prozess folgt. Ein streng stationärer Prozess ist ein stochastischer Prozess, dessen endlich dimensionale Verteilung gegenüber Zeitverschiebung invariant ist (Vogel 2015, S. 26). Eine weniger restriktive Forderung, die jedoch für viele Anwendungen ausreichend ist, ist die schwache Stationarität. Dabei gelten drei Bedingungen:

- Konstanter Erwartungswert
- Konstante Varianz
- Autokovarianz hängt nur von der Zeitdifferenz ab

Folglich soll es keine Abhängigkeit zwischen der Betrachtungszeit und den empirischen Werten geben. Im Folgenden werden Prozesse, die schwach stationär sind, und Zeitreihen, die solchen Prozessen folgen, als stationär bezeichnet. Zeitreihen können durch unterschiedliche Transformationen in den stationären Zustand überführt werden. Das Ziel ist es, die Komponenten, welche die Stationarität verletzen, zunächst zu separieren. Dazu kann die Zeitreihendekomposition verwendet werden, sodass die Zeitreihe in die verschiedene Komponenten *Trendkomponente*, *Konjunkturkomponente*, *Saisonkomponente* und *Restkomponente* (Residuen) zerlegt und isoliert analysiert werden (Schlittgen und Streitberg 2001, S. 9; Mertens und Rässler 2012, S. 40). Dabei ist es das Ziel, möglichst viel der Zeitreihen durch die Komponenten zu erklären, sodass die Restkomponente (auch Rauschen genannt) möglichst klein ist. Denn die Restkomponente stellt einen reinen Zufallsprozess dar, mit dem Erwartungswert $E(X_t) = 0$ und einer konstanten Varianz σ^2.

Die Prüfung einer Zeitreihe auf Stationarität kann durch die grafische Betrachtung eines Zeitreihenplots oder durch mathematische Berechnungen erfolgen. Sind Trends oder saisonale Schwankungen in der grafischen Zeitreihendarstellung zu erkennen, ist die vorliegende Zeitreihe nicht stationär. Weiterhin kann die Veränderung der empirischen Werte, wie z. B. des arithmetischen Mittels, betrachtet werden. Darüber hinaus stehen zur Überprüfung der Stationarität auch automatisierte Verfahren, wie z. B. der augmented Dickey-Fuller (ADF)-Test (vgl. (Vogel 2015, S. 125)) zur Verfügung.

2.1.3 Grundlagen der Zeitreihenprognose

Im nächsten Schritt muss in Abhängigkeit der analysierten Zeitreiheneigenschaft und dem Prognoseziel die richtige Prognosemethode ausgewählt werden. Zur Nachfrageprognose wurden in der Vergangenheit eine Vielzahl von Verfahren und Methoden veröffentlicht. Eine einheitliche Klassifizierung der Verfahren zur Zeitreihenprognose geht aus der Literatur nicht hervor. Jedoch lassen sich häufig genannte Kategorien in Abhängigkeit der Vorgehensweise wie folgt ableiten (Parmezan et al. 2019, S. 303; Hyndman und Athanasopoulos 2018):

- Klassische Methoden der Statistik
- Autoregressive Verfahren
- Verfahren des Machine Learnings

Nachfolgend werden aufgrund der hohen Anzahl an Veröffentlichungen sowie Weiter-(Entwicklungen) nur die grundlegenden Vorgehensweisen innerhalb der genannten Gruppen kurz erläutert. Im Bereich der **klassischen Methoden** kommen häufig einfache (gleitende-) *Mittelwert*- oder *Regressionsanalysen* zum Einsatz. Die lineare Regression wird oft verwendet, um einen Trend darzustellen. Mithilfe der multiplen linearen Regression können mehrere numerische Parameter mit in das Modell einbezogen werden (Islam et al. 2012, S. 155 ff.). Darüber hinaus wurde in den 50er Jahren die *exponentielle Glättung* durch BROWN entwickelt. Hierbei können durch einen Gewichtungsfaktor α aktuelle Datenpunkte stärker berücksichtigt werden als weiter zurückliegende. Anschließend folgten einige Erweiterung, beispielsweise durch die Verfahren nach HOLT sowie nach HOLT-WINTERS. Wobei das Modell zunächst um die Trendkomponenten (exponentielle Glättung 2. Ordnung) und später durch die Saisonkomponente (exponentielle Glättung 3. Ordnung) erweitert wurde. (Winters 1960, S. 324 ff.) Anschließend wurden zahlreiche Weiterentwicklungen und Kombination der exponentiellen Glättung veröffentlicht.

Die **autoregressiven Verfahren** sind zeitdiskrete Modelle für stochastische Prozesse. Das ARMA-Modell stellt das grundlegende autoregressive Modell dar. Dieses Modell setzt sich aus dem AR-Prozess (Auto-Regression), welcher die Selbstähnlichkeit der Zeitreihe beschreibt, und dem MA-Prozess (Moving Average) zusammen. Das ARIMA-Modell erweitert das ARMA-Modell, um den Parameter I (Integrated). Hierbei wird die Zeitreihe differenziert, sodass der Trend von trendbehafteten Zeitreihen eliminiert und Stationarität hergestellt werden kann. (Deng et al. 1997, S. 247) Das ARIMA-Verfahren wurde zum ersten Mal von BOX und JENKINS beschrieben (Box und Jenkins 1970). Die Literatur beschreibt neben ARIMA noch weitere spezialisierte Modelle mit autoregressiver Komponente, darunter zum Beispiel ARIMAX, ARMAX, ARX, SARIMA, ARIMAX, NARMAX, ARMAX, ARX und NARX (Vogel 2015, S. 148; Alasali et al. 2018, S. 223 ff.; Dantas et al. 2017, S. 116). Viele dieser Modelle können zusätzlich exogene

Variablen, Trends oder weitere Komponenten zur Modellbildung berücksichtigen (Lennox et al. 1995, S. 3274 ff.).

Modelle des **Machine Learnings** haben sich in den letzten Jahren als Konkurrenten zu klassischen statistischen Modellen im Bereich der Prognose etabliert. Die Forschung begann in den Achtzigerjahren mit der Entwicklung von neuronalen Netze. In Folge wurde das Konzept auf andere Modelle, wie z. B. Support-Vektor-Machine, Entscheidungsbäume und andere, die zusammenfassend als Modelle des Machine Learnings bezeichnet werden, ausgedehnt. (Hastie et al. 2017, S. 417 ff.)

Die *künstlichen neuronalen Netze* als ein Bereich des Machine Learnings, finden immer stärkere Beachtung in unterschiedlichen Anwendungsdomänen (Alasali et al. 2018, S. 223 ff.; Dzakiyullah et al. 2014, S. 2129 ff.). Während anfangs neuronale Netze nicht die damals aktuellen Methoden der Zeitreihenvorhersage übertrafen (Steurer 1996, S. 124), liefern heutzutage die neuronalen Netze fallspezifisch gute Ergebnisse (Rathipriya et al. 2023, S. 1946 ff.). Unter dem Begriff der neuronalen Netze werden zahlreiche verschieden geartete Netze geführt. Dazu zählen klassische neuronale Netze mit keinen oder mehreren versteckten Schichten sowie rekurrente neuronale Netze. Eine prominente Art der rekurrenten neuronalen Netze sind die sogenannten Long short-term memory Netze (LSTM). Tiefergehende Erläuterung zur Nachfrageprognose mittels Methoden des Machine Learnings können u. a. in CRONE 2010 nachgeschlagen werden (Crone 2010, S. 67 ff.)

Darüber hinaus werden häufig *Ensemble Methoden* im Bereich von Machine Learning zur Nachfrageprognose eingesetzt, wobei unterschiedliche Vorhersagealgorithmen kombiniert, hinsichtlich ihres Einsatzzwecks angepasst und parametrisiert werden (Li et al. 2016, 22 ff.). Beispielhafte Implementierungen von Ensemble-Methoden sind XGBOOST oder Random Forest Regression (Liao et al. 2019, S. 675 ff.).

Zusammenfassend lässt sich sagen, dass die Entwicklungen im Bereich der Zeitreihenprognosen vielfältig sind und in der jüngsten Vergangenheit die Forschungsaktivitäten im Bereich des Machine Learnings und der künstlichen neuronalen Netze stark zu genommen hat. Ein umfassender Überblick über unterschiedliche Prognosemethoden wird in PETROPOULOS et al. gegeben Petropoulos et al. 2022, S. 705 ff.. Zur Bewertung der jeweiligen Prognosemethoden und -ergebnisse werden Prognosegütemaße (Prognosefehler) verwendet, welche nachfolgend detaillierter dargestellt werden.

Prognosegütemaße

Die Güte einer Prognose wird mittels eines Prognosefehlers ausgedrückt. Dabei gibt der Prognosefehler ex post die Abweichung zwischen dem prognostizierten Wert und dem

tatsächlichen Wert an. Im Allgemeinen kann zwischen drei Prognosefehler (Abweichungen) unterschieden werden, die *systematische Über*- oder *Unterschätzung* sowie der *schwankende Fehler* (Rauschen) (Thonemann 2015, S. 68 ff.; Trigg 1964, S. 271). Beim Prognosefehler **Rauschen** treten sowohl positive wie auch negative Abweichungen zwischen dem Prognosewert und der tatsächlichen Nachfrage im Zeitverlauf auf. Beinhaltet die Nachfrage bzw. die Zeitreihe Strukturbrüche oder wird eine falsche Prognosemethode verwendet, so entstehen häufig systematische Über- oder Unterschätzungen. Die **systematische Überschätzung stellt** eine im Mittel positive Abweichung der prognostizierten Nachfrage dar. Wohingegen bei der **systematischen Unterschätzung** die Nachfrage im Durchschnitt als zu gering prognostiziert wird (Thonemann 2015, S. 71). Die systematischen Fehlerarten werden in der nachfolgenden Abbildung visualisiert.

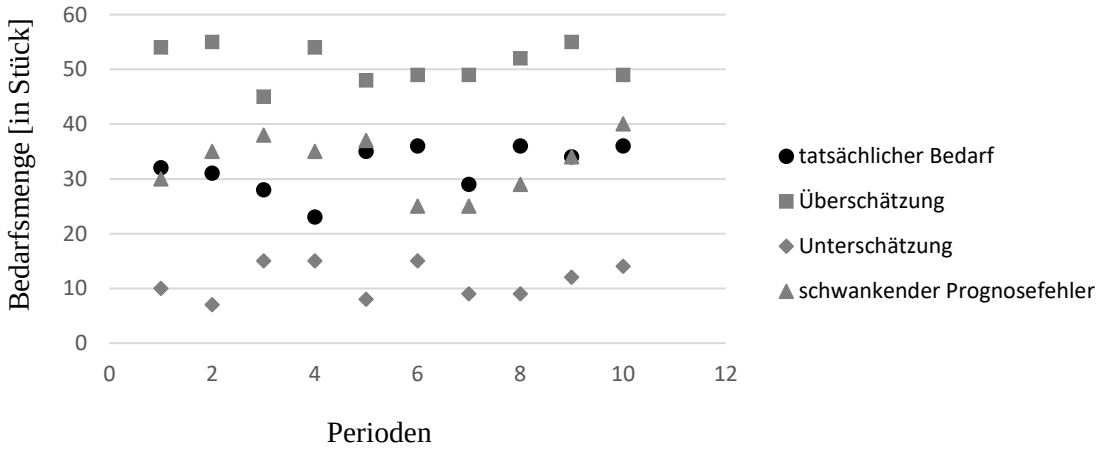

Abbildung 2-6: Systematische Prognosefehler i. A. a. (Thonemann 2015, S. 71)

Zur Quantifizierung der Prognosefehler wurden im Laufe der Jahre verschiedene Gütebewertungskriterien entwickelt, welche eine Bewertung der Prognosequalität ermöglichen (Jose 2017, S. 200). In diesem Rahmen muss das richtige Gütebewertungskriterium, in Abhängigkeit des Prognoseziels ausgewählt werden (Thonemann 2015, S. 68).

Zu den etablierten Gütebewertungskriterien gehören beispielsweise der mittlere quadratische Fehler (MSE), die mittlere absolute Abweichung (MAD), der mittlere absolute prozentuale Fehler (MAPE) und die Wurzel des mittleren quadratischen Fehlers (RMSE) (Ma et al. 2019, S. 2). Bei den zahlreichen Gütemaßen sind viele dieser Maße eng miteinander verwandt. Allgemein lassen sich Prognosefehler in symmetrisch und unsymmetrische unterscheiden. Darüber hinaus können durch eine sogenannte Kompressionsfunktion Prognosefehler zusammengefasst werden. Mögliche Kompressionsfunktionen sind unter anderem der geometrische Durchschnitt, der arithmetische Durchschnitt und der Median. Eine weiterführende Erläuterung der unterschiedlichen Prognosefehler kann

JOSE sowie HYNDMAN & KOEHLER entnommen werden (Jose 2017, S. 200 ff.; Hyndman und Koehler 2006, S. 680 ff.). Die nachfolgende Tabelle 2-1 gibt einen Überblick über die häufig eingesetzten Prognosefehler i. A. a. (Hyndman und Koehler 2006, S. 680 ff.; Jose 2017, S. 200 ff.):

Tabelle 2-1: Übersicht unterschiedlicher Prognosefehler

Prognose-fehler	**Symmetrisch**	**Symmet-riefunktion**	**Relativ nach Kompression**	**Kompression**	**Relativ vor Kompres-sion**
Generalized absolute percentage error (GAPE)					
MAPE	nein	$\lvert e_t \rvert$	nein	mean	prozentual
MdAPE	nein	$\lvert e_t \rvert$	nein	median	prozentual
sMAPE	explizit	$\lvert e_t \rvert$	nein	mean	prozentual
sMdAPE	explizit	$\lvert e_t \rvert$	nein	median	prozentual
Generalized squared percentage error (GSPE)					
SPE	nein	e^2_t	nein	nein	prozentual
RMSPE	nein	$\sqrt{(e^2_t)}$	nein	mean	prozentual
Generalized relative absolute error					
MdRAE	nein	$\lvert et \rvert$	nein	median	relative
GMRAE	ja	$\lvert et \rvert$	nein	geometrischer mean	relative
Generalized root squared error					
RMSE	nein	$\sqrt{(e^2_t)}$	nein	ja	nein
RRMSE	ja	$\sqrt{(e^2_t)}$	relativ	mean	nein
NRMSE	ja	$\sqrt{(e^2_t)}$	normalized	mean	nein
andere					
MAE	ja	$\lvert et \rvert$	nein	mean	nein
MSE	ja	e^2_t	nein	mean	nein
MAD	ja	$\lvert e_t \rvert$	nein	mean	nein
SSE	ja	e^2_t	nein	sum	nein
AE	ja	$\lvert e_t \rvert$	nein	nein	nein

Rollierende Bedarfsplanung

Mit zunehmenden Prognosehorizont sinkt die Prognosegenauigkeit bzw. steigt der Prognosefehler (vgl. Abbildung 2-7). Der Prognosehorizont ist definiert als die zeitliche Länge der Prognose für die, die zukünftige Nachfragewerte prognostiziert werden (Zhao et al. 2019, S. 4067).

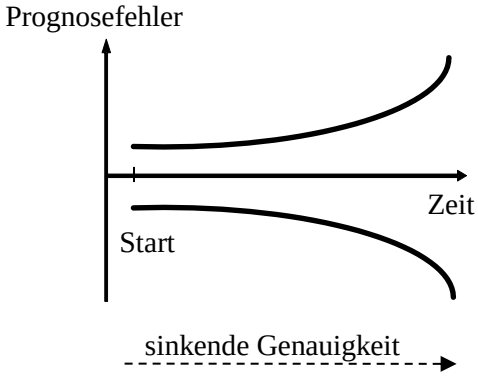

Abbildung 2-7: Wirkzusammenhang - Prognosefehler und -horizont

Zur Minimierung des steigenden Prognosefehlers bei zunehmenden Prognosehorizont wird in der betrieblichen Praxis rollierend geplant (Sahin et al. 2013, S. 5413). Die rollierende Planung stellt eine Vorgehensweise dar, bei welcher nach einem bestimmten Zeitintervall die bereits erfolgte Prognose aktualisiert wird. Diese Vorgehensweise fußt auf der Annahme, dass die Daten- und Informationsgrundlage für Perioden, welche in der nahen Zukunft liegen, höher und präziser ist als für die späteren Perioden. Dieses Vorgehen wird in der Abbildung 2-8 veranschaulicht. Dabei nimmt die Genauigkeit der Nachfrageprognose mit zunehmendem zeitlichem Abstand zum Planungszeitpunkt ab. Dementsprechend wird in einem fest definierten Abstand eine erneute Prognose auf Basis der aktualisierten Daten und Informationen durchgeführt. Durch die regelmäßige Aktualisierung der Prognose können die Unsicherheiten sowie der Prognosefehler geringgehalten werden. (Kistner und Steven 2001, S. 214)

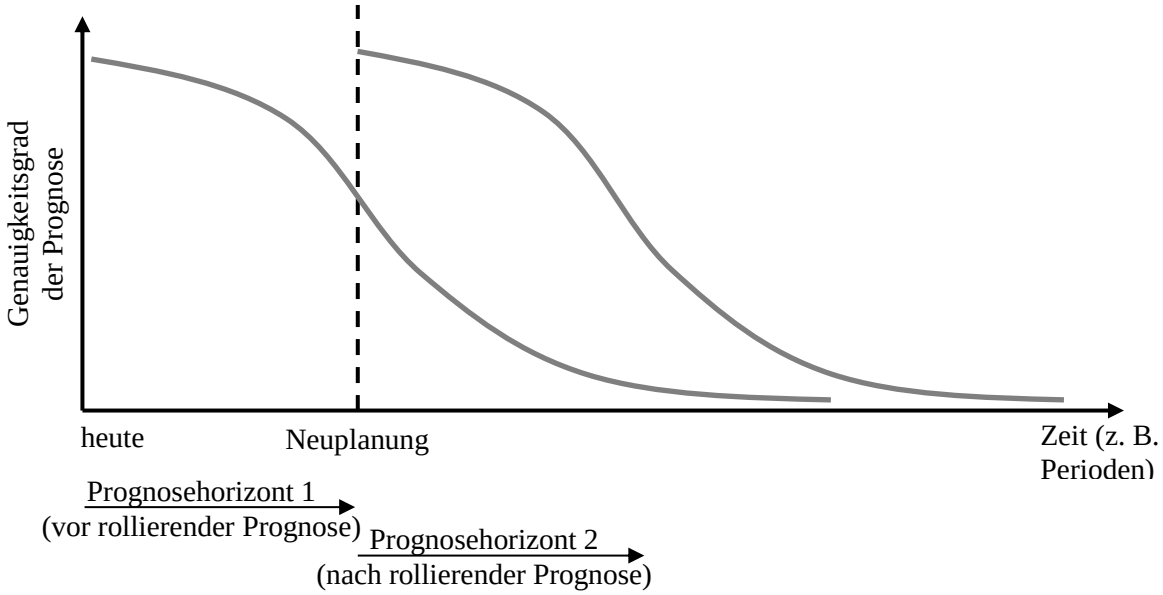

Abbildung 2-8: rollierende Nachfrageprognose

2.2 Planung der Sicherheitsbestände

Zur Gewährleistung der maximal zulässigen Fehlmengen werden in der betrieblichen Praxis Sicherheitsbestände verwendet. Der *Sicherheitsbestand* kann im Rahmen der Bestandsplanung als Absicherung gegen Unsicherheiten im Lagerzugang sowie -abgang definiert werden (Dangelmaier 2017, S. 546). Die Unsicherheiten im Lagerabgang stehen in engem Zusammenhang mit der Bedarfs- und Bestellplanung, sodass sie einerseits die Unsicherheiten auf unvorhersehbare Nachfrageschwankungen oder eine mangelnde Prognosegüte zurückzuführen sind (Wagner 2003, S. 73 f.). Andererseits können die Unsicherheiten durch schwankende Wiederbeschaffungszeiten und mengenmäßige Abweichungen von Bestell- und Liefermenge entstehen (Wassermann und Schwarzer 2012, S. 139). Die Abbildung 2-9 zeigt, innerhalb der idealisierten Sägezahnkurve, die Unsicherheiten im Lagerzugang und -abgang sowie den Zusammenhang zum Sicherheitsbestand auf.

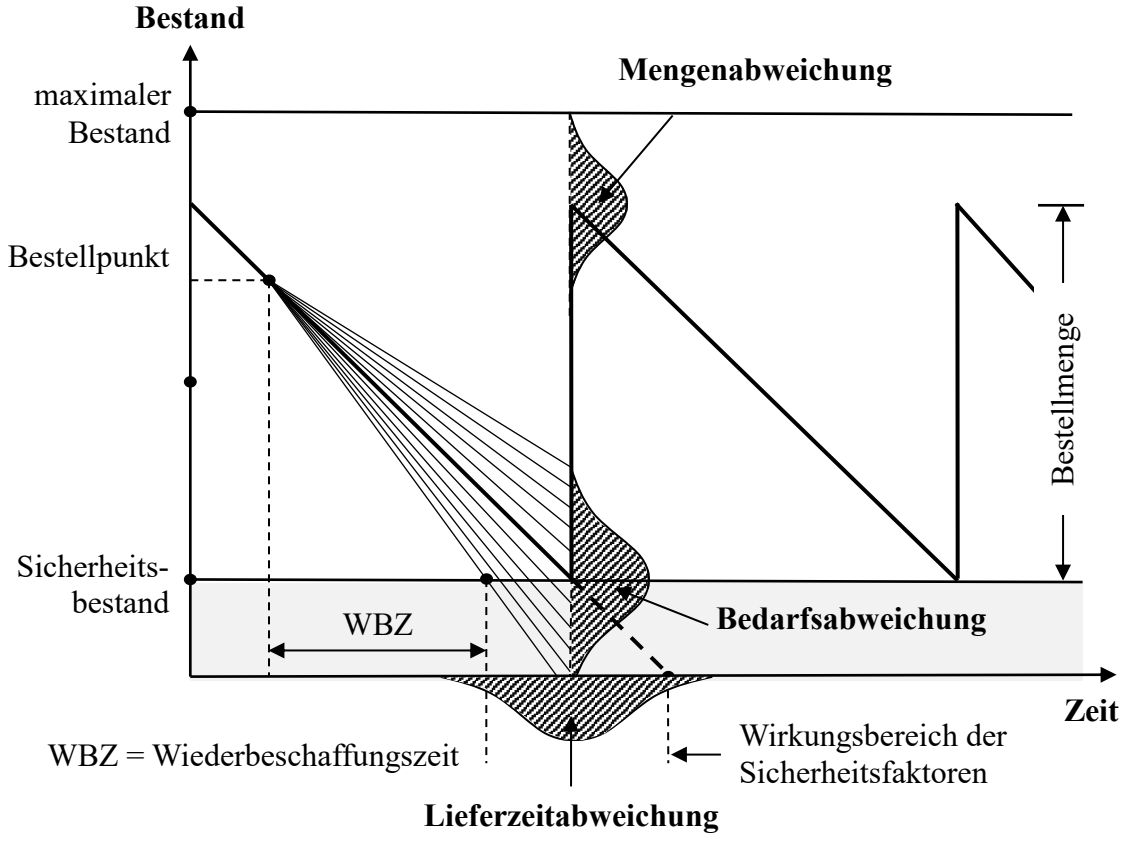

Abbildung 2-9: Unsicherheiten bei der Auslegung des Sicherheitsbestands i.A.a. (Hartmann 2005, S. 424)

Die Unsicherheiten *Lieferzeit-*, *Bedarfs-* und *Mengenabweichung* sind in Form einer Normalverteilung (vgl. Kapitel 3.2.3) dargestellt (vgl. Abbildung 2-9). Im Falle einer bestehenden Unsicherheit in Bezug auf den Lagerzugang besteht die Möglichkeit eines vorzeitigen oder verspäteten Lagereingangs, welcher durch eine Abweichung der Lieferzeit charakterisiert ist. Darüber hinaus können auch mengenmäßige Abweichungen beim Lagerzugang (Mengenabweichung) auftreten, wodurch weniger Materialien oder Produkte geliefert werden als vorab bestellt. Weiterhin wird die Unsicherheit beim Lagerabgang mittels der unterschiedlich stark fallenden Geradenverläufe des Bedarfs (Bedarfsabweichung) visualisiert.

Die aufgezeigten Unsicherheiten unterliegen stochastischen Schwankungen, die nicht immer durch unternehmensinterne Maßnahmen vollständig eliminiert werden können. Dies führt zur Notwendigkeit eines Sicherheitsbestandes (Gleißner und Femerling 2008, S. 150). In der betrieblichen Praxis werden primär Sicherheitsbestände vorgehalten, um sich vor allem gegen Unsicherheiten im Lagerabgang abzusichern. (Wagner 2003, S. 74)

2.2.1 Lieferservicegrad

Demzufolge ist ein Sicherheitsbestand in der betrieblichen Praxis unablässig, um Fehlkosten zu senken. Bei der Dimensionierung des Sicherheitsbestands muss der Trade-Off zwischen den Lagerhaltungs- und Fehlmengenkosten beachtet werden (Korponai et al. 2017, S. 69). Diesen Zusammenhang stellt die nachfolgende Abbildung 2-10 dar.

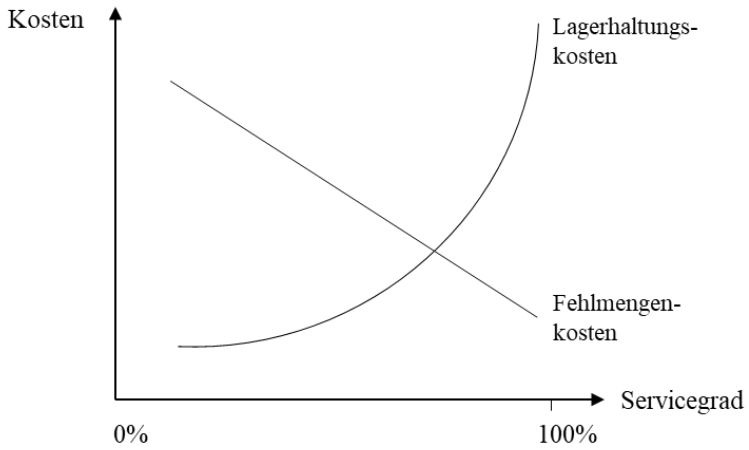

Abbildung 2-10: Kostenorientierte Dimensionierung des Sicherheitsbestands i.A.a. (Bretzke 2015, S. 139)

Eine Erhöhung des Servicegrades hat einen exponentiellen Anstieg des Sicherheitsbestands und den damit einhergehenden Lagerhaltungskosten zufolge (vgl. Abbildung

2-10). Wodurch häufig in der betrieblichen Praxis ein Servicegrad von 100 % nicht wirtschaftlich ist, denn so muss ein Sicherheitsbestand für jegliche Unsicherheit dauerhaft vorgehalten werden. Auf der anderen Seite würde aufgrund der Unsicherheiten ein Betrieb ohne Sicherheitsbestand hohe Fehlkosten verursachen. Folglich gilt es das Kostenminimum zwischen Lagerhaltungskosten und Fehlmengenkosten zu bestimmen (vgl. Abbildung 2-10).

Der (Liefer-)Servicegrad gibt im Allgemeinen den Anteil der mengen- und zeitmäßigen Bedarfserfüllung an. In der Literatur werden unterschiedliche Servicegrade beschrieben. Die am häufigsten verwendeten Servicegrade sind *α-Servicegrad*, *β-Servicegrad* und *γ-Servicegrad* (Friese 2022, S. 30 f.).

Der **α-Servicegrad** stellt eine ereignisorientierte Kennzahl dar und gibt den Anteil der Perioden im Betrachtungszeitraum an, in denen der Bedarf direkt bedient werden kann. Die Berechnung des α-Servicegrades zeigt die Formel (2-7).

$$\alpha = P\ (Bedarf\ einer\ Periode \leq Bestand\ zu\ Beginn\ der\ Periode) \qquad (2\text{-}7)$$

Folglich beträgt beispielsweise der periodenbezogene α-Servicegrad 90 %, wenn von 100 Perioden in 90 Perioden der Bedarf vollständig bedient werden kann. Dabei finden Teillieferungen keine Berücksichtigung.

Der **β-Servicegrad** stellt eine mengenorientierte Kennzahl dar und gibt den Anteil der lieferbaren Bedarfsmenge vom Gesamtbedarf im Betrachtungszeitraum an. Demzufolge wird nicht nur gemessen, ob eine Fehlmenge auftritt, sondern auch die Höhe der Fehlmenge berücksichtigt. (Bäck et al. 2008, S. 174) Die nachfolgende Formel (2-8) zeigt die Berechnung des β-Servicegrades.

$$\beta = 1 - \frac{E\ (Fehlmenge\ pro\ Periode)}{E(Bedarf\ pro\ Periode)} \qquad (2\text{-}8)$$

Die zeit- und mengenmäßige Berücksichtigung erfolgt beim **γ-Servicegrad**, sodass der durchschnittliche Fehlbestand pro Periode mit dem durchschnittlichen Bedarf pro Periode ins Verhältnis gesetzt wird (Bantel 2014, S. 11) (vgl. (2-9)).

$$\gamma = 1 - \frac{E(Fehlbestand\ pro\ Periode)}{E\ (Bedarf\ pro\ Periode)} \qquad (2\text{-}9)$$

Im Vergleich zum β-Servicegrad wird beim γ-Servicegrad nicht nur der Fehlbestand unmittelbar vor der Wiederauffüllung betrachtet, sondern auch die Fehlbestandsentwicklung in den davorliegenden Perioden (Tempelmeier 2018, S. 23). In der betrieblichen Praxis findet der γ-Servicegrad nur selten Anwendung (Tempelmeier 2018, S. 24).

2.2.2 Bestimmung des Sicherheitsbestands

Zur Bestimmung des Sicherheitsbestands kann unterschiedlich vorgangen werden. Im Allgemeinen können die Vorgehensweisen in drei Kategorien unterteilt werden (Biedermann 2008, S. 43):

- Sicherheitszuschläge
- heuristische Methoden
- mathematische Methoden

Werden **Sicherheitszuschläge** verwendet, bestimmt sich der Sicherheitsbestand aus einem prozentualen Zuschlag, welcher beispielsweise mit der durchschnittlichen Bedarfsmenge innerhalb der Wiederbeschaffungszeit multipliziert wird. Der Vorteil von statischen Sicherheitszuschlägen ist die schnelle und einfach Handhabung, jedoch resultieren daraus vielfach hohe Sicherheitsbestände und Lagerhaltungskosten, aufgrund des statischen Vorgehens oder in Folge eines erhöhten Sicherheitsdenkens der Mitarbeiter (Biedermann 2008, S. 44).

Die **heuristischen Methoden** berechnen den Sicherheitsbestand auf Basis von Erfahrungs- sowie Vergangenheitswerten und stellen in der Praxis eine vielfach verwendete Vorgehensweise dar. Ein häufig verwendeter heuristischer Ansatz ist der *reichweitenorientierte Sicherheitsbestand*. (Becker 2016, S. 25 f.) Hierfür wird erfahrungsbasierend eine Sicherheitszeit festgelegt, innerhalb welcher der Sicherheitsbestand die Versorgungssicherheit gewährleisten soll. In diesem Rahmen errechnet sich der Sicherheitsbestand *SB* durch die Multiplikation von Sicherheitszeit *SZ* und dem durchschnittlichen Tagesbedarf *d* (vgl. (2-10)) (Bichler et al. 2010, S. 108):

$$SB = SZ * d \tag{2-10}$$

Im Rahmen der **mathematischen Verfahren** zur Sicherheitsbestandsplanung kommen in den üblichen ERP-Systemen die statisch-stochastischen Verfahren zum Einsatz (Biedermann 2008, S. 45; Babai und Dallery 2009, S. 415). Im Gegensatz zu den zuvor auf-

gezeigten heuristischen Ansätzen wird bei den statisch-stochastischen Verfahren die Planungsunsicherheiten durch eine Verteilungsfunktion berücksichtigt (Becker 2016, S. 28). Dabei muss zunächst die richtige Verteilungsfunktion des Bedarfs bestimmt werden. Die Verteilungsfunktion kann zum einen geschätzt oder durch Anpassungstests auf Basis historischer Bedarfsdaten bestimmt werden (Hartung 2012, S.182 ff.). In der Praxis wird häufig die Normalverteilung aufgrund der einfachen Handhabung und ausreichenden Genauigkeit angenommen (Stadtler 2010, S. 183; Biedermann 2008, S. 69). Im einfachsten Fall werden ausschließlich die Bedarfsschwankungen berücksichtigt. Demzufolge berechnet sich der Sicherheitsbestand *SB* als Produkt aus dem Sicherheitsfaktor *k* und der Standardabweichung der Lagerabgänge σ (vgl. (2-11)) (Schuh und Stich 2013, 101 f.):

$$SB = k * \sigma \tag{2-11}$$

Der Sicherheitsfaktor *k* kann mittels der Normalverteilung in Abhängigkeit des Servicegrades bestimmt und einem Tabellenwerk entnommen werden. Wird der Sicherheitsfaktor mit null definiert, so wird ein Servicegrad von 50 % erreicht. Wird ein höherer Servicegrad angestrebt, so erhöht sich der Sicherheitsfaktor und führt zu einem überproportionalen Anstieg des erforderlichen Sicherheitsbestands.

Auf Grundlage des überproportionalen Anstiegs der Sicherheitsbestände ist ein Servicegrad von 100 % nicht wirtschaftlich (Hartmann 2017, S. 119 ff.). Werden Nachfrageprognosen eingesetzt, so sollte die Auslegung des Sicherheitsbestands auf Basis der Prognosefehlerverteilung und nicht aufgrund der Nachfrageverteilung durchgeführt werden (Meyer und Sander, S. 55).

Statisch-stochastische Verfahren weisen mehrere grundlegende Einschränkungen auf. Zum einen werden Bestandsgrößen für einen längeren Planungshorizont als konstante Zielwerte bestimmt, sodass zeitliche Dynamiken und kurzfristige Schwankungen im System nur unzureichend berücksichtigt werden. Eine rollierende Fortschreibung der relevanten Einflussgrößen kann diese Problematik zwar abmildern, hebt sie jedoch nicht vollständig auf. Zum anderen beruht der Ansatz typischerweise auf der Annahme normalverteilter Bedarfs- bzw. Prognosefehler, was in realen Anwendungskontexten häufig nicht gegeben ist. Insgesamt resultiert hieraus in der Regel eine tendenzielle Überdimensionierung der Sicherheitsbestände. (Becker 2016, S. 28 ff.)

Über die statisch-stochastischen Verfahren hinaus wurden weitere mathematische Verfahren entwickelt, welche eine dynamische Sicherheitsbestandsplanung oder auch unterschiedliche Erweiterungen ermöglichen. Erweiterte Modelle berücksichtigen beispielsweise Terminabweichungen während der Wiederbeschaffungszeit oder Transport- und Sicherheitsbestandskosten (Tempelmeier und Bantel 2015, S. 101 ff.). Diese Modelle

weisen eine immer stärkere Spezialisierung auf. Daher wird an dieser Stelle für einen tieferen Einblick im Bereich der Sicherheitsbestandsplanung auf GONÇALVES et al. und BARROS et al. verwiesen. In der Arbeit von GONÇALVES et al. wird ein systematischer Literaturüberblick über die Verfahren zur Auslegung des Sicherheitsbestands aus dem Bereich Operations Research gegeben (Gonçalves et al. 2020, S. 3 ff.). Wohingegen in BARROS et al. die Verfahren der Sicherheitsbestandsplanung unter stochastischen Bedingungen mit Fokus auf die Beschaffungsrisiken aufgezeigt werden (Barros et al. 2021, S. 2 ff.).

2.3 Bestellplanung

Das Ziel der Bestellplanung ist es, die Bestellzeitpunkte und -mengen kostenoptimal auszulegen (Thommen et al. 2016, S. 159 f.). Dabei bildet die Bedarfsplanung eine zentrale Eingangsgröße, sodass Bestellentscheidungen eng mit dem Wissen über die zukünftigen Bedarfe verbunden sind (Kilger und Wagner 2008, S. 135; Engelmeyer 2016, S. 29).

Im Rahmen der Bestellplanung fokussiert die vorliegende Arbeit die Bestellmengenplanung als einen Bereich der Bestellplanung. Die Bestellmengenplanung zeichnet sich durch eine hohe praktische Relevanz und einer Vielzahl unterschiedlicher Methoden mit einem ähnlichen Einsatzzweck aus. Bereits 1966 deutete VEINOTT daraufhin, dass eine unüberschaubare Anzahl unterschiedlicher Modelle zur Bestellmengenplanung entwickelt wurde (Veinott 1966, S. 745 ff.). Daher ist ein vollständiger Literaturüberblick aufgrund der Vielzahl an Publikationen nicht zielführend (Bushuev et al. 2015, S. 284). Infolgedessen fokussiert die vorliegende Arbeit die Verfahren der Bestellmengenplanung, welche eine praktische Relevanz sowie eine Allgemeingültigkeit im Bezug zum Einsatzzweck aufweisen.

2.3.1 Bestellmengenplanung

Die gängigen ERP-Systeme (z. B. SAP) bieten die Möglichkeit die Bestellmengenplanung mittels *statischer*, *periodischer* und *optimierender Verfahren* sowie mittels des *Meldepunktverfahrens* zu planen (SAP 2022) (vgl. Abbildung 2-11).

Die **statischen Verfahren** basieren auf einfachen Entscheidungsregeln. Beim Lot-for-lot Verfahren, als statisches Verfahren entspricht die Bestellmenge immer dem anfallenden Bedarf für jede Periode. Diese Methode wird häufig für hochpreisige Beschaffungseinheiten sowie für Beschaffungseinheiten mit unregelmäßigem Bedarf verwendet. Werden hingegen statische Bestellmengen definiert, welche in regelmäßigen Zeitabständen bestellt werden, wird von festen Bestellmengen gesprochen. Die feste Bestellmenge kommt häufig bei Mindestbestellmengen oder konstanten Bedarfen zum Einsatz. Darüber hinaus

können die Bestellmengen ebenso durch das zyklische Auffüllen bis zum Höchstbestand geplant werden. Hierbei wird zu definierten Zeitpunkten bis zur Lagerkapazitätsgrenze nachbestellt (Schuh und Schmidt 2014, S. 159). Dieses Verfahren ist besonders für günstige Artikel mit geringen Lagerkosten von Vorteil. Im Bereich der **periodischen Verfahren** werden die Bedarfsmengen periodisch unter Berücksichtigung der Verfügbarkeits- und Liefertermine zusammengefasst und entsprechend wird eine Bestellung veranlasst. Wird das **Meldepunktverfahren** verwendet, so wird ein Meldepunkt (definierte Bestandsmenge) definiert. Wird der Meldepunkt erreicht bzw. unterschritten wird eine Bestellung unter Beachtung der Zielreichweite und des Wunschverfügbarkeitstermins ausgelöst (Burggräf 2021, S. 230). Abschließend stehen in den heutigen ERP-Systemen die **optimierenden Verfahren** der Bestellmengenplanung zur Verfügung, welche in Abhängigkeit der Bestell- und Lagerhaltungskosten die (kosten-)optimale Bestellmenge bestimmen. Dabei werden mehrere Bedarfsperioden zu einer Bestellung zusammengefasst, sodass die Summe aus den Bestell- und Lagerhaltungskosten minimal ist. In diesem Rahmen stehen die vier Heuristiken der dynamischen Bestellmengenplanung, *gleitende wirtschaftliche Losgröße (LUC)*, *Verfahren nach Groff (GR)*, *Stück-Perioden-Ausgleich-Verfahren (PP)* und *dynamische Planungsrechnung (SM)* zur Verfügung. Die vier Heuristiken ermitteln mittels der gleichen Eingangsparameter auf unterschiedlicher Weise die Bestellmengen und -zeitpunkte (vgl. Kapitel 2.3.3). Die beschriebene Unterteilung der Bestellmengenverfahren in den heutigen ERP-Systemen ist in der nachfolgenden Abbildung ersichtlich.

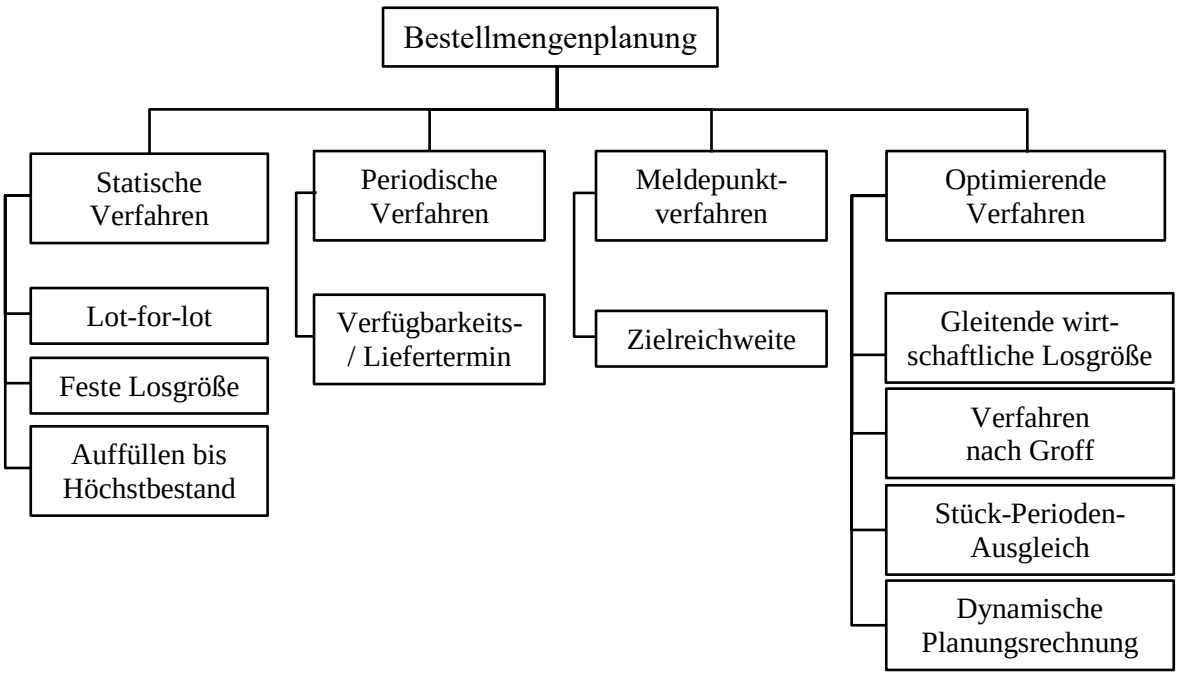

Abbildung 2-11: Klassifizierung der Bestellmengenmethoden in den heutigen ERP-Systemen i.A.a. (Wolf 2021, S. 451 ff.)

Darüber hinaus werden die vier beschriebenen Vorgehensweisen zur Bestellmengenplanung, um die unternehmensspezifischen Eingangsgrößen und Gegebenheiten ergänzt. Dabei werden die relevanten Parameter im ERP-System hinterlegt, wie z. B. Wiederbeschaffungszeit oder Servicegrad (Wolf 2021, S. 451 ff.).

Zusammenfassend bieten die statischen Verfahren eine einfache Möglichkeit die Bestellmengen zu planen, wobei nur wenig Flexibilität sowie keine Kostenorientierung hinsichtlich der Bestellmengenplanung ermöglicht wird. Demgegenüber bieten die periodischen Verfahren, durch flexible Bestellmengen sowie das Meldepunktverfahren, durch flexible Bestellzeitpunkte mehr Flexibilität. Folglich kann auf Bedarfsschwankungen besser reagiert werden als bei den statischen Verfahren. Jedoch haben die relevanten Kosten der Bestellmengenplanung keinen expliziten Einfluss auf die Auslegung von Bestellzeitpunkten und -mengen, wodurch eine kostenorientierte Planung nicht möglich ist. Demnach bieten die optimierenden Verfahren, aufgrund der flexiblen Bestellmengen und -zeitpunkte sowie der Kostenorientierung das größte Potenzial, um die Kosten innerhalb der Bestellmengenplanung zu senken. Im Rahmen der optimierenden Verfahren stellt sich die Frage, wann welches der vier Verfahren ausgewählt werden sollte. Diese übergeordnete Fragestellung stellt den Kern der vorliegenden Arbeit dar. Um die Fragestellung im Detail zu untersuchen werden im nachfolgenden Kapitel zunächst die entscheidungsrelevanten Kostenarten erläutert.

Kostenarten der Bestellmengenplanung

Generell stellt die Minimierung der Kosten eine wichtige Zielgröße der dynamischen Bestellmengenplanung dar, sodass die Bedarfe hinsichtlich Bestellzeitpunkt und -menge möglichst kostenoptimal bestellt werden sollen. Die Gesamtkosten der Bestellmengenplanung können im Wesentlichen auf die drei nachfolgenden Kostenarten zurückgeführt werden (Schulte 2001, S. 28 ff.; Hutzschenreuter 2015, S. 227; Kämmerer 2017, S. 35 ff.):

- Bestellkosten (oder Beschaffungskosten)
- Lagerhaltungskosten (oder Bestandskosten)
- Fehlmengenkosten

Die **Bestellkosten** K_B umfassen alle Kosten, die für die Materialbeschaffung von Bedeutung sind und lassen sich in *Bestellfixkosten* K_{fB} und *variable Bestellkosten* K_{vB} unterteilen (Hutzschenreuter 2015, S. 227).

Die **Bestellfixkosten** K_{fB} umfassen alle Beschaffungskosten, die sich auf die Bestellung an sich zurückführen lassen. Demnach sind die Kosten unabhängig von der Bestellmenge und lassen sich in interne und externe Kosten untergliedern. Externe Kosten sind häufig

verhandelbar, beispielsweise Transportkosten von externen Dienstleistern wie Speditionen (Krobath 2010, S. 38). Die internen Bestellkosten werden durch unternehmensinterne Prozesse und Aufwände erzeugt und umfassen alle Kosten, welche bei der Bestellvorbereitung, -abschluss und -abwicklung entstehen (Hartmann 2005, S. 393) und beinhalten Personal- und Sachmittelkosten für administrative, planende und operative Tätigkeiten. Die Bestellfixkosten sind je nach Fall individuell zu bewerten, sodass diejenigen Kosten identifiziert werden müssen, welche einen Einfluss auf die Bestellplanung haben (Adam 1998, S. 218). Die Bestellfixkosten unterliegen dem Fixkostendegressionseffekt, wobei die Bestellfixkosten pro Einheit, bei steigender Bestellmenge sinken (Vojdani und Knop 2014, S. 3). Diesen Effekt zeigt die nachfolgende Abbildung auf.

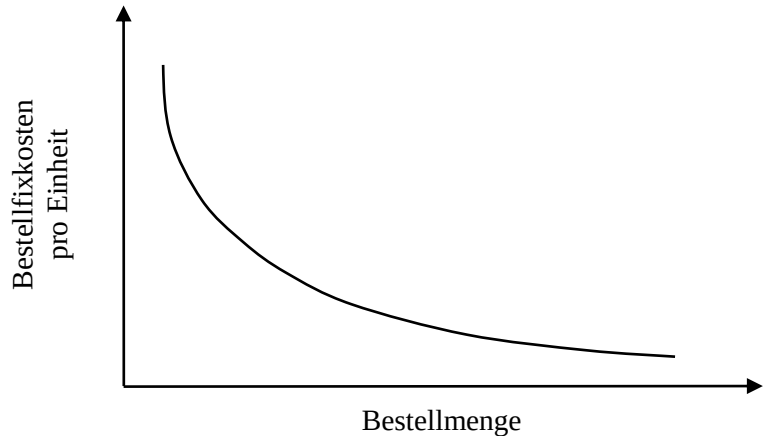

Abbildung 2-12: Degressionseffekt der Bestellfixkosten i.A.a. [KROBATH 2010, S. 40]

Demzufolge werden die Bestellfixkosten durch die zunehmende Bestellmenge getragen. Jedoch lässt sich der aufgezeigte Zusammenhang nicht immer generalisieren, sodass auch sprungfixe Bestellkosten auftreten können. Sprungfixe Kosten verhalten sich innerhalb eines Intervalls fix und steigen oder sinken bei einer Über- oder Unterschreitung des Intervalls. (Mumm 2015, S. 29) Im Rahmen der Bestellplanung entstehen solche Kosten, wenn z. B. die Ladekapazität eines Ladeträgers überschritten wird und ein zweiter Ladungsträger benötigt wird, welcher zusätzliche Kosten verursacht. Treten keine sprungfixen Kosten auf, so können die Bestellfixkosten K_{fB} mittels der Anzahl der Bestellungen n und der Kosten pro Bestellung k_{fB} berechnet werden (vgl. (2-12)).

$$K_{fB} = n * k_{fB} \tag{2-12}$$

Neben den Bestellfixkosten kommen in der Bestellplanung die **variablen Bestellkosten** K_{vB} (oder Anschaffungskosten) zum Tragen (Arnolds et al. 2016, S. 7 f.). Die variablen Bestellkosten steigen mit zunehmender Bestellmenge *Q* und sind in erster Linie vom Einkaufspreis abhängig (Lasch 2019, S. 196). Zwischen dem Einkaufspreis und der Bestellmenge wird häufig ein proportionaler Zusammenhang angenommen, sodass ein linearer Kostenanstieg bei steigender Bestellmenge auftritt. Werden jedoch Mengenrabatte seitens des Lieferanten gewährt, entsteht ein degressiver Verlauf der Einkaufskosten. Ist ein proportionaler Zusammenhang der Einkaufskosten und der Bestellmenge vorherrschend, haben die variablen Bestellkosten keinen Einfluss auf die Bestellmengen und die jeweilige Bestellpolitik (Krobath 2010, S. 42 f.). Die variablen Bestellkosten K_{vB} können aus der Bedarfsmenge im Betrachtungszeitraum B_{BZ} multipliziert mit dem Einstandspreis p berechnet werden (vgl. (2-13)).

$$K_{vB} = B_{BZ} * p \tag{2-13}$$

Die **Lagerhaltungskosten** K_{LH} (oder Bestandskosten) stellen einen weiteren zentralen Einflussfaktor für die Auslegung von Bestellzeitpunkt und -menge dar. Die Lagerhaltungskosten umfassen alle Kosten, die durch die Bevorratung von Produkten und Materialien entstehen und setzen sich aus den *Kapitalbindungs*- und *Lagerkosten* zusammen (Arnolds et al. 2016, S. 8).

Kapitalbindungskosten K_{KB} sind variable Kosten und umfassen die anfallenden Kosten, die sich aus den eingelagerten Produkten und Materialien ergeben (Krobath 2010, S. 47). Zur Ermittlung der Kapitalbindungskosten wird i. d. R. ein Lagerhaltungssatz z verwendet. Dieser Zinssatz beschreibt die prozentuale Verzinsung des gebundenen Kapitals, wenn das Kapital anderweitig genutzt wird (Opportunitätskosten). (Hartmann 2005, S. 396) Die Opportunitätskosten stellen in diesem Rahmen den entgangenen Gewinn bzw. Nutzen eines alternativen Kapitaleinsatzes dar. Demzufolge lassen sich die Kapitalbindungskosten K_{KB} durch die Multiplikation des durchschnittlichen Bestands m_d mit dem Einstandspreis p sowie dem kalkulatorischen Zinssatz z berechnen (vgl. Formel (2-14)).

$$K_{KB} = m_d * p * z \tag{2-14}$$

Demgegenüber sind die **Lagerkosten** K_L fixe Kosten, denn diese sind von den Lagerungsmengen unabhängig und umfassen die Infrastruktur, Instandhaltung und Personalkosten, welche für die Lagerung und für den Materialfluss beansprucht werden (Dyckhoff

2006, S. 316). Darüber hinaus können im Bezug zur Lagerkapazität jedoch auch sprungfixe Kosten auftreten, beispielsweise wenn durch Kapazitätsengpässe die Lagerkapazität angepasst werden muss (Biedermann 2008, S. 41). Bei einer mittel- und kurzfristigen Planung sind i. d. R. die fixen Lagerkosten nicht beeinflussbar und finden keine Berücksichtigung bei der Bestimmung der Bestellzeitpunkte und -mengen. Sind die Lagerkosten im Rahmen der Bestellplanung entscheidungsrelevant, so muss zusätzlich zum Lagerhaltungssatz ein Lagersatz l berücksichtigt werden. (Krobath 2010, S. 47 ff.) Mit dem Lagersatz wird das gebundene Kapital ermittelt, welches durch die fixen Lagerkosten verursacht wird.

Resultierend ergeben sich die Lagerhaltungskosten aus der Summe von Kapitalbindungskosten und Lagerkosten, welche durch die nachfolgende Formel (2-15) mathematisch beschrieben werden können.

$$K_{LH} = K_{KB} + K_L = m_d * p * z + m_d * p * l = m_d * p * (z + l) \qquad (2\text{-}15)$$

Das Ziel der Bestellplanung ist es die unternehmensinternen und -externen Kundenbedarfe hinsichtlich Menge und Zeit zu erfüllen. Die nicht Erreichung des Ziels lässt sich innerhalb der Bestandsplanung auf zwei Ursachen zurückführen. Zum einen die *Überschätzung* und zum anderen die *Unterschätzung* eines Kundenbedarfs. Die Überschätzung des Bedarfs hat eine erhöhte Bevorratung von Materialien und Produkten zur Folge und mündet in zu hohen Kapitalbindungskosten. Bei einer Unterschätzung können die Kundenbedarfe zeit- oder mengenbezogen nicht erfüllt werden. Die negative Abweichung zum tatsächlichen Bedarf wird **Fehlmenge** genannt. (Weber 2012, S. 170)

Die resultierenden Kosten aufgrund von Fehlmengen werden als **Fehlmengenkosten** bezeichnet und lassen sich im Allgemeinen in die drei folgenden Kategorien unterteilen: *zusätzliche Kosten*, *Erlösschmälerung* und *entgangener Deckungsbeitrag* (Schulte 2001, S. 32; Holderied 2005, S. 269). *Zusätzliche Kosten* entstehen, wenn teurere Transportmittel oder Materialien verwendet werden müssen oder ein Mehraufwand durch Überstunden entsteht, um eine fristgerechte Lieferung zu gewährleisten. Können Fehlmengen nicht vermieden werden, so können Schadenersatzzahlungen oder Preisnachlässe zu *Erlösschmälerungen* führen. Im Falle *entgangener Deckungsbeiträge* führt die Lieferunfähigkeit eines Unternehmens zum Verlust von Kunden. (Schulte 2001, S. 31 f.)

Die kurzfristigen Fehlmengenkosten in Form von zusätzlichen Kosten oder Erlösschmälerungen lassen sich i. d. R. gut quantifizieren. Wohingegen die Bestimmung von gravierenderen Fehlmengenkosten, wie z. B. durch Vertrauensschäden oder Reputationsverluste, in der Praxis häufig nicht möglich ist (Lödding 2016, S. 41). Demzufolge ist die Quantifizierung der Fehlmengenkosten mit einer großen Unsicherheit verbunden und

muss kritisch betrachtet werden (Krobath 2010, S. 55). Die vollständige Vermeidung von Fehlmengen ist in den meisten Fällen nicht wirtschaftlich, sodass ein gewisses Ausmaß von Fehlmengen zulässig ist (Kämmerer 2017, S. 49).

2.3.2 Wirkzusammenhänge der Bestellmengenplanung

Einige Methoden der Bestellmengenplanung wurden ursprünglich für die Optimierung der Losgröße im Produktionsbereich konzipiert, welche jedoch später auf die Beschaffung übertragen wurde. Eine Abgrenzung entsprechend dem vom jeweiligen Autor vorgesehenen Einsatzzweck, Produktion oder Beschaffung erfolgt aufgrund der Äquivalenz der beiden Problemstellungen und -strukturen im weiteren Verlauf nicht. (Zelewski et al. 2008, S. 319 ff.)

Die Bestellmengenplanung lässt sich ursprünglich auf einen 1913 publizierten Artikel im „Magazine of Management“ von HARRIS zurückführen (Harris 1913, S. 947 ff.). Aufbauend legte ANDLER im Jahr 1929 den Grundstein der Bestellmengenplanung im deutschsprachigen Raum (Andler 1929). Heute sind diese Modelle sowie Weiterentwicklungen auch als Economic Order Quantity (EOQ) bekannt (Ma et al. 2019, S. 1491).

Das **Economic Order Quantity** Modell stellt die klassische Bestellmengenformel dar (Klaus et al. 2012, S. 301). Eine Vielzahl neuerer Modelle zur Bestellmengenplanung basieren auf dem EOQ-Modell oder sind eine Weiterentwicklung des Modells. Das EOQ-Modell basiert auf dem Prinzip der Minimierung von Lagerhaltungs- und Bestellfixkosten bei bekannten (deterministischen) Bedarfen (Meyer und Sander 2009, S. 66.).

In diesem Rahmen liegen der EOQ-Formel einige Modellannahmen und -vereinfachungen zugrunde (Bichler et al. 2010, S. 96):

- Alle Parameter sind bekannt und konstant
- Die Bestellmengen sind nicht beschränkt
- Die Lieferzeit beträgt Null
- Teillieferungen sind nicht möglich
- Eine Verbunddisposition ist nicht zulässig
- Es können keine Fehlmengen auftreten
- Lagerkapazität sowie Kapital sind unbeschränkt vorhanden

Unter Einhaltung der Prämissen wird die Lösung optimal bestimmt, weswegen es sich um ein optimierendes Verfahren handelt (Krobath 2010, S. 85). Im Allgemeinen können kleine Bestellmengen in kurzen Zeitabständen oder große Bestellmengen in längeren zeitlichen Abständen beschafft werden. In diesem Rahmen macht sich das EOQ-Modell die gegensätzlichen Kurvenverläufe von Bestell- und Lagerhaltungskosten zunutze. Mit dem

Ziel, das optimale Verhältnis zwischen den Bestell- und Lagerhaltungskosten zu ermitteln, sodass das Gesamtkostenminimum erreicht wird (vgl. Abbildung 2-13). (Wiendahl 2014, S. 305)

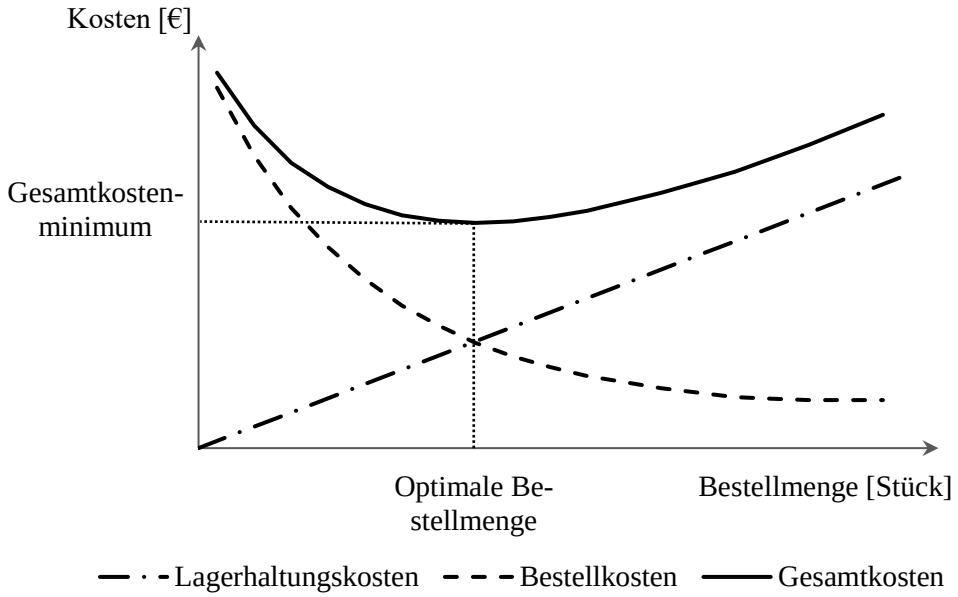

Abbildung 2-13: Gesamtkostenminimum aus Lagerhaltungs- und Bestellkosten (Stölzle et al. 2004, S. 86)

Die Lagerhaltungskosten werden analog zu Gleichung (2-15) berechnet, indem der durchschnittliche Bestand m_d mit dem Einstandspreis p, dem Lagerhaltungssatz l sowie dem kalkulatorischen Zinssatz z multipliziert wird. Aufgrund der Modellannahme konstanter Bedarfe im EOQ-Modell ergibt sich, dass nach Ablauf der Hälfte des Zeitintervalls t bereits die Hälfte der Bestellmenge verbraucht ist, während sich die verbleibende Hälfte weiterhin im Lager befindet. Demnach beträgt der durchschnittliche Bestand m_d die Hälfte der Bestellmenge Q (Çalışkan 2021, S. 1373) (vgl. Abbildung 2-14).

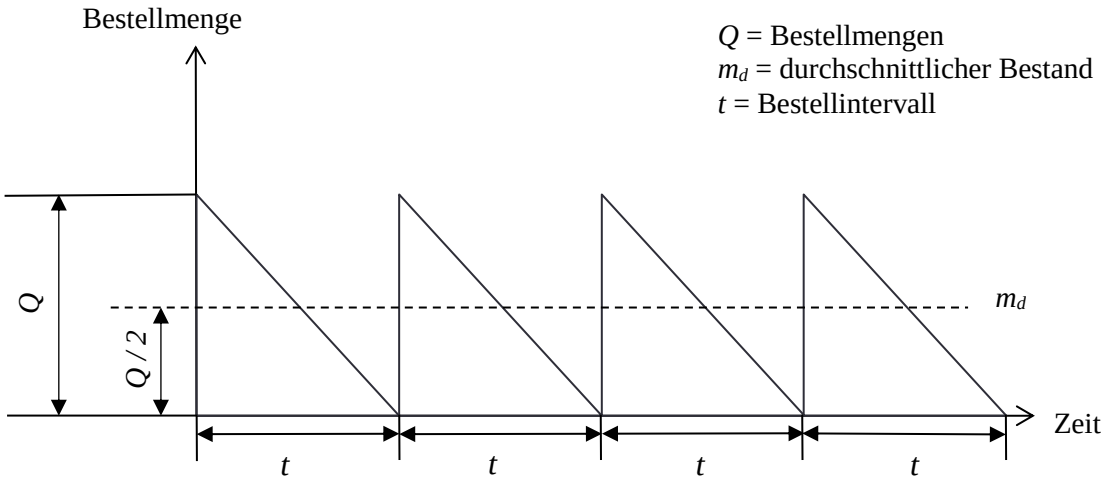

Abbildung 2-14: Zyklische Bestellmengen

Resultierend lassen sich die Lagerhaltungskosten mittels der nachfolgenden Formel ermitteln.

$$K_{LH} = \frac{Q}{2} * p * (z + l) \tag{2-16}$$

Die Bestellkosten K_B setzen sich, wie zuvor erläutert, aus den variablen und fixen Bestellkosten zusammen und lassen sich mittels den Formeln (2-12) und (2-13) berechnen, sodass sich die nachfolgende Formel für die Bestellkosten K_B ergibt.

$$K_B = K_{vB} + K_{fB} = B_{BZ} * p + n * k_{fB} \tag{2-17}$$

Auf Basis der Modellannahmen, dass in konstanten und gleichbleibenden Zeitabschnitten die Bestellmengen bestellt werden, kann die Anzahl der Bestellungen n aus dem Quotienten des Bedarfs im Betrachtungszeitraum B_{BZ} und der Bestellmenge Q bestimmt werden. Demzufolge lassen sich die Bestellkosten K_B wie folgt beschreiben.

$$K_B = K_{vB} + K_{fB} = B_{BZ} * p + \frac{B_{BZ}}{Q} * k_{fB} \tag{2-18}$$

Zusammenfassend ergeben sich die Gesamtkosten K_{ges} auf Basis der Gleichung (2-16) und (2-18).

$$K_{ges} = \frac{Q}{2} * p * (z + l) + B_{BZ} * p + \frac{B_{BZ}}{Q} * k_{fB} \quad (2\text{-}19)$$

Die optimale Bestellmengen der EOQ-Formel Q_{EOQ} ergibt sich, durch die Differenzierung der Formel (2-19) nach Q sowie Auflösung nach Q (Seeck 2010, S. 98). Die nachfolgende Formel zeigt die optimale Bestellmenge.

$$Q_{EOQ} = \sqrt{\frac{2 * B_{BZ} * k_{fB}}{p * (z + l)}} \quad (2\text{-}20)$$

Anschließend kann aus dem Quotienten der optimalen Bestellmenge Q_{EOQ} und dem Bedarf im Betrachtungszeitrum B_{BZ} die optimale Bestellhäufigkeit n_{EOQ} berechnet werden (vgl. Formel (2-21)) (Stölzle et al. 2004, S. 86).

$$n_{EOQ} = \frac{Q_{EOQ}}{B_{BZ}} \quad (2\text{-}21)$$

Darüber hinaus kann die durchschnittliche Bestellfrequenz TBO (Time Between Order) durch den Quotienten der optimalen Bestellmenge Q_{EOQ} und des durchschnittlichen Bedarf pro Periode B_P ermittelt werden (Mohd-Lair et al. 2013, S. 735). Diese Berechnung zeigt die nachfolgende Formel auf.

$$TBO = \frac{Q_{EOQ}}{B_P} \quad (2\text{-}22)$$

Aus der EOQ-Formel (2-20) geht hervor, dass im Kostenminimum die Lagerhaltungskosten den Bestellkosten gleichen. Daraus folgt, umso kostenintensiver ein Artikel ist, desto niedriger fällt die optimale Bestellmenge aus bzw. umso höher die Bestellfixkosten k_{fB} sind, umso größer ist die optimale Bestellmenge. (Kern 1996, S. 1141 ff.)

Obwohl das EOQ-Modell einigen Modellannahmen unterliegt und eine vereinfachte Überlegung zum Ausgleich der Bestell- und Lagerhaltungskosten darstellt, wird das Modell in der Praxis aufgrund der Robust- und Einfachheit häufig zur Grobplanung der Bestellmengen verwendet (Hartmann 2005, S. 405). Dennoch müssen die Ergebnisse, aufgrund der Vielzahl von Modellannahmen und -vereinfachungen kritisch betrachtet werden (Seeck 2010, S. 100). Viele der Modellparameter sind in der Praxis nicht konstant und verändern sich im zeitlichen Verlauf oder sind nicht weit genug im Vorfeld bekannt (Kluck 2008, S. 106 f.). Daher existieren vielfältige Modellerweiterungen, welche aufbauend auf der EOQ-Formel realitätsnaher gestaltet sind (Arnold et al. 2008, S. 156; Meyer und Sander 2009, S. 66). Eine zentrale Erweiterung stellt die Berücksichtigung einer dynamischen Nachfrage dar. Dadurch hat sich die dynamische Bestellmengenplanung mit den vier praxisrelevanten Planungsheuristiken entwickelt. Diese Entwicklung wird im nachfolgenden Kapitel näher diskutiert.

2.3.3 Dynamische Bestellmengenplanung

Der Grundgedanke der dynamischen Bestellmengenplanung folgt dem des EOQ-Modells, sodass das Kostenminimum dann erreicht wird, wenn die Summe aus Bestell- und Lagerhaltungskosten minimal ist (Richards und Grinsted 2016, S. 147). Gegenüber dem EOQ-Modell gehen die dynamischen Bestellmengenmodelle nicht von einer über mehrere Perioden hinweg gleichbleibenden Nachfrage aus, sondern von mehreren, aufeinander folgenden Perioden (z. B. Tagen, Wochen, Monaten) mit schwankender Nachfrage aus. Ebenso wird dabei ein endlicher Zeithorizont angenommen. Die Zeitspanne bis zur letzten betrachteten Periode wird als Planungshorizont bezeichnet. Die Berücksichtigung der Zeitdimension und der Bedarfsänderungen in den einzelnen Perioden führt zur Bezeichnung der "dynamischen" Bestellmengenplanung. (Vahrenkamp 2008, S. 165)

Im Allgemeinen liegen den entwickelten Verfahren der dynamischen Bestellmengenplanung folgenden Prämissen zugrunde (Kern 1996, 1027 f.):

- Zeitlich unregelmäßige Nachfrage
- Endlicher Zeithorizont
- Vollständig Bestellmengen zum Periodenbeginn
- Zu Anfang und zu Ende des Planungszeitraums ist der Bestand bekannt und ist Null
- Mengenproportionale Kapitalbindungskosten am Periodenende
- Alle Kostensätze sind periodenunabhängig

Das Grundmodell der dynamischen Bestellmengenplanung stellt das **Wagner-Whitin** Verfahren (WW-Verfahren) dar. Dieses Planungsmodell wurde 1958 von den beiden namensgebenden Erfindern entwickelt und stellt ein Verfahren der dynamischen Optimierung dar (Wagner und Whitin 1958, S. 89 ff.). Die Vorgehensweise zur Bestimmung der optimalen Bestellmengen wird nach SCHÖNSLEBEN wie folgt beschrieben (Schönsleben 2016, S. 515):

1. Die erste Bestellmenge ist auf den Beginn der ersten Periode anzusetzen.
2. In jeder weiteren Periode ist als Alternative eine neue Bestellmenge anzusetzen. Die Anfangskosten bestimmen sich aus dem Minimum der Gesamtkosten für alle bisherigen Varianten in der vorhergehenden Periode plus Bestellkosten für das Ansetzen einer neuen Bestellung für die laufende Periode.
3. Das Kostenminimum ist der minimale Wert der Gesamtkosten in der letzten Periode.
4. Ausgehend von diesem minimalen Wert wird die Zusammenstellung der Bestellungen „rückwärts“ bestimmt, indem der Weg gesucht wird, auf dem dieses Minimum erreicht wurde.

Demzufolge ist das Verfahren nach WAGNER und WHITIN eine exakte Methode, wobei die optimale Bestellmenge erst bestimmt wird, wenn über den gesamten Lösungsraum alle möglichen Lösungen berechnet und anschließend eine Rekursion durchgeführt wurde (Meyer und Sander 2009, S. 67 f.).

Treten keine Unsicherheiten bei der Bestellmengenplanung auf, so liefert das Optimierungsverfahren nach WAGNER und WHITIN die besten Ergebnisse. Jedoch lassen sich diese Bedingungen nur selten in der Praxis aufgrund von z. B. Nachfrageunsicherheiten antreffen. Folglich stellen die ermittelten optimalen Bestellmengen für den Planungshorizont nicht mehr zwangsläufig die optimale Lösung dar (Kazan et al. 2000, S. 1326). Ebenso benötigt das WW-Verfahren im Vergleich zu den Heuristiken ein Vielfaches der Rechenzeit (Beck et al. 2015, S. 32), was vor allem bei einem großen Produktspektrum, langen Planungshorizonten und häufigen Anpassungsplanungen rechenintensiv sein kann.

Daher wurden in der Vergangenheit unterschiedliche Heuristiken zur Planung der Bestellmengen entwickelt (Ganas und Papachristos 1997, S. 129). In den üblichen ERP-Systemen haben sich die vier Heuristiken *gleitende wirtschaftliche Losgröße* (Least-Unit-Cost Verfahren genannt), *dynamische Planungsrechnung* (Silver-Meal Verfahren genannt), *Stückperiodenausgleichsverfahren* (Part-Period Verfahren genannt) und *Verfahren nach Groff* zur Planung der Bestellmengen etabliert. (Hoppe 2012, S. 609 ff.; Beck et al. 2015, S. 32; Baltes et al. 2017, S. 182; Wolf 2021, S. 460) Diese vier etablierten und

praxisrelevanten Heuristiken der dynamischen Bestellmengenplanung werden nachfolgend näher erläutert.

Heuristiken der dynamischen Bestellmengenplanung

Die vier etablierten Heuristiken der dynamischen Bestellmengenplanung nutzen die geringe Sensitivität der Gesamtkostenkurve, der zufolge Abweichungen von der optimalen Bestellmenge lediglich zu einem moderaten Anstieg der Gesamtkosten führen. (Wiendahl 2014, S. 319). Dabei wird die Bestellmenge durchlaufend iterativ nach bestimmten Regeln ermittelt. Es wird nicht zwangsweise das globale Gesamtkostenminimum bestimmt, sondern eine Näherungslösung, welche in Anbetracht der Bedarfsunsicherheiten und der notwendigen sowie fortlaufenden rollierenden Planung hinreichend genau ist. Darüber hinaus sind die Heuristiken der Bestellmengenplanung als praxisnahe Verfahren zur Ermittlung geeigneter Bestellmengen zu verstehen. Im Vergleich zu exakten Verfahren zeichnen sie sich durch eine geringere Komplexität sowie eine höhere Nachvollziehbarkeit aus. (Kluck 2008, S. 117)

Die vier etablierten Heuristiken der dynamischen Bestellmengenplanung basieren auf denselben Annahmen wie das Wagner-Whitin Verfahren. Darüber hinaus wird im Allgemeinen bei den Heuristiken sukzessiv die Bestellmenge um die Bedarfe der Folgeperioden ergänzt, bis das Abbruchkriterium (Bündelung zu einer Bestellung) erreicht ist. Das Abbruchkriterium wird in den unterschiedlichen Verfahren verschieden definiert. Folglich führen die verschiedenen Verfahren zu unterschiedlichen Bestellmengen und -zeitpunkten sowie Gesamtkosten. (Hartmann 2017, S. 112 f.) Im Folgenden werden die vier etablierten Heuristiken vorgestellt.

Least-Unit-Cost Verfahren (LUC)

Das Least-Unit-Cost Verfahren wurde 1968 entwickelt und geht davon aus, dass das Gesamtkostenminimum genau dann erreicht wird, sobald die Summe aus Bestell- und Lagerhaltungskosten pro Stück minimal ist (Kistner und Steven 2001, S. 64). Daher ist das Ziel des Verfahrens die Minimierung der Stückkosten. Dabei werden iterativ die Bedarfe der einzelnen Perioden sowie die entsprechenden Bestell- und Lagerhaltungskosten aufsummiert. Die Stückkosten ergeben sich anschließend aus der Division der anfallenden Kosten durch die aufsummierte Bestellmenge. Dieses Vorgehen wird periodenweise durchgeführt, solange wie die Stückkosten sinken. (Tempelmeier 2006, S. 153 f.; Bichler et al. 2010, S. 102 f.) Daraus ergibt sich die folgende Gleichung (2-23):

$$\frac{K_B + K_{LH} * \sum_{\tau=i+1}^{j}(\tau - i) * d_\tau}{\sum_{\tau=i}^{j} d_\tau} \leq k_{i,j-1}^{Stück} \tag{2-23}$$

mit ki,j: Stückkosten für die Bedarfe der Perioden i bis j [€/Stück]
Ki,j: Gesamtkosten für die Bedarfe der Periode i bis j [€]
xi,j: Bestellmenge von Periode i bis Periode j
K_B: Bestellkosten [€]
K_{LH}: Lagerhaltungskosten [€/Stück]
(τ-i): Anzahl der Perioden, die der jeweilige Bedarf gelagert werden muss
dτ: Bedarf der Periode τ
i: Periodenindex, in der die Bestellung durchgeführt wird
j: Periodenindex
τ: Periodenindex, mit $(\tau = 1, 2, \dots, T)$
$k_{i,j}^{Stück}$: Gesamtkosten pro Stück für die Bedarfe der Perioden i bis j [$\frac{€}{Stück}$]
$k_{i,j}^{Per}$: Gesamtkosten pro Bedarfsperiode zwischen i und j [$\frac{€}{Periode}$]

Silver-Meal Verfahren (SM)

Als weiteres kostenoptimierendes Verfahren der Bestellmengenplanung wurde die Heuristik nach Silver-Meal (SM) im Jahr 1973 entwickelt. Das Ziel des Verfahrens ist die Optimierung der durchschnittlichen Kosten pro Periode. Dabei werden die Bestellmengen solange zusammengefasst, bis sich die Kosten pro Periode nicht mehr reduzieren lassen (Lukesch und Kellner 2019, S. 161 ff.). Demzufolge werden die Kosten pro Periode von zwei aufeinanderfolgenden Perioden verglichen und geprüft, ob die Kosten der aktuellen Perioden geringer sind als der vorherigen Perioden. Dieses Vorgehen wird solange iteriert bis das Abbruchkriterium $k_{i,j-1}^{Per} > k_{i,j}^{Per}$ erreicht ist. Ist das Abbruchkriterium erreicht, werden alle vorherigen Bedarfsmengen zu einer Bestellmenge summiert. (Herrmann 2011, S. 272 f.) Daraus ergibt sich die nachfolgende Formel (2-24):

$$\frac{K_B + K_{LH} \cdot \sum_{\tau=i+1}^{j}(\tau - i) \cdot d_\tau}{j - i + 1} \leq k_{i,j-1}^{Per} \tag{2-24}$$

Groff Verfahren (GR)

Bei dem Verfahren nach Groff (GR), welches erstmals 1979 vorgestellt wurde, wird die Bestellmenge solange erweitert, bis die Lagerhaltungskosten, den Grenzbestellkosten entsprechen. Demnach werden die einzelnen Periodenbedarfe solange zu einer Bestellmenge zusammengefasst, bis der Grenzwert überschritten wird. (Herrmann 2018, S. 137) Der

Grenzwert wird durch die Division der zweifachen Bestellkosten mit den Lagerhaltungskosten ermittelt. Die nachfolgende Formel (2-25) zeigt die Berechnung nach dem Groff-Verfahren auf, wobei der Grenzwert durch die rechte Seite der Gleichung beschrieben ist.

$$d_j \cdot (j - i) \cdot (j - i + 1) \leq \frac{2 \cdot K_B}{K_{LH}} \tag{2-25}$$

Part-Period Verfahren (PP)

Abschließend steht in den üblichen ERP-Anwendungen das Part-Period Verfahren zur Verfügung. Nach dem Stückkostenausgleichsverfahren wird das Kostenminimum dann erreicht, wenn die Bestellkosten den Lagerhaltungskosten entsprechend. Daher besteht der Grundgedanke des Verfahrens darin, dass die Periodenbedarfe so lange aufsummiert und zu einer Bestellmenge gebündelt werden, bis die damit verbundenen Lagerhaltungskosten die Bestellkosten übersteigen. (Ortmann und Siebeking 2000, S. 15 ff.) Das Vorgehen wird durch die nachfolgenden Formel (2-26) veranschaulicht:

$$\sum_{\tau=i+1}^{j} (\tau - i) \cdot d_\tau \leq \frac{K_B}{K_{LH}} \tag{2-26}$$

2.4 Stand der Forschung: kostenorientierte Auswahl von dynamischen Bestellmengenverfahren

In diesem Kapitel wird der Forschungsstand zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren dargelegt. Hierzu werden zunächst die zentralen und praxisorientierten Anforderungen an den Stand der Forschung definiert, welche sich aus den drei Bereichen der Bestandsplanung sowie aus den vier Heuristiken der dynamischen Bestellmengenplanung ableiten lassen. In diesem Rahmen wird explizit der Fokus auf die allgemeingültigen Anforderungen und Einflussfaktoren gelegt, um die Übertragbarkeit der Ergebnisse zu gewährleisten. Anschließend werden die relevanten Arbeiten in diesem Bereich beschrieben und anhand der Anforderungen evaluiert.

Im Kapitel 2.3.3 wurden bereits die vier **dynamischen Bestellmengenverfahren** identifiziert, welche in den gängigen ERP-Systemen zur Verfügung stehen. Folglich sollen die existierenden Arbeiten alle vier Heuristiken der dynamischen Bestellmengenplanung berücksichtigen.

Diese Verfahren versuchen auf unterschiedliche Weise das Gesamtkostenminimum mittels der Minimierung von Bestell- und Lagerhaltungskosten zu bestimmen (vgl. Kapitel

2.3.3). Folglich besitzen die Bestell- und Lagerhaltungskosten einen signifikanten Einfluss auf die Bestellmengenplanung, sodass diese Kostenarten bei der Auswahl des richtigen Verfahrens berücksichtigt werden müssen. Auf Basis des Zusammenhangs der beiden Kostenverläufe ist das Verhältnis der beiden Kostenarten in Abhängigkeit der durchschnittlichen Nachfrage entscheidungsrelevant. Eine isolierte Betrachtung der beiden Kostenarten ist nicht notwendig. Um die Übertragbarkeit der Ergebnisse zu gewährleisten, müssen unterschiedliche Verhältnisse der beiden Kostenarten betrachtet werden, welche mittels unterschiedlichen **Time Between Order (TBO)** ausgedrückt werden können (vgl. Formel (2-22)).

Darüber hinaus unterliegen die Verfahren der dynamischen Bestellmengenplanung einigen Annahmen, welche häufig im unternehmerischen Umfeld nicht vorherrschen. Aus diesem Grund müssen zur kostenorientierten Auswahl der Verfahren weitere Einflussfaktoren berücksichtigt werden. Diese Einflussfaktoren lassen sich aus den drei Bereichen der Bestandsplanung ableiten und werden nachfolgend erläutert.

Im Bereich der Bedarfsplanung werden üblicherweise die Bedarfe auf Basis von Zeitreihenprognose geplant. Bei der Zeitreihenprognose ist einerseits die historische Nachfrage, welche die vergangenen Bedarfe beschreibt zentral. Andererseits stellt der **Prognosefehler**, welcher infolge einer Prognose entsteht, einen weiteren Einflussfaktor dar. Demzufolge müssen aus der Bedarfsplanung die zwei Einflussfaktoren *Nachfrage* und *Prognosefehler*, zur Auswahl des richtigen Bestellmengenverfahrens berücksichtigt werden. (vgl. Kapitel 2.1)

Auf Grundlage von Nachfrageunsicherheiten sind in der betrieblichen Praxis Sicherheitsbestände unablässig, um hohe Fehlkosten zu vermeiden. Bei der Dimensionierung des Sicherheitsbestands muss der Trade-Off zwischen den Lagerhaltungs- und Fehlmengenkosten beachtet werden. Zur Dimensionierung des Sicherheitsbestands wird i. d. R. der (Liefer-)Servicegrad verwendet. Der **(Liefer-)Servicegrad** gibt im Allgemeinen den Anteil der mengen- und zeitmäßigen Bedarfserfüllung an. (vgl. Kapitel 2.2) Daher sollten die existierenden Arbeiten den Einfluss des Sicherheitsbestands bzw. Servicegrades auf die Auswahl der Bestellmengenverfahren analysieren.

Weiterhin berücksichtigen die vier Heuristiken keine **Wiederbeschaffungszeit (WBZ)** bzw. es wird die WBZ als null angenommen, d. h. dass ein Wareneingang direkt bei einer Bestellung erfolgt. Diese Annahme entspricht nicht der Unternehmenspraxis, sodass zwischen einer Bestellung und dem Wareneingang eine Wiederbeschaffungszeit berücksichtigt werden muss.

Über die aufgezeigten Einflussfaktoren hinaus existieren noch eine Vielzahl weiterer Faktoren, welche jedoch unternehmensindividuell sind und welche aus den Unternehmensspezifika hervorgehen. Die vorliegende Arbeit fokussiert jedoch die allgemeingültigen

und übertragbaren Faktoren. Resultierend müssen die fünf praxisorientierten und allgemeingültigen Einflussfaktoren für die kostenorientierte Auswahl von dynamischen Bestellmengenverfahren berücksichtigt werden. Demzufolge ergeben sich die sechs nachfolgenden Anforderungen an den Stand der Forschung zur Auswahl von dynamischen Bestellmengenverfahren:

- Praxisrelevante dynamische Bestellmengenverfahren
- Kostenstruktur (TBO)
- Nachfrage
- Prognosefehler
- Sicherheitsbestand / Servicegrad
- Wiederbeschaffungszeit (WBZ)

Dabei gilt es in den existierenden Arbeiten zu evaluieren, ob die sechs beschriebenen Einflussfaktoren mit unterschiedlichen Faktorstufen berücksichtigt wurden, damit die Übertragbarkeit der Ergebnisse auf verschiedene Anwendungsfälle gewährleistet wird. Eine Beschreibung der gesamten Literatur zur dynamischen Bestellmengenplanung ist aufgrund der Vielzahl von Arbeiten nicht zielführend. Daher sei an dieser Stelle für einen breiteren Überblick über die Entwicklungen innerhalb der dynamischen Bestellmengenplanung auf die Arbeiten von ANDRIOLO et al. (Andriolo et al. 2014, S. 16 ff.), BRAHIMI et al. (Brahimi et al. 2017, S. 838 ff.), Ma et al. (Ma et al. 2019, S. 1490 f.) sowie VIDAL (Vidal 2023, S. 29 ff.) verwiesen. Nachfolgend soll der Fokus auf die Arbeiten gelegt werden, welche sich auf den Vergleich und die Bewertung von dynamischen Bestellmengenverfahren konzentrieren.

Mit der Entwicklung der unterschiedlichen Verfahren zur dynamischen Bestellmengenplanung wurden auch die ersten Arbeiten zur Bewertung und zum Vergleich dieser Verfahren veröffentlicht. DE BODT et al. zeigte bereits im Jahr 1984 auf, dass die Auswahl einer geeigneten Heuristik zur Bestellmengenplanung nicht einfach ist und von den jeweiligen Kosten und der Nachfragevariablität abhängt (Bodt et al. 1984, S. 169).

BLACKBURN und MILLEN zeigen auf, dass bei unterschiedlichen Kostenstrukturen und unterschiedlichen Nachfrageverläufen sowie kurzen Planungshorizonten die Silver-Meal (SM) Heuristik bessere Ergebnisse liefert als das Optimierungsverfahren nach WAGNER-WHITIN (WW) und die Part-Period Heuristik (PP) (Blackburn und Millen 1980, S. 691 ff.). Einschränkend ist, dass es sich nur um eine kleine Simulationsstudie handelt und kein Servicegrad, Wiederbeschaffungszeit und Prognosefehler berücksichtigt werden. Im Jahr 1985 erweitern BLACKBURN und MILLEN ihre Arbeit um die Heuristiken nach GROFF, FREELAND-COLLEY, und BOE-YILMAZ, bei sonst gleichen Faktoren. Die Studie zeigt auf, dass das GR-Verfahren nur marginale Vorteile gegenüber dem SM-Verfahren aufweist

und die Auswahl des richtigen Bestellmengenverfahrens dem Anwender überlassen wird. (Blackburn und Millen 1985, S. 433 ff.)

CALLARMAN und HAMRIN vergleichen die Verfahren WW, PP und EOQ bei Bedarfsunsicherheiten, unter Berücksichtigung von unterschiedlichen Nachfrageverläufen, TBO und Servicegraden. Vernachlässigt wurden dabei Wiederbeschaffungszeiten und Perioden ohne Bedarf in der Nachfrage. Unter diesen Bedingungen weist bei den meisten Faktorkombinationen das PP-Verfahren die besten Ergebnisse auf. Nur bei niedrigen TBO und hohen Prognosefehlern schneidet das EOQ-Verfahren besser ab. (Callarman und Hamrin 1984, S. 39 ff.)

Ein weiterer Vergleich von dynamischen Bestellmengenverfahren (einschließlich LUC, PP, SM und GR) mit numerischer Auswertung wurde von ZOLLER und ROBRADE veröffentlicht. In dieser Simulationsstudie werden nur die Faktoren Nachfrage und TBO betrachtet. Unter diesen Umständen weisen die Verfahren PP und GR bessere Ergebnisse als SM und LUC auf. (Zoller und Robrade 1987, S. 125 ff.)

WEMMERLÖV und WHEYBARK zeigen auf, dass der Prognosefehler einen signifikanten Einfluss auf die Auswahl des richtigen Bestellmengenverfahrens hat. Dabei wurden u. a. die vier praxisrelevanten Verfahren berücksichtigt. Wird der Prognosefehler nicht berücksichtigt, liefert das Optimierungsverfahren nach WAGNER-WHITIN die besten Ergebnisse. Wohingegen unter stochastischen Bedingungen das PP-Verfahren im Durchschnitt am besten abschneidet. (Wemmerlöv und WHYBARK 1984, S. 467 ff.) Im Jahr 1989 erweitert WEMMERLÖV diese Arbeit und ergänzte einen Servicegrad von 100 %, sodass infolge des Prognosefehlers ein Sicherheitsbestand zur Vermeidung von Fehlmengen berücksichtigt wird. Die Ergebnisse dieser Studie unterstützen die Ergebnisse aus den vorherigen Arbeiten, sodass das PP-Verfahren bei Prognosefehlern und das WW-Verfahren ohne Prognosefehler am besten geeignet ist. (Wemmerlöv 1989, S. 37 ff.)

JEUNET und JONARD bewerten die Robustheit von Bestellmengenverfahren, einschließlich der Verfahren WW, PP und LUC, unter Berücksichtigung stochastischer Nachfrage. In der Arbeit wird aufgezeigt, dass die Verfahren PP und SM robustere Ergebnisse als neuere Verfahren liefern. Eine Auswahl der Verfahren hinsichtlich der zuvor definierten praxisrelevanten Anforderungen wird nicht gegeben. (Jeunet und Jonard 2000, S. 197 ff.)

SIMPSON bewertet u. a. die Verfahren WW, GR, SM und LUC in einer Simulationsstudie ausführlich. Dabei wird der Planungshorizont, der Bedarf und die TBO mit mehreren Faktorstufen berücksichtigt und es wird geschlussfolgert, dass das WW-Verfahren die besten Ergebnisse liefert, obwohl das PP-Verfahren häufig in der Literatur genannt wird. Diese gegensätzliche Schlussfolgerung lässt sich vermutlich auf die Vernachlässigung des Prognosefehlers, Servicegrad und WBZ zurückführen. (Simpson 2001, S. 899 ff.)

FILDES und KINGSMAN entwickelten eine umfassende Simulationsstudie, in welcher das Verhalten unterschiedlicher Bestellmengenverfahren (SM, PP, LUC, EOQ, WW) in Abhängigkeit verschiedener Systemfaktoren und -stufen (Prognosefehler, Nachfrage, Servicelevel und TBO) analysiert wird. Die Ergebnisse bestätigen frühere Untersuchungen, sodass die besten Bestellmengenverfahren im deterministischen Umfeld die schlechtesten sind, wenn die Nachfrage ungewiss ist. Die Größe des Prognosefehlers spielt dabei eine wichtige Rolle. Darüber hinaus sind die Abweichungen vom Servicegrade und der Kostenstruktur abhängig. Die Stückkosten für ein bestimmten Servicegrad steigen exponentiell an, wenn die Bedarfsunsicherheit zunimmt. Konkrete Entscheidungsregeln zur kostenorientierten Auswahl der Bestellmengenverfahren werden nicht gegeben, vielmehr werden die unterschiedlichen Systemfaktoren quantifiziert. Ebenso werden bei der Nachfrage nicht explizit Perioden ohne Nachfrage sowie nicht die Groff-Heuristik berücksichtigt. (Fildes und Kingsman 2011)

In BACIARELLO et al. werden die vier praxisrelevanten Bestellmengenverfahren hinsichtlich unterschiedlicher Prognosefehler, Nachfragen und Kostenstrukturen analysiert. Als Messkriterium wird die Abweichung zum Wagner-Whitin Verfahren, unter deterministischen Bedingungen ermittelt. Die Ergebnisse werden über alle Simulationsläufe gemittelt. Eine situative Auswahl des kostengünstigsten Verfahrens wird nicht aufgezeigt. (Baciarello et al. 2013, S. 1 ff.)

BECK et al. vergleichen die Bestellmengenverfahren WW, GR und LUC sowie zusätzlich das Verfahren nach LEINZ–BOSSERT–HABE (LBH) unter deterministischen Bedingungen. Demzufolge werden in der Arbeit keine Prognosefehler oder Servicegrade berücksichtigt. Der Vergleich der Verfahren erfolgt anhand der Gesamtkosten. Die Ergebnisse unterstützen die vorherigen Arbeiten unter der Prämisse, dass keine Prognosefehler auftreten. Demzufolge liefert das WW-Verfahren die besten Ergebnisse. Wohingegen das LUC-Verfahren die schlechtesten Ergebnisse aufweist. (Beck et al. 2015, S. 20 ff.)

Zwei Vergleichsarbeiten auf Basis von realen Unternehmensdaten werden in DJUNAIDI et al. und YALÇINER präsentiert. In DJUNAIDI et al. werden die Verfahren WW und SM verglichen. Die zugrunde liegenden Daten stammen von einem indonesischen Unternehmen für Innen- und Außenmöbeln. Dabei werden unterschiedliche Produkte betrachtet. Zur Bestimmung der Nachfrage werden verschiedene Prognosemethoden verwendet und retroperspektiv die Prognosefehler bestimmt. Die Kostenstruktur ist durch den realen Anwendungsfall definiert. Zusammenfassend stellen Djunaidi et al. fest, dass in Abhängigkeit der Systemfaktoren das richtige Bestellmengenverfahren ausgewählt werden muss und es nicht ein Verfahren gibt, welches in allen Fällen die geringsten Kosten aufweist. (Djunaidi et al. 2019, S. 1 ff.) In YALÇINER wird ein industrieller Use-Case vorgestellt, in dem zehn Bestellmengenverfahren (u. a. SM, PP, LUC und WW) zur Planung von drei

unterschiedlichen Kaugummi-Produkten eingesetzt werden. Dabei werden die Bestellmengenverfahren anhand der verursachten Gesamtkosten bewertet. Nicht berücksichtigt werden Prognosefehler und Wiederbeschaffungszeiten. (Yılmaz Yalçıner 2021, S. 33 ff.) Auf Basis der geringen Anzahl unterschiedlicher Produkte und miteinhergehenden Faktorstufen lassen sich die Ergebnisse der zwei Arbeiten nicht verallgemeinern.

KIAN et al. unterscheiden beim Vergleich der Bestellmengenverfahren SM, LUC und WW tiefergehend in konvexe und konkave Kosten. Weiterhin werden systematische Nachfrageverläufe (stationär, steigend, fallend und saisonal) und verschiedene Prognosehorizonte berücksichtigt. Die Ergebnisse zeigen auf, dass unterschiedliche Prognosehorizonte keinen Einfluss auf die Rangordnung der unterschiedlichen Bestellmengenverfahren haben. Eine situative Auswahl des richtigen Verfahrens wird nur nach den unterschiedlichen Kostenstrukturen gegeben. (Kian et al. 2021, S. 2294 ff.)

In STADTLER wird ein neues Groff-Null-Kriterium abgeleitet und mittels einer Simulationsstudie mit den bestehenden Verfahren SM, Groff und WW verglichen. Im Rahmen der Nachfrage wird zwischen konstanten, systematischen und sprunghaften Zeitreihen unterscheiden. Weiterhin werden drei unterschiedliche Anteile von Perioden ohne Bedarf angenommen. Die Leistungsbewertung erfolgt anhand der relativen Abweichung zu den minimalen möglichen Kosten. Eine Berücksichtigung von Wiederbeschaffungszeiten und Prognosefehler findet nicht statt. Auf Grundlage der deterministischen Planungsumgebung weist das WW-Verfahren die besten Ergebnisse auf und wird gefolgt von Groff-Null-Kriterium- und SM-Verfahren. Das klassische Groff-Verfahren schneidet deutlich schlechter ab. (Stadtler 2023, S. 3 ff.)

In der Tabelle 2-2 werden die zuvor beschriebenen Arbeiten zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren zusammengefasst und hinsichtlicher der praxisrelevanten Einflussfaktoren bewertet. Die Bewertung der Arbeiten erfolgt anhand von drei unterschiedlich gefüllten Kreise, Anforderungen nicht erfüllt (Kreis nicht gefüllt), Anforderungen teilweise erfüllt (Kreis teilweise gefüllt) und Anforderungen voll erfüllt (Kreis vollständig gefüllt). Die zweite Spalte der Tabelle 2-2 zeigt, ob in den existierenden Arbeiten alle praxisrelevanten Bestellmengenverfahren berücksichtigt und analysiert wurden. In den weiteren fünf Spalten werden die fünf praxisrelevanten Einflussfaktoren bewertet, ob diese allgemeingültig modelliert und mit ausreichend Faktorstufen berücksichtigt wurden.

Tabelle 2-2: Stand der Forschung zur Auswahl von dynamischen Bestellmengenverfahren

Veröffentlichungen / Anforderungen		Bestellmengenverfahren	Nachfrage	Prognosefehler	Kostenstruktur	WBZ	Servicegrad	Entscheidungsregeln
1984	WEMMERLÖV & WHYBARK	●	◑	◑	◑	◑	○	◑
1984	CALLERMANN & HAMRIN	◑	◑	◑	●	○	●	◑
1985	BLACKBURN & MILLEN	◑	●	○	●	○	○	◑
1988	ZOLLER & ROBRADE	●	◑	○	●	○	○	◑
1989	WEMMERLÖV	●	◑	◑	◑	◑	◑	◑
2000	JEUNET & JONARD	◑	◑	○	○	○	○	◑
2001	SIMPSON	●	●	○	●	○	○	○
2011	FILDES & KINGSMAN	◑	◑	◑	●	●	●	◑
2013	BACIARELLO	●	◑	○	◑	○	○	○
2015	BECK ET AL.	●	◑	○	●	○	○	◑
2019	DJUNAIDI ET AL.	◑	○	◑	◑	◑	○	○
2021	KIAN ET AL.	◑	◑	◑	●	○	○	○
2021	YILMAZ YALÇINER	◑	◑	○	◑	○	○	◑
2023	STADTLER	◑	●	○	◑	○	○	○

○ nicht erfüllt ◑ teilweise erfüllt ● erfüllt

Insgesamt lässt sich sagen, dass häufig die vier praxisrelevanten Bestellmengenverfahren berücksichtigt werden. Nur in wenigen Fällen werden zwei der vier Verfahren berücksichtigt. Ein ähnliches Bild zeigt sich im Bereich der Kostenstruktur (TBO). Häufig wird eine ausreichende Bandbreite von unterschiedlichen Faktorstufen, Verhältnissen zwischen Bestell- und Lagerhaltungskosten berücksichtigt. Nur wenige Arbeiten führen keine ausreichende Faktorvariation der Kostenstruktur durch oder basierten auf realen Unternehmensdaten (vgl. (Djunaidi et al. 2019, S. 4 ff.) und (Yılmaz Yalçıner 2021, S. 37 ff.)). In diesen Fällen ist die Allgemeingültigkeit durch das geringe Spektrum nicht gegeben.

Darüber hinaus werden in allen Arbeiten unterschiedliche Nachfrageverläufe berücksichtigt. Jedoch wird in den meisten Fällen die Nachfrage nur mittels unterschiedlicher Streuung (Standardabweichung oder Variationskoeffizient) ohne eine definierte Anzahl an Perioden ohne Bedarf modelliert (vgl. (Wemmerlöv 1989, S. 38 ff.), (Baciarello et al. 2013, S. 4 f) oder (Beck et al. 2015, S. 26 f)). Obwohl die Nachfrage mittels einer Wahrscheinlichkeit modelliert wird, wird die Nachfrage häufig nicht ausreichend repliziert, um valide und übertragbare Ergebnisse zu erhalten (vgl. Kapitel 3.2.4).

Darüber hinaus berücksichtigen nur wenige Arbeiten eine Wiederbeschaffungszeit. Nur FILDES und KINGSMAN berücksichtigen die Wiederbeschaffungszeit mit mehr als zwei Faktorstufen (vgl. (Fildes und Kingsman 2011, S. 490)). Wird explizit keine WBZ berücksichtigt, so wird die WBZ gleich null angenommen, d. h. alle Produkte und Materialien müssen bei Verwendung im Lager vorrätig sein oder mit einer Beschaffungsgeschwindigkeit von unendlich beschafft werden. Diese Vereinfachung bildet, vor allem in der Beschaffung nicht das praktische Umfeld ab.

Die beiden Systemfaktoren Prognosefehler und Servicegrad sind eng miteinander verbunden. Infolge von Prognosefehlern können Bedarfsunterdeckungen entstehen, um dem entgegenzuwirken müssen die geforderten Servicegrade durch Sicherheitsbestände gewährleistet werden. Nur bei ca. der Hälfte der Arbeiten werden Prognosefehler betrachtet. Dabei werden die Prognosefehler i. d. R. als Rauschen modelliert (vgl. (Wemmerlöv und WHYBARK 1984, S. 474), (Callarman und Hamrin 1984, S. 41) und (Fildes und Kingsman 2011, S. 490 ff.)), d. h. mit positiven und negativen Bedarfsabweichungen mit dem Erwartungswert gleich null. Spezifische Prognosefehler, wie beispielsweise systematische Über- oder Unterschätzungen werden nicht betrachtet. Werden Prognosefehler berücksichtigt, so werden nur in drei Arbeiten Sicherheitsbestände in Abhängigkeit des geforderten Servicelevels berücksichtigt. In WEMMERLÖV wird der Sicherheitsbestand vorgehalten, um ein Servicelevel von 100 % zu erreichen. Wohingegen in CALLERMANN & HAMRIN (Callarman und Hamrin 1984, S. 41) sowie FILDES und KINGSMAN (Fildes und Kingsman 2011, S. 488 f) unterschiedliche Höhe des Sicherheitsbestands und der Servicegrade untersucht wurden.

Ebenso lässt sich feststellen, dass für den Vergleich und die Analyse der Bestellmengenverfahren i. d. R. synthetische Daten verwendet werden, welche mittels Simulationsstudien generiert werden. Darüber hinaus besitzen die Ergebnisse vieler Arbeiten nur einen beschreibenden Charakter, sodass selten konkrete Entscheidungsmodelle abgeleitet werden.

Zusammenfassend kann festgehalten werden, dass die vorgestellten Arbeiten verdeutlichen, dass es deutliche Kostenunterschiede zwischen den verschiedenen Verfahren der dynamischen Bestellmengenplanung gibt. Eine kontextspezifische Auswahl birgt daher ein großes Potenzial zur Kosteneinsparung. Demgegenüber adressieren die relevanten Arbeiten bereits einige Anforderungen zur kostenorientierten Auswahl, jedoch werden in keiner der Arbeiten alle praxisrelevanten Systemfaktoren mit einer ausreichenden Anzahl von Faktorstufen modelliert. Vor allem die Systemfaktoren Prognosefehler, Sicherheitsbestand / Servicegrad und Wiederbeschaffungszeit werden nicht praxisorientiert modelliert oder nur mit wenigen Faktorstufen berücksichtigt. Auch konkrete Entscheidungsmodelle zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren werden unzureichend entwickelt. Auf Basis der abgeleiteten Anforderungen sowie der existierenden Arbeiten ergibt sich die Forschungslücke sowie der Forschungsbedarf zur Erzeugung einer allgemeingültigen Datenbasis, auf Grundlage welcher valide und übertragbare Entscheidungsmodelle zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren entwickelt werden können. Um den Forschungsbedarf zu adressieren und gleichzeitig die Forschungslücke zu schließen werden die zuvor definierten Forschungsfragen FF.2 bis FF.4 (vgl. Kapitel 1.2) in den nachfolgenden Kapiteln 3 bis 5 erforscht.

2.5 Zwischenfazit: Konzeptionelle Grundlagen und Stand der Forschung

Das zweite Kapitel widmet sich der ersten Forschungsfrage: „*Welche Anforderungen an die Bestellmengenplanung charakterisieren das unternehmerische Umfeld?*“. Zur Beantwortung dieser Frage wird zunächst die Bestandsplanung definiert und abgegrenzt. Die Bestandsplanung nimmt in Unternehmen eine Querschnittsfunktion ein, indem sie sowohl Produktions-, Beschaffungs- als auch Instandhaltungsprozesse umfasst. Die vorliegende Arbeit fokussiert dabei insbesondere die Schnittstelle zur Beschaffung. In der wissenschaftlichen Literatur existiert keine einheitliche Definition der Bestandsplanung, üblicherweise werden die Aufgabenbereiche Bedarfsplanung, Sicherheitsbestandsplanung und Bestellplanung hervorgehoben. Die Bedarfsplanung dient der Ermittlung des Materialbedarfs, die Sicherheitsbestandsplanung adressiert Schwankungen und Unsicherheiten im Materialfluss, um Fehlmengen zu vermeiden und die Bestellplanung legt Bestellzeitpunkte und -mengen fest. Aus dem Zusammenspiel dieser drei Bereichen und deren Wirkzusammenhängen lassen sich zentrale Anforderungen (Faktoren) ableiten, die für die kostenorientierte Auswahl dynamischer Bestellmengenverfahren von Bedeutung sind und im weiteren Verlauf der Untersuchung berücksichtigt werden.

In der betrieblichen Praxis stehen häufig vier dynamischen Bestellmengenverfahren zur Verfügung: Silver-Meal Verfahren (SM), Groff Verfahren (GR), Least-Unit-Cost Verfahren (LUC) und Part-Period Verfahren (PP) zur Verfügung. Der aktuelle Forschungsstand zur kostenorientierten Auswahl dieser Verfahren zeigt, dass zwischen den vier Heuristiken deutliche Kostenunterschiede bestehen können. Allerdings berücksichtigen bisherige Studien nicht alle praxisrelevanten Anforderungen (Systemfaktoren) mit hinreichender Differenzierung der Faktorstufen, die für eine kontextspezifische Auswahl der Heuristik notwendig wären. Insbesondere die Aspekte Prognosefehler, Sicherheitsbestand beziehungsweise Servicegrad und Wiederbeschaffungszeit wurden bislang entweder vernachlässigt oder nur unzureichend untersucht. Darüber hinaus fehlen konkrete und allgemeingültige Entscheidungsmodelle, die eine kostenorientierte Auswahl dieser Verfahren ermöglichen.

Zusammenfassend lässt sich auf Basis der abgeleiteten Anforderungen und des Forschungsstandes die Forschungslücke sowie der Forschungsbedarf ableiten. Dieser besteht in der Schaffung einer allgemeingültigen Datenbasis, die als Grundlage für die Entwicklung valider und übertragbarer Entscheidungsmodelle zur kostenorientierten Auswahl dynamischer Bestellmengenverfahren dient.

3 Synthetische Daten zur dynamischen Bestellmengenplanung

Zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren muss zunächst eine allgemeingültige und übertragbare Datenbasis erhoben werden, um die unterschiedlichen Bestellmengenverfahren hinsichtlich ihrer Gesamtkosten zu analysieren und zu bewerten. Im Bereich der Bestellmengenplanung können Daten aus betrieblichen Abläufen (reale Daten) und aus Simulationsexperimenten (synthetische Daten) gewonnen werden. Zur Gewährleistung der Reproduzierbarkeit, Generalisierbarkeit und Abbildungsgüte müssen eine Vielzahl unterschiedlicher Daten erhoben werden, sodass eine große Bandbreite unterschiedlicher Faktoren und Faktorstufen berücksichtigt werden. Eine derart große Datenbasis kann nur schwer aus realen Systemen erhoben werden. Simulationsexperimente bieten die Möglichkeit, Daten in ausreichender Quantität und Qualität zu erzeugen, um eine neue oder erweiterte Datenbasis für solche Verfahren zu schaffen (Rueda und Fink 2020, S. 10066 ff.). Demzufolge ist für die kostenorientierte Analyse von dynamischen Bestellmengenverfahren eine synthetische Datenerzeugung vorteilhaft. An dieser Stelle setzt die zweite Forschungsfrage an: „*FF 2.: Wie lassen sich synthetische Daten zur dynamischen Bestellmengenplanung erzeugen?*“

Zur Beantwortung der zweiten Forschungsfrage wird zunächst die theoretische Verankerung zur synthetischen Datenerzeugung und die konzeptionellen Grundlagen zur Modellierung und Simulation erläutert, welche den Rahmen für die weiteren Arbeiten in diesem Kapitel darstellen. Anschließend erfolgt die Modellierung und Simulation der dynamischen Bestellmengenplanung. Dabei werden die Ziele und Aufgaben der Simulationstudie beschrieben. Aufbauend wird das formale Modell, der Experimentplan und die Experimentdurchführung entwickelt und vorgestellt. Im Anschluss erfolgt die Verifizierung und Validierung der Simulationsstudie. Abschließend werden die Ergebnisse des Kapitels zusammengefasst und die Forschungsfrage beantwortet.

3.1 Simulation und Modellierung von logistischen Systemen

In diesem Abschnitt werden zunächst die zentralen Aspekte der Systemtheorie und der Modelltheorie (vgl. Kapitel 3.1.1) erläutert, welche grundlegend für die Arbeit sind. Aufbauend werden die notwendigen Vorgehensweisen und Methoden zur Modellierung und Simulation von logistischen Systemen (vgl. Kapitel 3.1.2) aufgezeigt.

3.1.1 Systemtheoretischer Ansatz zur synthetischen Datenerzeugung

Die Systemtheorie hat ihren Ursprung in der Biologie und wurde durch BERTALANFFY als allgemeine Systemtheorie begründet (Bertalanffy und Ludwig von 1949, S. 114 ff.). Die allgemeine Systemtheorie stellt die Grundlage für zahlreiche Weiterentwicklungen in unterschiedlichen Disziplinen dar, wie z. B. die Kybernetik (Wiener und Trawny 2022, S. 25 ff.), das Operations Research (Churchman et al. 1957, S. 7 ff.) oder die Soziologie (Luhmann 2018, S. 30 ff.). Auch wenn sich keine einheitliche Definition zur Systemtheorie durchgesetzt hat, so unterscheiden sich die verschiedenen Definitionen nicht wesentlich (Patzak 1982, S. 11).

Die **allgemeine Systemtheorie** definiert, Systeme als strukturierte Gesamteinheiten von Elementen, die durch Beziehung miteinander verknüpft sind. Sie analysiert, wie sich Systeme und Elemente durch Interaktionen, Rückkopplungsmechanismen und Selbstorganisation verhalten und verändern, unabhängig davon, ob sie biologischer, technischer, sozialer oder wirtschaftlicher Natur sind. Somit liefert die Systemtheorie das theoretische Fundament für das Verständnis von Systemen im Allgemeinen.

Demgegenüber konkretisiert die **Systemtechnik** die Systemtheorie durch eine strukturierte Vorgehensweise und methodische Ansätze, um (komplexe) Systeme effizient und zielgerichtet zu gestalten (Schmidt 2018, S. 13). Diese gestalterische und anwendungsorientierte Sichtweise der Systemtheorie und aufbauend der Systemtechnik ist von besonderer Bedeutung für die Gestaltung der dynamischen Bestellmengenplanung, als zugrunde liegendes System für die Erzeugung der notwendigen synthetischen Daten.

Gemäß DIN IEC 60050-351 ist ein **System** als eine Menge miteinander in Beziehung stehender Elemente, die in einem bestimmten Zusammenhang als Ganzes betrachtet werden und von ihrer Umgebung abgegrenzt sind (DIN IEC 60050-351:2013, S. 21). Gutenschwager et al. konkretisieren diesen Systembegriff anhand der folgenden Eigenschaften eines Systems (Gutenschwager et al. 2017, S. 11):

- Ein System ist grundsätzlich begrenzt und durch sogenannte Systemgrenzen gegenüber der Umwelt (Systemumgebung) in seinem Umfang festgelegt. Über definierte Schnittstellen kann ein System an den Systemgrenzen Materie, Energie und Information (Ein- und Ausgangsgrößen) austauschen. Sofern der Austausch von der Umwelt in das System erfolgt, wird von Eingangsgrößen, ansonsten von Ausgangsgrößen gesprochen. Bei den in Produktion und Logistik betrachteten Systemen sind die Systemschnittstellen und damit die Einflüsse der Systemumgebung auf das System in vielen Fällen von hoher Relevanz.
- Ein System besteht aus Systemelementen (Synonym: Systemkomponenten), die bei detaillierter Betrachtung selbst wieder Systeme darstellen (Subsysteme) oder

aber als nicht weiter zerlegbar angesehen werden. So kann eine Maschine Bestandteil eines Produktionssystems sein oder selber als System mit seinen Maschinenbestandteilen betrachtet werden.

- Die Struktur eines Systems ergibt sich aus den Beziehungen zwischen den Elementen eines Systems. Diese Beziehungen können durch unterschiedliche Einflüsse der Systemumgebung oder der Systemfunktionalität bestimmt sein.
- Jedes Systemelement besitzt Eigenschaften, die über konstante und variable Attribute (die sogenannten Zustandsgrößen) abgebildet werden. Die jeweiligen Zustände eines Systemelementes werden durch die Werte der konstanten und variablen Attribute zu einem Zeitpunkt beschrieben. Die Zustände der Systemelemente zu einem bestimmten Zeitpunkt definieren den Systemzustand.
- Die Zustände der Elemente können Änderungen von einer oder mehreren Zustandsgrößen unterliegen (Zustandsänderung, Zustandsübergang aufgrund des in dem System ablaufenden Prozesses). Als Prozess wird in diesem Zusammenhang „die Gesamtheit von aufeinander einwirkenden Vorgängen in einem System, durch die Materie, Energie oder Informationen umgeformt, transportiert oder gespeichert wird“, bezeichnet (DIN IEC 60050-351:2013, S. 33). Prozesse können sequenziell, parallel, nebenläufig, unabhängig voneinander, und synchronisiert zusammenwirken (VDI 4465, S. 44).
- Die einzelnen Elemente enthalten eine eigene Ablaufstruktur, die durch spezifische Regeln hinsichtlich der Zustandsgrößen und der Zustandsübergänge charakterisiert wird.

Die nachfolgende Abbildung veranschaulicht die aufgezeigten Elemente und Eigenschaften eines Systems.

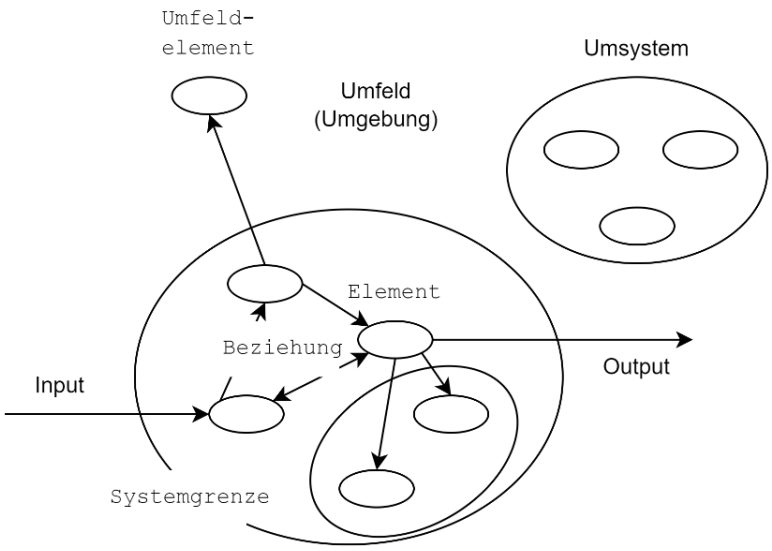

Abbildung 3-1: Zusammenwirken der Systemelemente (Schmidt 2018, S. 15)

Auf Basis einer Vielzahl in Wechselwirkung stehender Faktoren innerhalb der dynamischen Bestellmengenplanung, teilweise mit stochastischen Verhalten können solche Systeme als komplexe Systeme bezeichnet werden. Ein komplexes System ist ein System, dessen Gesamtverhalten trotz zahlreicher Informationen über die Einzelkomponenten und deren Wechselwirkungen nicht vollständig vorhersehbar ist. Die Komplexität entsteht durch die Faktorvielfalt, deren gegenseitige Abhängigkeit, konträren Zielsetzungen sowie die Unsicherheit einiger Faktoren. (Johnson 2004, S. 19 f.)

3.1.2 Konzeptionelle Grundlagen zur Simulation und Modellierung von logistischen Systemen

Ein **Modell** wird definiert als eine vereinfachte Nachbildung eines geplanten oder bestehenden Systems, einschließlich seiner Prozesse in einem alternativen begrifflichen oder gegenständlichen System. Dabei weicht es nur innerhalb eines durch das Untersuchungsziel bestimmten Toleranzrahmens von seinem Original ab. (VDI 3633 2014, S. 3)

Das Untersuchungsziel stellt eine zentrale Determinante eines Modells dar. Abhängig von der jeweiligen Zielsetzung können Modelle desselben Objekts variieren. Die Überführung eines realen Systems in ein Modell wird als Modellierung bezeichnet. Ein Modell muss die wesentlichen Charakteristika des Originals in hinreichender Genauigkeit erfassen, sodass die aus dem Modell gewonnenen Erkenntnisse auf das reale System übertragbar sind. (Schenk et al. 2014, S. 220)

Eine zentrale Herausforderung der Modellierung besteht in der Bestimmung eines angemessenen Detaillierungsgrades, der den Anforderungen des jeweiligen Analyseziels entspricht. Ein Modell sollte nur so detailliert wie nötig gestaltet sein, um eine möglichst geringe Modellkomplexität zu erhalten. (Gutenschwager et al. 2017, S. 20)

Zur Abstraktion nennen GUTENSCHWAGER et al. die Techniken der Reduktion, die das Weglassen von nicht untersuchungsrelevanten Details umfasst, sowie die Idealisierung, die eine vereinfachte Darstellung relevanter Merkmale bezeichnet (Gutenschwager et al. 2017, S. 20 f.). Auf Basis der hohen Komplexität logistischer Systeme ist ein hoher Grad an Reduktion erforderlich. Eine solche Reduktion ist auch im Bereich der dynamischen Bestellmengenplanung notwendig, um mit vertretbarem Aufwand das System zu analysieren (Grundstein 2017, S. 31). Gleichzeitig muss ein Modell die wesentlichen Eigenschaften des Systems hinreichend genau abbilden, sodass die durch die Simulation gewonnenen Erkenntnisse auf die Realität übertragbar sind (Schenk et al. 2014, S. 220).

Im Allgemeinen lassen sich zwei methodische Ansätze zur Modellierung unterscheiden, die Top-down-Methode und die Bottom-up-Methode. Bei der Top-down-Methode wird ein System zunächst in seiner Gesamtheit betrachtet und anschließend schrittweise detailliert, bis die erforderliche Genauigkeit erreicht ist. Im Gegensatz dazu basiert die Bottom-up-Methode auf der Modellierung einzelner, überschaubarer Teilsysteme, die sukzessive zu einem Gesamtmodell zusammengeführt werden. (Schenk et al. 2014, S. 221)

Aufbauend auf der Modellierung wird häufig die Simulation zur Untersuchung von dynamischen Zusammenhängen in logistischen Systemen verwendet. Nach lässt sich die Simulation als das

> *„Nachbilden eines Systems mit seinen dynamischen Prozessen in einem experimentierbaren Modell, um zu Erkenntnissen zu gelangen, die auf die Wirklichkeit übertragbar sind; insbesondere werden die Prozesse über die Zeit entwickelt" definieren (VDI 3633 2014, S. 3).*

Ergänzend dazu kann die Simulation als das Vorbereiten, Durchführen und Auswerten gezielter Experimente mit einem Simulationsmodell verstanden werden. Demzufolge werden Simulationsexperimente definiert als „gezielte empirische Untersuchung des Verhaltens eines Modells durch wiederholte Simulationsläufe mit systematischer Parameter- oder Strukturvariation" (VDI 3633 2014, S. 3).

In Abhängigkeit des Simulationsziels können unterschiedliche Arten von Simulationsmodellen eingesetzt werden. Dabei können Modelle hinsichtlich ihrer zeitlichen (statisch oder dynamisch) und zufallsabhängigen (deterministisch oder stochastisch) Eigenschaften klassifiziert werden. *Statische Modelle* betrachten nur einen Zeitpunkt oder ignorieren den Zeitverlauf, während *dynamische Modelle* zeitliche Veränderungen einbeziehen. *Deterministische Modelle* enthalten keine zufallsabhängigen Komponenten, wohingegen

stochastische Modelle durch Zufallsverteilungen beeinflusst werden. Zudem unterscheiden sich Modelle in der Zeitdarstellung. *Kontinuierliche Modelle* erfassen den Zeitverlauf fortlaufend, während *diskrete Modelle* den Zeitfortschritt ereignisgesteuert in Sprüngen abbilden (Law 2015, S. 3 ff.). Die nachfolgende Abbildung fasst die aufgezeigte Ausprägung zusammen:

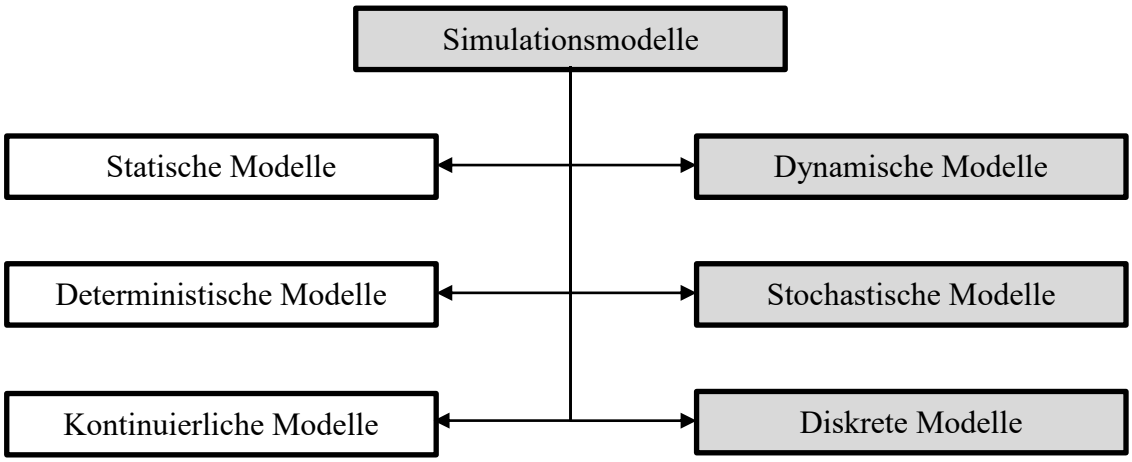

Abbildung 3-2: Einteilungen von Simulationsmodellen i.A.a. (Besenfelder 2018, S. 49)

Die sorgfältige und systematische Durchführung von Simulationsstudien ist von zentraler Bedeutung, weshalb ein strukturiertes Vorgehensmodell unerlässlich ist. Im deutschsprachigen Raum hat sich insbesondere das von RABE et al. etablierte Modell zur Modellformalisierung und Durchführung von Simulationsstudien durchgesetzt (Rabe et al. 2008, S. 4 ff.). Wie in Abbildung 3-3 dargestellt, ist dieses Vorgehensmodell durch verschiedene Phasen mit jeweils zugehörigen Ergebnissen charakterisiert, wodurch eine strukturierte und transparente Vorgehensweise gewährleistet wird.

Die Simulationsstudie gliedert sich in die folgenden Phasen *Aufgabendefinition*, *Systemanalyse*, *Modellformalisierung*, *Implementierung* sowie *Experimente und Analyse* (Rabe et al. 2008, S. 6). Parallel zur Modellbildung erfolgt die *Datenbeschaffung* und *-aufbereitung*, um eine fundierte Basis für die Modellierung und Analyse zu schaffen.

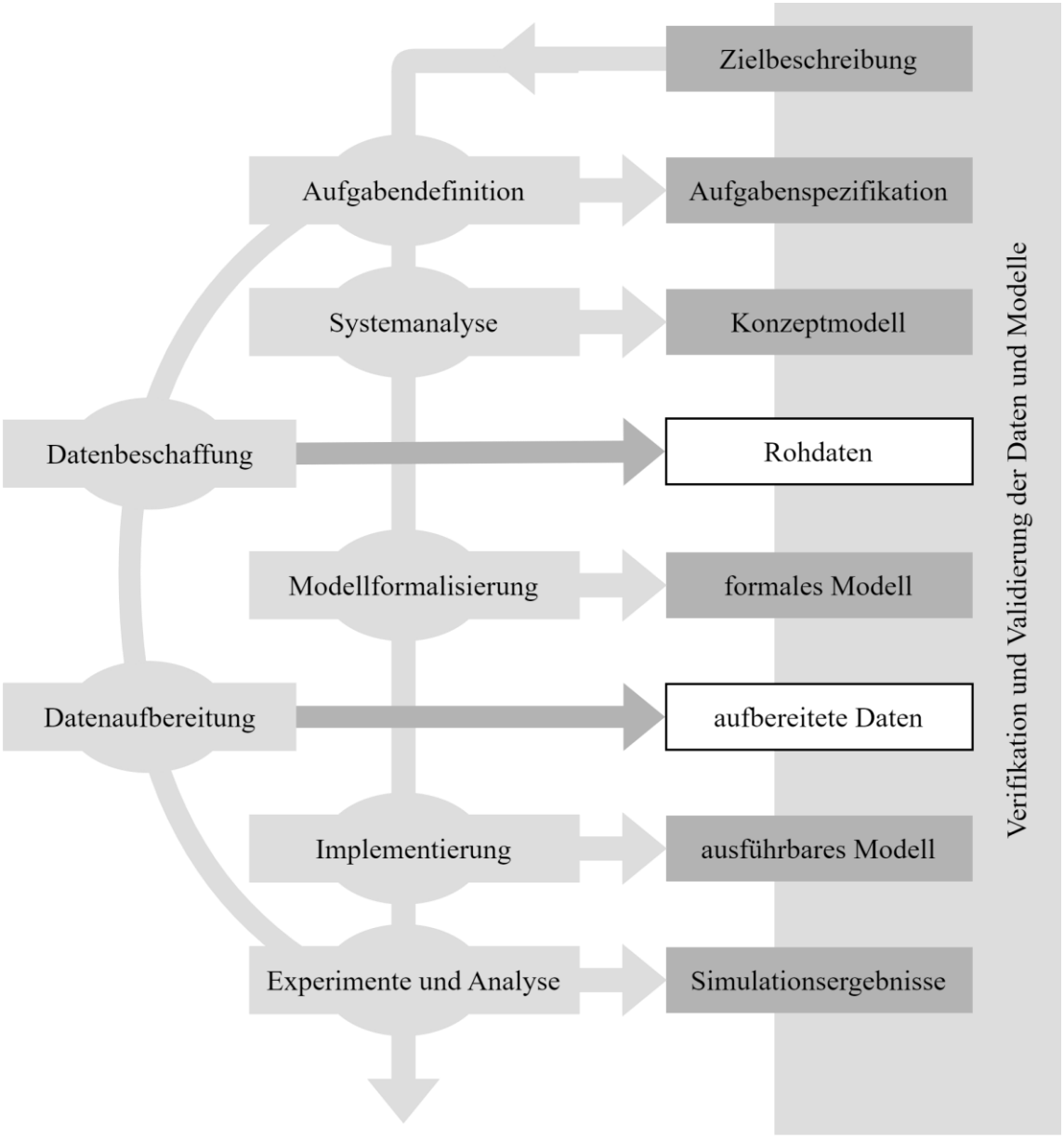

Abbildung 3-3: Vorgehensmodell zur Modellierung und Simulation i.A.a. (Rabe et al. 2008, S. 5)

Das Vorgehensmodell zur Durchführung von Simulationsstudien wird grundsätzlich sequenziell durchlaufen. Rückkopplungsschleifen zu vorangegangenen Phasen sind möglich, falls die Eignung, Plausibilität oder Vollständigkeit der Modellierung oder der zugrunde liegenden Daten als unzureichend bewertet werden (Wenzel et al. 2008, S. 8).

Eine Simulationsstudie beginnt mit der *Aufgabendefinition*, in der die mittels Simulation zu untersuchende Fragestellung präzisiert und ein entsprechendes Zielsystem festgelegt wird (VDI 3633 2014, S. 21 f.). In den anschließenden Phasen wird das System schrittweise in drei aufeinander aufbauenden Abstraktionsstufen modelliert (Rabe et al. 2008, S. 5 ff.):

- *Systemanalyse* → Erstellung eines *Konzeptmodells*
- *Modellformalisierung* → Entwicklung *eines formalen Modells*
- *Implementierung* → Umsetzung eines *ausführbaren Modells*

Die *Datenbeschaffung und -aufbereitung* erfolgt dabei parallel zu den drei Phasen und dient der Erfassung oder Extraktion relevanter Informationen aus bestehenden Datenquellen. Ziel ist es, die für die Modellierung erforderlichen Daten zu strukturieren und für die Simulation nutzbar zu machen (VDI 3633 2014, S. 33 f.).

Bei der *Implementierung* des Modells ist die Auswahl eines geeigneten Simulationswerkzeugs oder Programmiersprache zentral und sollte auf Grundlage der spezifischen Aufgabenstellung sowie der Kompetenz der zuständigen internen oder externen Simulationsexperten erfolgen, die für die Modellimplementierung verantwortlich sind (Wenzel et al. 2008, S. 87 ff.).

In der abschließenden Phase der Simulationsstudie werden *Experimente* durchgeführt und die Ergebnisse analysiert. Diese Simulationsexperimente bestehen in der Regel aus mehreren Simulationsläufen, bei denen Messwerte an definierten Punkten des Modells erfasst werden. Diese Daten sind anschließend im Kontext der definierten Aufgabenstellung zu interpretieren (VDI 3633 2014, S. 35 ff.). Die Ergebnisse werden für die abschließende Auswertung und Ableitung von Erkenntnissen aufbereitet, wobei weiterhin eine kritische Interpretation und Rückführung auf die ursprüngliche Fragestellung erforderlich ist (VDI 3633 2014, S. 37).

In jeder Phase kommen Methoden der Verifikation (Überprüfung der korrekten Modellbildung) und der Validierung (Bewertung der Modelltauglichkeit hinsichtlich der definierten Zielsetzung) zum Einsatz. Nachfolgend werden einige beispielhafte Methoden nach RABE et al. aufgezeigt (Rabe et al. 2008, S. 95 ff.):

- Trace-Analyse: Während der Simulation werden Trace-Dateien genutzt, um Zustandsänderungen im Zeitverlauf nachzuverfolgen.
- Sensitivitätsanalyse: Die Untersuchung analysiert die Auswirkungen von Änderungen der Eingangsdaten auf die Ausgangsdaten. Es sollte eine geringe Sensitivität der Ausgangsdaten gegenüber kleinen Eingangsvariationen gegeben sein, um stabile Ergebnisse zu gewährleisten. Andernfalls ist die Verlässlichkeit der Ergebnisse zu hinterfragen.
- Grenzwertbetrachtung: Beim Grenzwerttest werden Parameter an den Grenzen ihrer möglichen Werte betrieben, um das tatsächliche mit dem erwarteten Modellverhalten zu vergleichen.

- Test von Teilmodellen: Es wird ein modularer Aufbau umgesetzt, sodass einzelnen Teile mit einer überschaubaren Komplexität getestet werden können.
- Vergleich mit anderen Modelle: Es werden Teilergebnisse mit ähnlichen Arbeiten oder einer manuellen Berechnung verglichen.

Im nachfolgenden Kapitel wird die Generierung der synthetischen Daten zur dynamischen Bestellmengenplanung beschrieben. Dabei stellt die allgemeine Systemtheorie und das beschriebene Modellverständnis die theoretische Verankerung dar. Bei der Modellierung und Simulation der dynamischen Bestellmengenplanung wird sich an dem etablierten Vorgehensmodell von RABE et al. orientiert (Rabe et al. 2008, S. 5).

3.2 Simulationsexperimente zur dynamischen Bestellmengenplanung

In diesem Kapitel wird die Generierung der synthetischen Daten zur dynamischen Bestellmengenplanung beschrieben. Dabei strukturiert sich das Kapitel anhand der Phasen des Vorgehensmodells zur Modellierung und Simulation von logistischen Systemen nach RABE et al..

3.2.1 Zielbeschreibung und Aufgabendefinition

In diesem Kapitel werden zunächst die Ziele und Aufgaben der Modellierung und Simulation der dynamischen Bestellmengenplanung beschrieben, um aufbauend das Konzeptmodell aufzuzeigen. Dazu wurde im Kapitel 2 die dynamische Bestellmengenplanung und ihre Systemelemente als Untersuchungsbereich dieser Arbeit beschrieben. Hieraus lässt sich die Zielsetzung und aufbauend das Konzeptmodell der Simulationsstudie ableiten.

Im Allgemeinen verfolgt die Simulationsstudie das Ziel, der synthetischen Datenerzeugung zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren. Infolgedessen soll eine allgemeingültige Datenbasis geschaffen werden, auf Grundlage welcher valide und übertragbare Entscheidungsregeln sowie ein Machine Learning Modell zur kostenorientierte Auswahl entwickelt werden kann.

Dazu soll ein Simulationsmodell entwickelt werden, welches die dynamische Bestellmengenplanung als zugrunde liegendes System abbildet. Dabei müssen sämtliche Systemelemente (Faktoren) sowie die vier Heuristiken modelliert werden. Darüber hinaus wird ebenso die Vorgehensweise zur dynamischen Bestellmengenplanung sowie die Wirkzusammenhänge zwischen den Faktoren modelliert, wodurch die Beziehungen zwischen den Systemelementen definiert werden. Die Eingangsdaten der Simulationen werden

durch die unterschiedlichen Faktorstufen abgebildet. Die Zielgröße wird durch die Gesamtkosten der Planungsheuristik bestimmt. Einen Überblick über die unterschiedlichen Faktoren, Heuristiken und die Zielgrößen gibt die Abbildung 3-4.

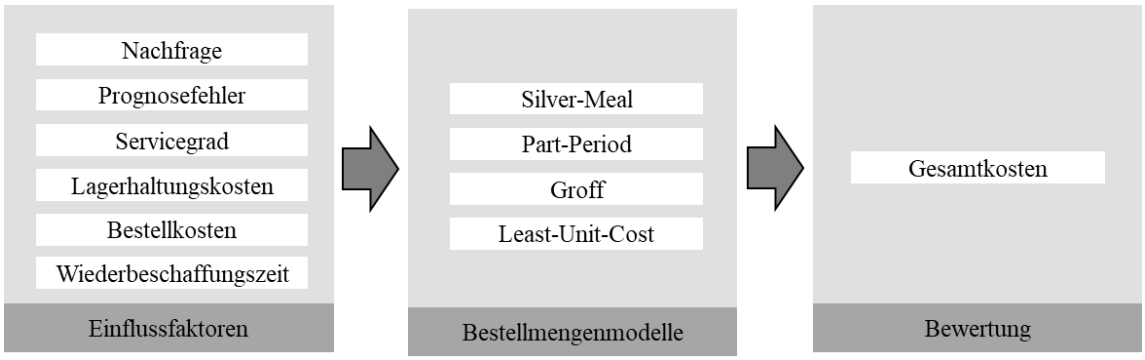

Abbildung 3-4: Konzeptmodell der dynamischen Bestellmengenplanung

Auf Grundlage des Konzeptmodells, in Form eines Systemmodells wird im nächsten Schritt das formale Modell entwickelt, welches später die Grundlage für die konkrete Implementierung des Simulationsmodells darstellt.

3.2.2 Entwicklung des formalen Modells

Das formale Modell dient der Beschreibung des Simulationsaufbaus sowie -ablaufes. Das formale Modell ergibt sich aus dem Aufbau und Ablauf der dynamischen Bestellmengenplanung in der Praxis. Die einzelnen Bestandteile und Wirkzusammenhänge wurden bereits im Kapitel 2 erläutert.

Zur synthetischen Datenerzeugung soll eine Vielzahl unterschiedlicher Faktorstufen der dynamischen Bestellmengenplanung berücksichtigt werden, sodass eine gute Übertragbarkeit gewährleistet wird. In jedem Simulationslauf dient eine Faktorkombination als Simulationsanstoß.

Beim Simulationsstart werden zunächst die ersten Perioden innerhalb der Wiederbeschaffungszeit (WBZ) auf Basis der realen Bedarfe geplant. Denn innerhalb der WBZ können die Bestellmengen nicht geändert werden, sodass sich diese Produkte im Zulauf befinden. Dieser Bereich wird eingefrorener Bereich genannt (Brahm et al. 2020, S. 89). Liegt die WBZ beispielsweise bei sieben Perioden und der Planungshorizont bei 20 Perioden, so können die Bestellmengen nur ab der achten Periode bis zur 20. Periode geplant und bestellt werden.

Anschließend wird der Bedarf im Planungshorizont ermittelt. Der Planungshorizont ist die Zeitspanne (Anzahl der Perioden) für die eine Nachfrageprognose und aufbauend eine Bestellmengenplanung erstellt wird. Der Planungshorizont sollte nicht zu groß und nicht

zu klein gewählt werden. Zum einen sinkt die Planungssicherheit mit zunehmenden Zeitabstand zur aktuellen Periode und zum anderen sollen ausreichend Perioden berücksichtigt werden, um kosteneffizient die Bestellmengen planen zu können. In diesem Kontext wurde der Planungshorizont auf 30 Perioden festgelegt, unter Berücksichtigung der maximalen WBZ von 14 Perioden (vgl. Kapitel 3.2.3 Abschnitt WBZ) und TBO von 7 Perioden (vgl. Kapitel 3.2.3 Abschnitt TBO).

Auf Basis der Nachfrageprognose für den Planungshorizont erfolgt die Prüfung ob Fehlmengen oder Bestände aus vorherigen Perioden vorliegen. Liegt eine Fehlmenge innerhalb des eingefrorenen Bereichs vor, so muss diese Fehlmenge zum nächstmöglichen Zeitpunkt bestellt werden. Dementsprechend muss der Bedarf ermittelt werden, in dem die Nachfrage in der ersten Periode des freien Planungshorizontes um die Fehlmenge erhört wird. Liegt ein ausreichender Bestand vor, so kann die Nachfrage ohne eine weitere Bestellung aus dem Bestand bedient werden.

Anschließend werden in beiden Fällen die Bestellmengen gespeichert und die Zeit t um eine Periode (t = t + 1) erhöht. In der nächsten Planungsperiode wird zunächst eine neue Nachfrageprognose für t+1 erstellt. Ebenso wird die reale Nachfrage zum Zeitpunkt t = 0 bekannt. Hieraus wird die Fehlmenge bzw. der Bestand ermittelt und in Kombination mit der Nachfrageprognose der Bedarf für den Planungshorizont berechnet. Anschließend erfolgt eine erneute Bestellmengenplanung für den freien Planungshorizont. Dieses Vorgehen wird solange iteriert bis der Simulationshorizont (Dauer der Simulation in Perioden) erreicht ist. Wurde der Simulationshorizont erreicht, so werden die Gesamtkosten der Planungsheuristik für jede der Heuristiken berechnet und gespeichert. Anschließend wird für jede Faktorkombination eine solche Simulation durchgeführt. Abschließend werden sämtliche Ergebnisse in einer CSV-Datei gespeichert.

Die Abbildung 3-5 zeigt den Simulationsablauf in Form des formalen Modells. Das formale Modell wurde mittels eines UML-Aktivitätsdiagramms modelliert. UML steht für Unified Modeling Language. Es handelt sich dabei um eine standardisierte und etablierte grafische Sprache zur Modellierung von Software-Systemen und anderen komplexen Systemen (Favre 2003, S. 2). UML wird hauptsächlich in der Softwareentwicklung verwendet, um die Struktur, das Verhalten und die Interaktionen eines Systems visuell darzustellen. Dabei modelliert das Aktivitätsdiagramm vor allem Arbeitsabläufe oder Prozesse (van Randen et al. 2016, S. 25).

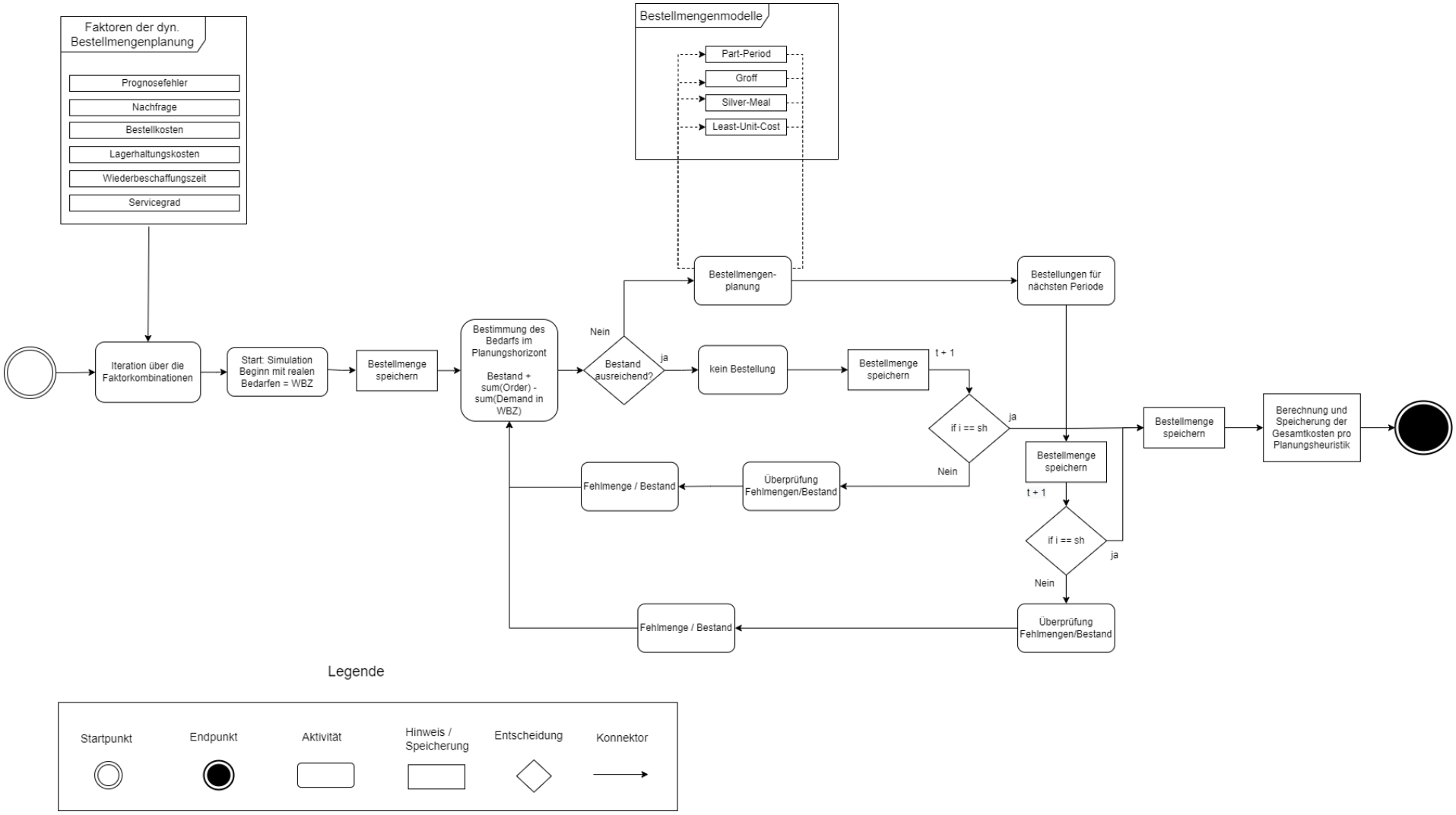

Abbildung 3-5: Formales Modell zur synthetischen Datenerzeugung

3.2.3 Modellierung der Einflussfaktoren der dynamischen Bestellmengenplanung

Im Allgemeinen können die Einflussfaktoren der dynamischen Bestellmengenplanung deterministisch und stochastisch modelliert werden. Unterliegt ein Einflussfaktor einer Unsicherheit, so wird er stochastisch genannt. Die stochastischen Einflussfaktoren müssen mittels geeigneten Wahrscheinlichkeitsfunktionen modelliert werden. In diesem Rahmen sind die Faktoren „Nachfrage“ und „Prognosefehler“ stochastisch. Wohingegen die Faktoren Bestellkosten, Lagerhaltungskosten, Servicegrad und Wiederbeschaffungszeit als deterministisch ohne Unsicherheit modelliert werden (Stadtler et al. 2015, S. 45; Fildes und Kingsman 2011, S. 488). Nachfolgend werden zunächst die zentralen Wahrscheinlichkeitsverteilungen erläutert, um darauf aufbauend die Nachfrage und die Prognosefehler zu modellieren.

Wahrscheinlichkeitsverteilungen zur Modellierung

Grundsätzlich kann zwischen diskreten und stetigen Wahrscheinlichkeitsverteilungen unterschieden werden. Diese Unterscheidung basiert auf der Art der Werte, die eine Zufallsvariable annehmen kann.

Diskrete Wahrscheinlichkeitsverteilungen beschreiben Zufallsvariablen, die abzählbar viele Ausprägungen besitzen. Typischerweise handelt es sich hierbei um ganze Zahlen, wie beispielsweise die Anzahl von Treffern in einem Experiment (z. B. Binomialverteilung), die Anzahl von Ankünften in einem Wartesystem (z. B. Poissonverteilung) oder die Anzahl der Versuche bis zum ersten Erfolg (z. B. geometrische Verteilung). Die Wahrscheinlichkeit, dass eine diskrete Zufallsvariable genau einen bestimmten Wert annimmt, kann direkt durch die sogenannte Wahrscheinlichkeitsfunktion angegeben werden. (Thomopoulos 2018, S. 28)

Im Gegensatz dazu beschreiben stetige Wahrscheinlichkeitsverteilungen Zufallsvariablen, die beliebige Werte innerhalb eines Intervalls annehmen können, typischerweise in reellen Zahlen. Solche Verteilungen eignen sich zur Modellierung von kontinuierlich messbaren Größen wie Zeit, Temperatur oder Gewicht. Da die Wahrscheinlichkeit, dass eine stetige Zufallsvariable genau einen bestimmten Wert annimmt, gleich null ist, wird hier die Wahrscheinlichkeitsdichtefunktion verwendet. Wahrscheinlichkeiten werden über Intervalle definiert und ergeben sich durch Integration der Dichtefunktion über das entsprechende Intervall. (Bonamente 2017, S. 19) Nachfolgend werden die zentralen Wahrscheinlichkeitsverteilungen zur Modellierung logistischer Systeme erläutert.

Bernoulli-Verteilung

Die Bernoulli-Verteilung ist eine diskrete Wahrscheinlichkeitsverteilung und beschreibt ein Zufallsexperiment mit genau zwei möglichen Ausgängen: "Erfolg" (1) mit Wahrscheinlichkeit p und "Misserfolg" (0) mit Wahrscheinlichkeit 1 − p, wobei 0 ≤ p ≤ 1. (Georgii 2009, S. 32 ff.)

Eigenschaften:

- Zufallsvariable: $x \in \{0,1\}$
- Wahrscheinlichkeitsfunktion:

$$P(X = x) = \begin{cases} p, & wenn\ x = 1 \\ 1 - p, & wenn\ x = 0 \\ 0, & sonst \end{cases}$$

- Erwartungswert: $E(X) = p$
- Varianz: $Var(X) = p(1 - p)$

Anwendung: Modellierung einfacher Ja/Nein-Entscheidungen wie Münzwurf oder Qualitätsprüfung

Binomialverteilung

Die Binomialverteilung generalisiert die Bernoulli-Verteilung auf den Fall von n unabhängigen Wiederholungen desselben Bernoulli-Experiments mit konstanter Erfolgswahrscheinlichkeit p. Sie beschreibt die Verteilung der Anzahl an Erfolgen in n Versuchen. (Bonamente 2017, S. 35 f.)

Eigenschaften:

- Zufallsvariable: $x \in \{0,1, \dots, n\}$
- Wahrscheinlichkeitsfunktion:

$$P(X = k) = \binom{n}{k} p^k\ (1 - p)^{n-k} \qquad für\ k = 0,1, \dots, n$$

- Erwartungswert: $E(X) = n * p$
- Varianz: $Var(X) = n * p(1 - p)$

Anwendung: Qualitätskontrolle (z. B. Anzahl defekter Produkte in einer Stichprobe), Umfragen (z. B. Anzahl der Befürworter), Wahrscheinlichkeitsrechnung (z. B. Anzahl der Sechsen beim Würfeln)

Poisson-Verteilung

Die Poisson-Verteilung beschreibt die Anzahl von Ereignissen, die in einem festen Zeit- oder Raumbereich bei gegebener durchschnittlicher Ereignisrate $\lambda > 0$ zufällig auftreten. Sie ist besonders geeignet für seltene Ereignisse. (Bonamente 2017, S. 45 f.)

Eigenschaften:

- Zufallsvariable: $x \in \mathbb{N}_0$
- Wahrscheinlichkeitsfunktion:

$$P(X = k) = \frac{\lambda^k * e^{-\lambda}}{k!} \qquad \text{für } k = 0,1,2, \ldots$$

- Erwartungswert: $E(X) = \lambda$
- Varianz: $Var(X) = \lambda$

Anwendung: Anzahl von Anrufen pro Minute in einem Callcenter, Defekte pro Meter Kabel, Verkehrsunfälle pro Tag

Geometrische Verteilung

Die geometrische Verteilung beschreibt die Anzahl der Bernoulli-Experimente bis zum ersten Erfolg. Dabei ist jedes Experiment unabhängig und die Erfolgswahrscheinlichkeit p bleibt konstant. (Thomopoulos 2018, S. 33 f.)

Eigenschaften:

- Zufallsvariable: $x \in \{1, 2, 3, \ldots\}$
- Wahrscheinlichkeitsfunktion:

$$P(X = k) = (1 - p)^{k-1} * p \qquad \text{für } k = 1,2, \ldots$$

- Erwartungswert: $E(X) = \frac{1}{p}$
- Varianz: $Var(X) = \frac{1-p}{p^2}$

Anwendung: Anzahl von Versuchen bis zum ersten Treffer (z. B. beim Glücksrad, Software-Testfehler, Kaufentscheidung)

Normalverteilung (Gauß-Verteilung)

Die Normalverteilung ist eine symmetrische, glockenförmige Verteilung, die in der Natur und in vielen statistischen Phänomenen häufig auftritt. (Thomopoulos 2018, S. 46 f.)

Eigenschaften:

- Zufallsvariable**:** $x \in \mathbb{R}$
- Parameter**:** Erwartungswert $\mu \in \mathbb{R}$; $Varianz\ \sigma^2 > 0$
- Dichtefunktion**:**

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

- Erwartungswert**:** $E(X) = \mu$
- Varianz**:** $Var(X) = \sigma^2$

Anwendung: Messfehler, Körpergrößen, IQ-Werte, Finanzmodelle, Naturprozesse

Exponentialverteilung

Die Exponentialverteilung modelliert die Zeit zwischen zwei aufeinanderfolgenden, zufällig eintretenden Ereignissen in einem Poisson-Prozess und unabhängig von der Vergangenheit. (Forbes et al. 2010, S. 88 ff.)

Eigenschaften:

- Zufallsvariable**:** $x \in [0, \infty)$
- Parameter**:** Rate $\lambda > 0$
- Dichtefunktion:

$$f(x) = \lambda e^{-\lambda x}, \qquad x \geq 0$$

- Erwartungswert**:** $E(X) = \frac{1}{\lambda}$
- Varianz**:** $Var(X) = \frac{1}{\lambda^2}$

Anwendung: Lebensdauer von Geräten (wenn kein Verschleiß), Wartezeiten, Modellierung seltener Ereignisse

Gammaverteilung

Die Gammaverteilung ist eine Verallgemeinerung der Exponentialverteilung und modelliert die Wartezeit bis zum k-ten Ereignis in einem Poisson-Prozess. Sie ist sehr flexibel und dient als Grundlage für viele andere Verteilungen. (Forbes et al. 2010, S. 109 ff.)

Eigenschaften:

- Zufallsvariable**:** x ∈ $[0, \infty)$
- Parameter**:** Formparameter $k > 0$, Skalenparameter $\theta > 0$ ($\alpha = k$, $\beta = 1/\theta$)

$$f(x) = \frac{x^{k-1} e^{-x/\theta}}{\theta^k \Gamma(k)},\ x \geq 0$$

- Erwartungswert**:** $E(X) = k * \theta$
- Varianz**:** $Var(X) = k * \theta^2$

Anwendung: Warteschlangentheorie, Lebensdauermodelle, Bayesianische Statistik, Meteorologie

Weibull-Verteilung

Die Weibull-Verteilung wird in der Zuverlässigkeitstheorie und Lebensdaueranalyse verwendet. Sie ist flexibel, um beispielsweise sowohl alterungsfreie als auch alterungsabhängige Prozesse zu modellieren. (Forbes et al. 2010, S. 193 ff.)

Eigenschaften:

- Zufallsvariable: $x \in [0, \infty)$
- Parameter: Formparameter $k > 0$, Skalenparameter $\lambda > 0$
- Dichtefunktion:

$$f(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^2}\ x \geq 0$$

- Erwartungswert: $E(X) = \lambda\Gamma(1+1/k)$
- Varianz: $Var(X) = \lambda^2[\Gamma\left(1 + \frac{2}{k}\right) - \left(\Gamma\left(1 + \frac{1}{k}\right)\right)^2]$

Anwendung: Lebensdauer von Maschinen, Zuverlässigkeitsanalysen, Materialfestigkeit

Die Tabelle 3-1 fasst die vorgestellten Wahrscheinlichkeitsverteilungen nach ihren Typ, Parameter und typischen Fragestellungen zusammen.

Tabelle 3-1: Übersicht relevanter Wahrscheinlichkeitsverteilungen

Verteilung	**Typ**	**Parameter**	**Typische Frage**
Bernoulli	diskret	p	Erfolg oder Misserfolg?
Binomial	diskret	n, p	Wie viele Erfolge in n Versuchen?
Poisson	diskret	λ	Wie viele Ereignisse in fester Zeit?
Geometrisch	diskret	p	Wie viele Versuche bis zum ersten Erfolg?
Normal (Gauß)	stetig	μ, σ^2	Wie ist die Verteilung um einen Mittelwert?
Exponential	stetig	λ	Wie lang ist die Zeit bis zum nächsten Ereignis?
Gamma	stetig	k, θ (oder α, β)	Wie lang bis zu k Ereignissen?
Weibull	stetig	λ, k	Wie lange hält ein Bauteil?

Modellierung: Nachfrageverläufe

Zur Modellierung der Nachfrage für die synthetische Datenerzeugung wird sich an den vier definierten Kategorien von Syntetos et al. orientiert (vgl. Kapitel 2.1.2). Um eine große Bandbreite an unterschiedlichen Nachfrageverläufen zu berücksichtigen, werden sowohl der $VarK^2$, als auch der *ADI* mit jeweils unterschiedlichen Werten variiert. Dabei wird als $VarK^2$ für einen regelmäßigen Bedarf der Wert 0,01, für einen schwankenden Bedarf der Wert 0,49 und für einen unregelmäßigen Bedarf der Wert 1 gewählt. Für den *ADI* werden 0 % Nullperioden bzw. 100 % nicht Null-Perioden (NNPerioden) mit einem *ADI* von 0, 80 % NNPerioden mit einem *ADI* von 1,25 und 60 % NNPerioden mit einem *ADI* von 1,67 modelliert. Wobei die Nullperioden die Anzahl der Perioden ohne Nachfrage beschreiben. Mittels dieser Parameterwahl wird jede der vier Bedarfsarten nach Syntetos et al. berücksichtigt (Syntetos et al. 2005, S. 499 ff.). Diese Klassifizierung zeigt die nachfolgende Abbildung 3-6 auf, wobei die neun Punkte die neun Nachfrageverläufe mit den entsprechenden Kriterienwerten darstellen.

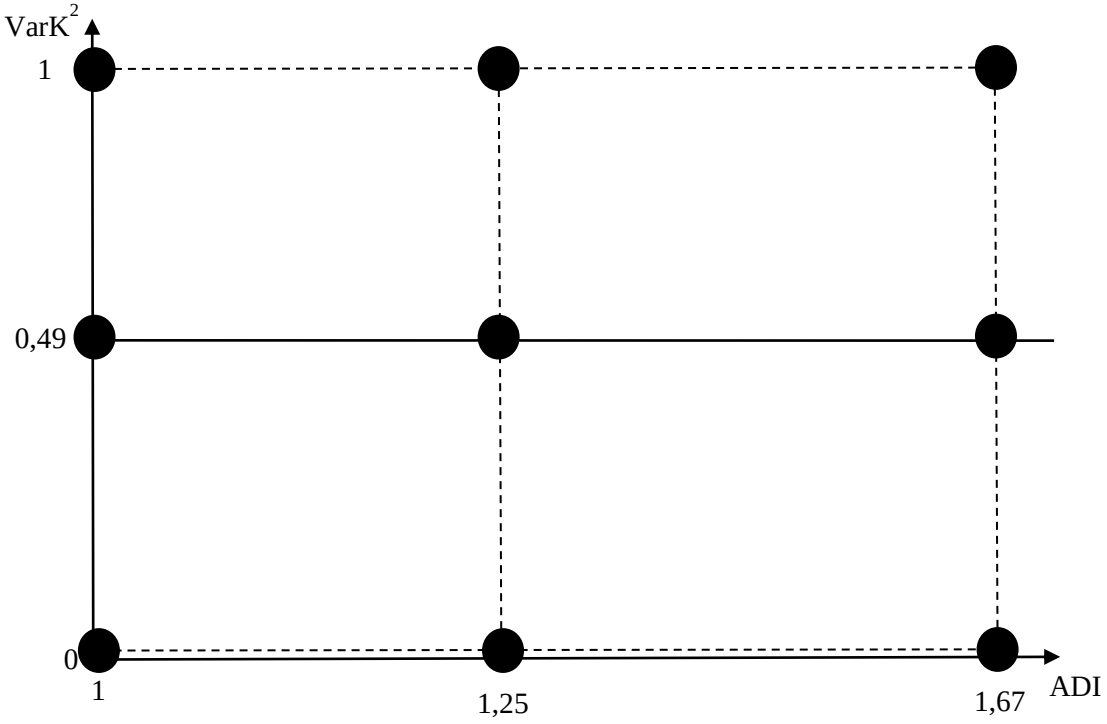

Abbildung 3-6: Klassifizierung der Nachfrageverläufe zur synthetischen Datenerzeugung

Nachdem definiert wurde, welche Nachfrageverläufe erzeugt werden sollen, um eine große Bandbreite und Diversität der Nachfrage zu modellieren, muss die passende Wahrscheinlichkeitsverteilung zur Modellierung ausgewählt werden. Im Allgemeinen ist die Auswahl der richtigen Wahrscheinlichkeitsverteilung ein häufig diskutiertes Problem in der Wissenschaft und Praxis (Thorstenson 2018, S. 153 ff.). Häufig wird die Normalverteilung angewandt. Diese ist vorwiegend für die Abbildung von Nachfrageverläufen mit kleinem Variationskoeffizient anwendbar (Tadikamalla 1984, S.575). Da die Normalverteilung sowohl negative Werte erlaubt, als auch symmetrisch ist, ist die Anwendung für die Modellierung von Nachfrageverläufen ungeeignet (Thorstenson 2018, S. 153 ff.; Ramaekers und Janssens 2008, S. 484; Lau und Lau 2003, S. 149 ff.). Zur Auswahl der passenden Wahrscheinlichkeitsverteilung $f\ (X)$ werden nachfolgend die zentralen Anforderungen aufgezeigt, welche eine Wahrscheinlichkeitsverteilung erfüllen muss, um die zuvor definierten Bedarfsverläufe zu modellieren.

1. Es sollen nur ganzzahlige positive Nachfragen in $\mathbb{N}$ möglich sein
2. Der Erwartungswert $E(X)$ soll definierbar sein
3. Der Variationskoeffizient VarK soll definierbar sein, ohne dass sich der Erwartungswert ändert. Somit muss für zwei Zufallsvariablen X_1 und X_2 mit Variationskoeffizienten VarK_1 und VarK_2 gelten: $E(X_1) = E(X_2)$
4. Die Anzahl der Nullperioden bzw. das Average demand interval soll definierbar sein.

Eine Wahrscheinlichkeitsverteilung, welche die Kriterien eins bis drei erfüllt und auch in der Literatur häufig anstelle einer Normalverteilung angewandt wird, ist die Gammaverteilung (Burgin 1975, S. 507 ff.; Snyder 1984, S. 373 ff.; Yeh et al. 1997, 1197 ff.; Axsäter 2015, S. 72; Silver et al. 2016, S. 275 f.).

Gammaverteilung

Wie bereits im Unterkapitel Wahrscheinlichkeitsverteilungen beschrieben, wird die Gammaverteilung durch den Formparameter $\alpha > 0$ und dem Skalenparameter $\beta > 0$ bestimmt. Der Erwartungswert und die Varianz der Gammaverteilung sind gegeben durch:

$$E(X) = \frac{\alpha}{\beta} \tag{3-1}$$

$$Var(X) = \frac{\alpha}{\beta^2} \tag{3-2}$$

Wobei es sich bei X um eine mit der Gammaverteilung verteilte Zufallsvariable handelt (Casella und Berger 2002, S. 99). Durch das Umstellen der Formeln für Erwartungswert und Varianz lassen sich anhand eines gegebenen Erwartungswertes und Variationskoeffizient die Werte für die beiden Parameter α und β bestimmen. So gilt:

$$\beta = \frac{\alpha}{E(X)} \tag{3-3}$$

$$\alpha = Var(X) * \beta^2 \tag{3-4}$$

Einsetzen von (3-4) in (3-3) liefert:

$$\beta = \frac{Var(X)\beta^2}{E(X)} \leftrightarrow E(X)\beta = Var(X)\beta^2 \leftrightarrow Var(X) = \frac{E(X)}{\beta} \tag{3-5}$$

$$\rightarrow \beta = \frac{E(X)}{Var(X)}$$

und Einsetzen von (3-5) in (3-3) ergibt:

$$\alpha = \frac{E(X)^2}{Var(X)} \tag{3-6}$$

Da der Variationskoeffizient frei wählbar sein soll, muss nun anhand des gewünschten Variationskoeffizienten auf die implizierte Varianz geschlossen werden. Auch dies ist durch Umstellen möglich:

$$VarK(X) = \frac{\sqrt{Var(X)}}{E(X)} \quad \leftrightarrow \quad VarK(X)E(X) = \sqrt{Var(X)} \tag{3-7}$$
$$Var(X) = \left(VarK(X)E(X)\right)^2$$

Einsetzen in die Formeln für α und β ergeben die finalen Formeln (3-8) und (3-9) zur Bestimmung von α und β anhand eines gegebenen Erwartungswertes $E(X)$ und Variationskoeffizienten $VarK$:

$$\alpha = \frac{E(X)^2}{(VarK * E(X))^2} \tag{3-8}$$

$$\beta = \frac{(VarK * E(X))^2}{E(X)} = VarK^2 * E(X) \tag{3-9}$$

Somit erfüllt die Gammaverteilung die ersten drei definierten Anforderungen zur Modellierung der Nachfrageverläufe.

Einbeziehen der Nullperioden

Für die Modellierung des ADIs muss die vierte Anforderung erfüllt werden, sodass eine definierte Anzahl an Perioden ohne Nachfrage modelliert werden kann. Um auch diese Nachfrageverläufe in den synthetischen Daten modellieren zu können, sollen die Daten für jeweils 60 %, 80 % und 100 % NNPerioden erzeugt werden. An dieser Stelle reicht eine Gammaverteilung nicht aus. Im Allgemeinen können zur Modellierung der Perioden ohne Nachfrage „hurdle“ und „zero-inflated“ Modelle verwendet werden (Hilbe 2014, S. 184).

Beim Hurdle-Modell wird das Wahrscheinlichkeitsmodell in zwei separate Wahrscheinlichkeitsmodelle aufgeteilt. Der erste Teil stellt die Hürde dar und definiert, ob eine Variable den Wert 0 oder einen positiven Wert annimmt. Nachdem die Hürde durch den ersten Teil überwunden wurde, modelliert die zweite Wahrscheinlichkeitsverteilung des Hurdle-Modells die Verteilung der nicht null Werte (Hilbe 2014, S. 184; Baetschmann und Winkelmann 2017, S. 7174 ff.).

Das Zero-inflated Modell besteht ebenfalls aus zwei Teilen, wobei der erste Teil ein Binärteil ist, welcher die Null-Daten modelliert, wohingegen der zweite Teil die restlichen Werte modelliert. Während beim Hurdle-Modell die Null-Daten ausschließlich durch den ersten Teil bestimmt werden, kann bei dem Zero-inflated Modell auch der zweite Teil Null-Daten generieren, wodurch die Wahrscheinlichkeit einer Null von beiden Verteilungen abhängt (Hilbe 2014, S. 197; Freitas Costa et al. 2021, S. 1131 ff.).

Im Allgemeinen können mit beiden Modellen die Anzahl der Perioden ohne Bedarfe modelliert werden. Die klassischen Hurdle-Modelle eignen sich besonders für sequenzielle Entscheidungsprozesse (Frees 2012, S. 355) und sind somit vorteilhaft, wenn eine exakte Anzahl an Perioden ohne Nachfrage modelliert werden soll. Folglich wird für die Modellierung der Perioden ohne Nachfrage ein Hurdle-Modell ausgewählt. Wobei die Hürden mit einer Binomialverteilung modelliert werden (siehe Unterkapitel Wahrscheinlichkeitsverteilungen – Binomialverteilung). Resultierend wird zunächst die Hürde mit der Binomialverteilung modelliert. Tritt eine Nachfrage auf, so wird anschließend die Gammaverteilung verwendet um die Nachfrage in der Periode zu modellieren. Nachfolgend wird die konkrete Modellierung der Nachfrage für die synthetische Datenerzeugung erläutert.

Modellierung: Nachfrage

Für die Modellierung wird eine mittlere Nachfrage von 2000 Einheiten unter Einsatz des quadrierten Variationskoeffizienten 0,01, 0,49 und 1, sowie 60 %, 80 % und 100 % NNPerioden verwendet. Für den Fall von 100 % NNPerioden wird die Gammaverteilung verwendet, während im Falle von 60 % und 80 % NNPerioden ein Hurdle-Modell aus einer Binominal- und einer Gammaverteilung verwendet wird.

Die Parameter α und β der Gammaverteilung werden anhand der Formeln (3-8) und (3-9) aus dem gewählten Erwartungswert und dem gewählten Variationskoeffizienten bestimmt. Bei 60 % und 80 % NNPerioden wird der Erwartungswert für die Bestimmung von α und β um folgende Formel angepasst:

$$E(X) = 2000 * \frac{1}{1-p}$$

Wobei p den gewählten prozentualen Anteil an Nullperioden entspricht. Somit werden die Parameter in Tabelle 3-2 für die Simulation verwendet.

Tabelle 3-2: Parameterkonfiguration der Gammaverteilung für die Nachfrage

ADI	**1 (100 %-NNPerioden)**		**1,25 (80 %- NNPerioden)**		**1,67 (60 %- NNPerioden)**	
VarK²	α	β	α	β	α	β
0,01	100	20	100	25	100	33,33
0,49	2,041	980	2,041	1225	2,041	1633,33
1	1	2000	1	2500	1	3333,33

Die Abbildung 3-7 (VarK² = 0,01 / ADI = 1) und Abbildung 3-8 (VarK² = 1 / ADI = 1,25) zeigen zwei exemplarische Nachfrageverläufe unter Verwendung der aufgezeigten Parameter (vgl. Tabelle 3-2). Das linke Diagramm in den Abbildungen zeigt den Zeitreihenverlauf, wobei die Nachfragemenge über die Zeit abgebildet ist. Das rechte Diagramm zeigt die Dichtefunktion des Nachfrageverlaufs in Form eines Histogrammes.

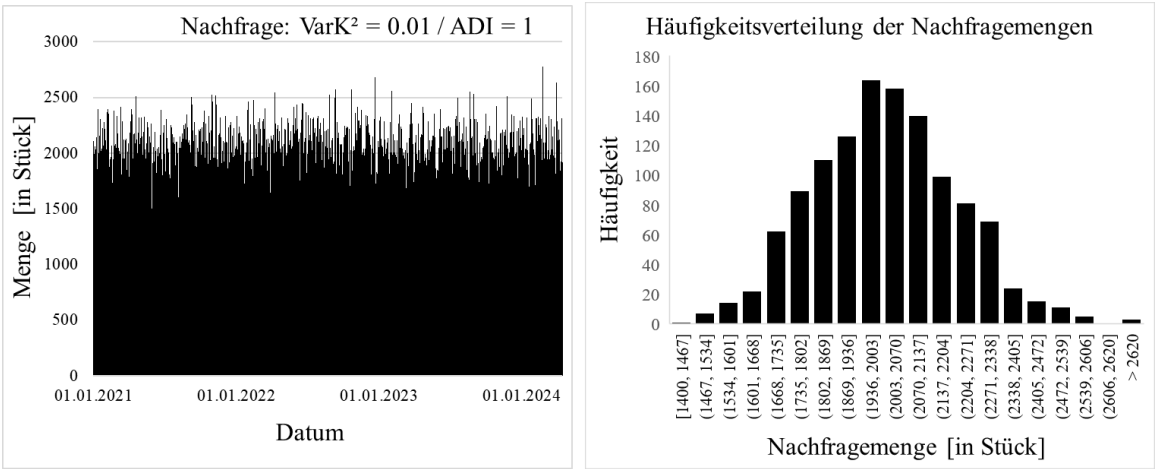

Abbildung 3-7: Synthetische Nachfrage - VarK² = 0,01 / ADI = 1

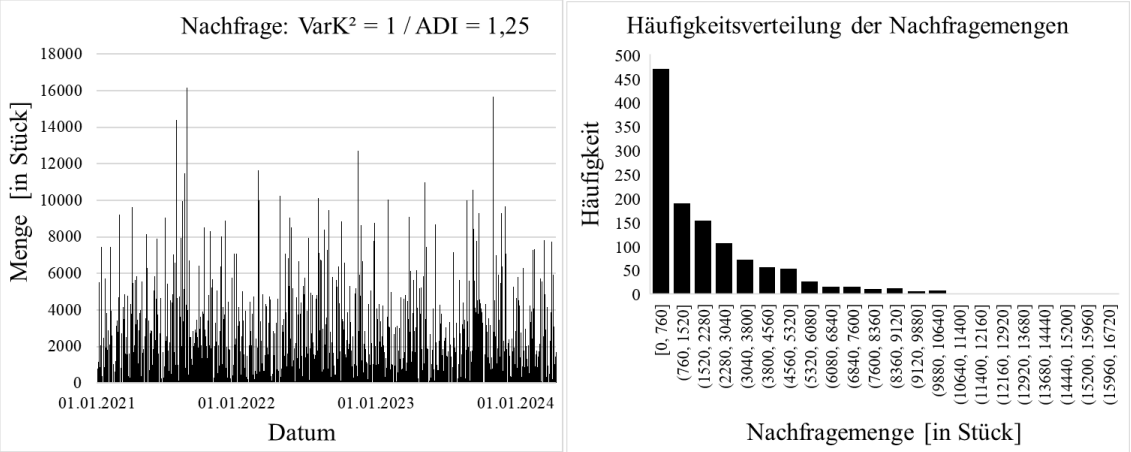

Abbildung 3-8: Synthetische Nachfrage: VarK² = 1 / ADI = 1,25

Modellierung: Prognosefehler

Die Prognose der zukünftigen Bedarfe stellt die Planungsgrundlage für die dynamische Bestellmengenplanung dar (vgl. Kapitel 2.1). Demzufolge wird im nächsten Schritt der Prognosefehler modelliert werden. Das Einbeziehen von Prognosefehlern in das Simulationsdesign ist ein wichtiger Aspekt, um realistische und aussagekräftige Ergebnisse zu erzielen.

Die grundlegende Klassifizierung von Prognosefehlern wurde bereits im Kapitel 2.1.1 beschrieben. Demzufolge sollen im Rahmen der synthetischen Datenerzeugung die drei systematischen Fehlertypen *Überschätzung*, *Unterschätzung* und *Rauschen* modelliert werden. Der Erwartungswert der Über- bzw. Unterschätzung wird als prozentuale Angabe definiert und wird durch die drei unterschiedlichen Faktorstufen 10 %, 20 % und 40 % abgebildet.

Demgegenüber hat der Prognosefehler Rauschen den Erwartungswert gleich null, sodass sich die Prognosefehler aus den einzelnen Perioden aufsummiert null ergeben. Zur Modellierung der Faktorstufen des Prognosefehlers Rauschen bietet sich der Variationskoeffizient an, sodass unterschiedliche Abweichung zu der realen Nachfrage auftreten. Damit eine große Bandbreite an unterschiedlichen Faktorstufen abgebildet wird, werden die Faktorstufen 0,1, 0,2 und 0,4 ausgewählt. Damit nicht in jeder Periode die gleiche Abweichung zur realen Nachfrage auftreten, muss der Prognosefehler einer Unsicherheit unterliegen. Wie auch bei der Modellierung der Nachfrage wird dazu eine Wahrscheinlichkeitsverteilung verwendet.

Für die Auswahl der Wahrscheinlichkeitsverteilung bei den Prognosefehlern **Über**- und **Unterschätzung** gelten ebenso die ersten drei Anforderungen von der Nachfrage (vgl. Abschnitt „Modellierung: Nachfrageverläufe“). Daher wird für die Modellierung der Prognosefehler Über- und Unterschätzung ebenso die Gammaverteilung verwendet. Zur

Abbildung der Prognosefehler *Überschätzung* und *Unterschätzung* werden definierte Zufallszahlen aus einer Wahrscheinlichkeitsverteilung gezogen. Diese werden im Fall der Überschätzung zu den realen Werten addiert und im Fall der Unterschätzung von diesen subtrahiert. Die Vorgehensweise zur Parametrisierung der Gammaverteilung erfolgt analog der Parametrisierung für die Nachfrage mittels der α- und β-Parameter. Wobei der α-Parameter konstant auf 1 gesetzt wird und der β-Parameter variiert wird, damit kleine Werte (Prognosefehler) häufiger als große Werte (Prognosefehler) auftreten. In diesem Fall gleicht die Gammaverteilung der Exponentialverteilung (vgl. Unterkapitel Wahrscheinlichkeitsverteilungen – Exponentialverteilung).

Durch die Subtraktion der Zufallszahl von der realen Nachfrage können bei einer Unterschätzung negative Werte auftreten. Dies ist aus praktischer Sicht nicht möglich. Daher müssen die Werte $x < 0$ auf null gesetzt werden. Die definierten Parameter der Gammaverteilung für die Prognosefehler Über- und Unterschätzung sind in der nachfolgenden Tabelle 3-3 ersichtlich, unter Berücksichtigung des Erwartungswertes $E(x) = 2000$ für die reale Nachfrage.

Tabelle 3-3: Parameter der Gammaverteilung für den Prognosefehler Über- und Unterschätzung

Prognosefehler	**Über- und Unterschätzung**	
	α	β
10 %	1	200
20 %	1	400
40 %	1	800

Demgegenüber müssen beim Prognosefehler **Rauschen** sowohl positive wie auch negative Abweichungen modelliert werden. Für den Prognosefehler Rauschen ergeben sich folgenden Anforderungen an die Wahrscheinlichkeitsverteilung:

1) Die Wahrscheinlichkeitsverteilung muss sowohl negative wie auch positive Zufallszahlen generieren können
2) Der Erwartungswert $\tilde{x}$ soll definierbar sein, sodass $E(X) = \tilde{x}$
3) Der Variationskoeffizient VarK soll definierbar sein, ohne dass sich der Erwartungswert ändert. Somit muss für zwei Zufallsvariablen X_1 und X_2 mit Variationskoeffizienten VarK_1 und VarK_2 gelten:

$$E(X_1) = E(X_2)$$

Diese Anforderungen können mit der Normalverteilung mit den beiden Parametern μ und σ² modelliert werden. Auch hier gilt, wie bei den systematischen Unterschätzungen, dass negative Nachfragewerte nicht zulässig sind und auf null gesetzt werden. Die definierten Parameter der Normalverteilung für die Prognosefehler Rauschen sind in der nachfolgenden Tabelle 3-4 ersichtlich, unter Berücksichtigung des Erwartungswertes E(x) = 2000 von der realen Nachfrage.

Tabelle 3-4: Parameter der Normalverteilung für den Prognosefehler Rauschen

Prognosefehler	**Rauschen**	
	μ	σ
0,1	0	200
0,2	0	400
0,4	0	800

Modellierung: Servicegrad

Auf Basis des Prognosefehlers können Fehlmengen in der dynamischen Bestellmengenplanung auftreten (vgl. Kapitel 2.2). In diesem Rahmen müssen Sicherheitsbestände berücksichtigt werden, welche i. d. R. über definierte Servicegrade dimensioniert werden. Für die synthetische Datenerzeugung wird der α-Servicegrad verwendet (siehe Kapitel 2.2.1). Die Servicegrade liegen typischerweise im Bereich zwischen 90 % bis 99 % (Warwas 2015, S. 344), wobei der optimale Servicegrad von mehreren Faktoren, wie den Lagerkosten, dem Kundenverhalten bei einer Fehlmenge oder der Wettbewerbssituation abhängt (Warwas 2015, S. 344). Im Rahmen der Simulationsexperimente werden Servicegrade von 97,5 %, 99 % und 100 % als Faktorstufen berücksichtigt. Der Servicegrad wird im Rahmen der vorliegenden Arbeit als deterministischer Faktor definiert, sodass ex post die Höhe des Sicherheitsbestands in Abhängigkeit des gewählten Servicegrads definiert wird.

Modellierung: Time Between Order (TBO)

Ein weiterer wichtiger Systemfaktor ist die Kostenstruktur, bestehend aus den Bestellkosten und Lagerhaltungskosten (vgl. Kapitel 2.2.1). Anstelle der isolierten Betrachtung der Kostenparameter wird das Verhältnis zwischen der durchschnittlichen Nachfrage, den

Bestellkosten und den Lagerhaltungskosten in Form der TBO angewandt. Da die durchschnittliche Nachfrage in dieser Analyse konstant ist, drückt die TBO das Verhältnis zwischen den Bestell- und den Lagerhaltungskosten aus. Interpretieren lässt sich die TBO hierbei als die Anzahl an Perioden, die eine Bestellmenge durchschnittlich abdeckt (vgl. Kapitel 2.2.1) (Callarman und Hamrin 1984, S. 41). Dabei basiert die Berechnung auf der Andler-Formel sowie auf einer konstanten Nachfrage. Nachfolgend ist die Formel zur Berechnung der TBO dargestellt:

$$TBO = \sqrt{\frac{2 * K_B}{d * K_{LH}}} \tag{3-10}$$

mit

d = Nachfrage
K_{LH} = Lagerhaltungskosten
K_B = Bestellkosten

Für die vorliegende Arbeit werden als TBO 2, 3, 5 und 7 Perioden gewählt. Dies entspricht der Verwendung von TBO in ähnlicher Literatur (Fildes und Kingsman 2011, S. 488; Jeunet 2006, S. 414). Aus den gewählten TBO ergeben sich die in Tabelle 3-5 dargestellten Verhältnisse zwischen den Lagerhaltungs- und den Bestellkosten.

Tabelle 3-5: TBO - Faktorstufen

TBO	**K_{LH}/K_B**
2 Perioden	0,01 / 40
3 Perioden	0,01 / 90
5 Perioden	0,01 / 250
7 Perioden	0,01 / 490

Modellierung: Wiederbeschaffungszeit

Als letzter Parameter wird die Wiederbeschaffungszeit (WBZ) betrachtet (vgl. Kapitel 2.2). Dabei werden Wiederbeschaffungszeiten von 2, 5, 7 und 14 Perioden gewählt, um typische Wiederbeschaffungszeiten abzudecken.

3.2.4 Experimentplanung und -durchführung

Resultierend ergibt sich aus den beschriebenen Faktoren und Faktorstufen der Experimentplan zur synthetischen Datenerzeugung der dynamischen Bestellmengenplanung. Ein Überblick über den vollständigen Experimentplan ist in der nachfolgenden Tabelle 3-6 ersichtlich.

Tabelle 3-6: Experimentplan - synthetische Datenerzeugung zur dynamischen Bestellmengenplanung

Streuung (Nachfrage)	**NNPerioden (Nachfrage)**	**Prognosefehler**	**Servicegrad**	**TBO (K_{LH} / K_B)**	**WBZ**	**Planungsverfahren**
1) $VarK^2 = 0{,}01$	1) 60% (ADI = 1,67)	1) VarK = 0,1	97,5 %	2 Perioden (0,01/40)	2 Perioden	Groff
2) $VarK^2 = 0{,}49$	2) 80% (ADI = 1,25)	2) VarK = 0,2	99 %	3 Perioden (0,01/90)	5 Perioden	Silver-Meal
3) $VarK^2 = 1$	3) 100% (ADI = 1)	3) VarK = 0,4	100 %	5 Perioden (0,01/250)	7 Perioden	Part-Period
		4) +10%		7 Perioden (0,01/490)	14 Perioden	Least-Unit-Cost
		5) +20%				
		6) +40%				
		7) -10%				
		8) -20%				
		9) -40%				

Aus dem Experimentplan folgen insgesamt 3.888 Faktorkombinationen für die vier Heuristiken. Zur Implementierung des Simulationsmodells wurde die Programmiersprache Python gewählt. Dabei wird das formale Modell zur synthetischen Datenerzeugung in ein ausführbares Modell mittels Python überführt. Neben dem Simulationsablauf werden die zuvor beschrieben Faktoren und Faktorstufen in das Simulationsmodell integriert.

Im nächsten Schritt muss die Dauer des Simulationsexperiments definiert werden. Dabei gilt es zwischen Aufwand und Nutzen abzuwägen. Denn eine hohe Simulationsdauer bietet eine hohe Robustheit gegen zufällige Ereignisse oder Ausreißern, jedoch erfordert dies mehr Rechenleistung (Wenzel et al. 2008, S. 139).

Zur Bestimmung eines ausreichenden Simulationshorizontes muss die Einschwingphase und die Abklingphase des Simulationsexperiments berücksichtigt werden. Die Einschwingphase ist der Zeitraum zwischen dem Start der Simulation und dem Erreichen eines stabilen, typischen Betriebsverhaltens (stationärer Zustand) (Wenzel et al. 2008, S. 140 f.). Wohingegen die Abklingphase das Ende der Simulation darstellt. Dieser Zeiträume werden bei der späteren Ergebnisauswertung nicht berücksichtigt. Zur Definition der Einschwing- und Abklingphase sowie der Simulationsdauer können statistische und grafische Methoden eingesetzt werden, wobei die grafischen Methoden in der Regel verlässlichere Ergebnisse liefert (Wenzel et al. 2008, S. 141 f.).

Zur Bestimmung der Einschwing- und Abklingphase wird zunächst ein Simulationshorizont von 1150 Perioden ausgewählt, um anschließend die Kosten im Zeitverlauf zu analysieren. Die Abbildung 3-9 zeigt die Kosten pro Perioden im Zeitverlauf anhand von zwei Simulationsexperimenten mit zwei unterschiedlichen Faktorkombination für die vier Heuristiken.

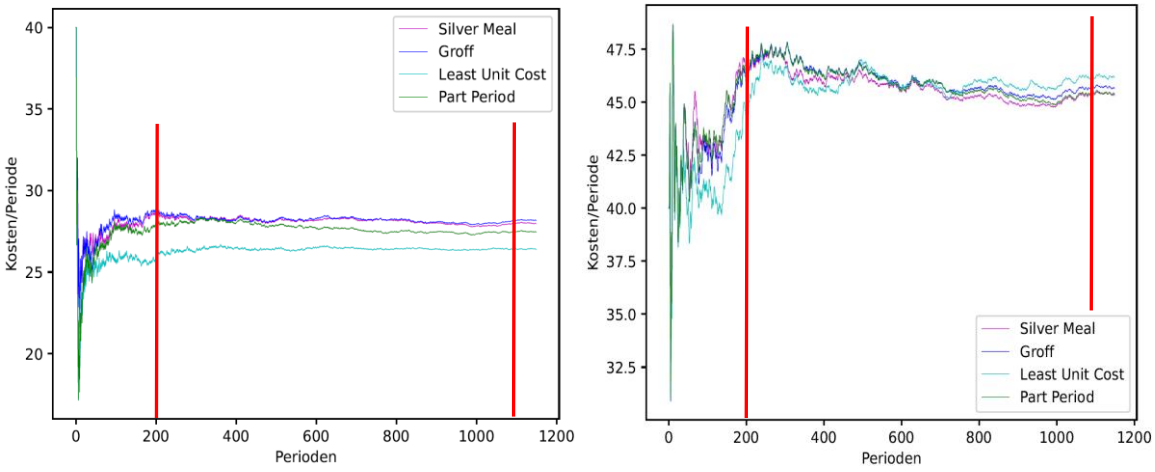

Abbildung 3-9: Kosten pro Periode – Definition von Einschwing- und Ausklangphase

Auf der Y-Achse sind die Kosten pro Periode aufgetragen. Auf der X-Achse die Simulationsdauer in Perioden. Die vier Heuristiken sind farblich gekennzeichnet. Wie in den beiden Diagrammen ersichtlich, befindet sich das System in etwa nach 200 Perioden im eingeschwungenen Zustand. Ab diesem Zeitpunkt weisen die Kosten pro Periode nur noch geringe Abweichung im Zeitverlauf auf. Eine Abklingphase ist in den beiden Diagrammen nicht ersichtlich. Daher wurden die Abklingphase mit 50 Perioden am Simulationsende definiert. Resultierend wurden 900 Perioden nach der Einschwingphase (200 Perioden) zur Analyse betrachtet. Somit ist ein Simulationshorizont mit 1150 Perioden ausreichend dimensioniert.

Die beiden Faktoren Nachfrage und Prognosefehler wurden stochastisch, mittels einer Wahrscheinlichkeitsverteilung modelliert. Aus diesem Grund müssen die beiden Faktoren repliziert werden, sodass belastbare Ergebnisse erzeugt werden. Die Anzahl der Replikationen kann sowohl grafisch wie auch analytisch, mittels Konfidenzintervallen ermittelt werden (Gutenschwager et al. 2017, S. 188).

Ein Konfidenzintervall gibt an, in welchem Bereich ein unbekannter Parameter (z. B. der wahre Mittelwert einer Population) mit einer bestimmten Wahrscheinlichkeit liegt. Je kleiner das Konfidenzintervall, desto präziser ist die Schätzung des zugrunde liegenden Parameters. Ein breites Konfidenzintervall hingegen weist auf eine hohe Unsicherheit der Schätzung hin und liefert entsprechend wenig Aussagekraft über den wahren Wert des Parameters. (Arrenberg 2020, S. 213 f.) Die obere und untere Konfidenzintervallgrenze kann mit nachfolgender Formel (3-11) berechnet werden (Arrenberg 2020, S. 214):

$$KI = \bar{x} \pm z * \frac{\sigma}{\sqrt{n}} \tag{3-11}$$

mit

$\bar{x}$ = Stichproben Mittelwert
σ = Standardabweichung
n = Stichprobenumfang
z = Kritischer Wert aus Standardnormalverteilung entsprechend Konfidenzniveau

Zur Bestimmung der Anzahl von Replikationen ist das Konfidenzintervall möglichst klein zu wählen, damit eine präzise Analyse, auch bei kleineren Abweichungen ermöglicht wird. Daher wird ein 99 %-Konfidenzintervall ausgewählt, d. h. mit 99 % Wahrscheinlichkeit enthält das Intervall den wahren (unbekannten) Wert des Parameters.

Zur Berechnung der Konfidenzintervalle werden die Simulationsexperimente mit gleicher Parametrisierung der Nachfrage zehn Mal repliziert und beim Prognosefehler 15-

mal repliziert, sodass insgesamt 150 Replikationen der 3.888 Faktorkombination simuliert werden.

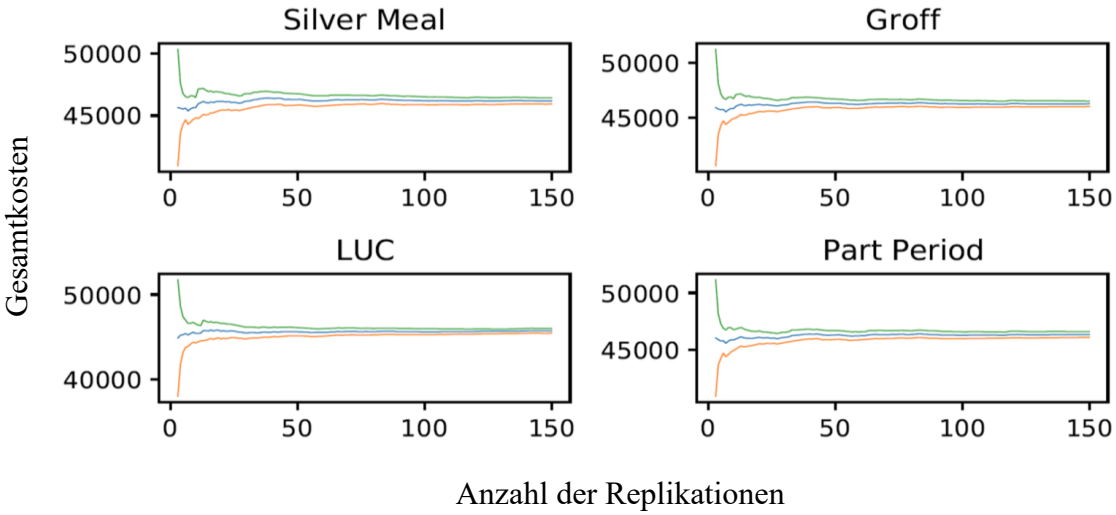

Abbildung 3-10: Obere und untere Konfidenzintervallgrenze - Anzahl Replikationen

In der Abbildung 3-10 sind die Konfidenzintervalle der vier Heuristiken für eine Faktorkombination dargestellt. Dabei sind die Gesamtkosten über die Anzahl der Replikationen aufgetragen. In den einzelnen Diagrammen zeigt die grüne Kurve die obere Intervallgrenze, die rote Kurve die untere Intervallgrenze und die blaue Kurve den Mittelwert. Wie in der Abbildung 3-10 ersichtlich, ist das Konfidenzintervall bereits nach ca. 50 Replikationen ausreichend klein, sodass nach der grafischen Analyse die 150 Replikationen ausreichend sind.

Darüber hinaus werden ebenfalls die Konfidenzintervalle analytisch untersucht. In diesem Rahmen liegen nach 150 Replikationen die obere und untere Konfidenzintervallgrenz bei einem 99 %-Konfidenzintervall über alle Faktorkombinationen durchschnittlich ca. 0,8 % auseinander. Dies bedeutet, dass sich die Gesamtkosten im Intervall von 0,8 % unterscheiden bei 99 % der Replikationen. Folglich kann gesagt werden, dass 150 Replikationen zu präzisen Ergebnissen führt und eine genaue Analyse der Simulationen ermöglicht. Resultierend werden die 3888 Faktorkombinationen mit 150 Replikationen simuliert, sodass 583.200 Simulationsexperimente für die vier Heuristiken durchgeführt werden. Die Auswertung findet anhand der Gesamtkosten sowie anhand der prozentualen Abweichung zur besten Heuristik statt.

3.2.5 Verifikation und Validierung der Modellierung und Simulation

Die Verifikation prüft, ob die Umsetzung eines Modells formal korrekt ist, d. h. ob ein Phasenergebnis korrekt in das nächste überführt wird. Wohingegen die Validierung bewertet, ob das Modell das reale System adäquat abbildet und somit für die gestellte Fragestellung geeignet ist. (Rabe et al. 2008, S. 15)

Die Verifikation und Validierung der Modellierung und Simulation von logistischen Systemen sind eine zentrale Aufgabe und als Querschnitt beziehungsweise als Bestandteil von allen Planungsschritten zusehen. Darüber hinaus sind sie essenzielle Verfahren zur Sicherstellung der Glaubwürdigkeit von Simulationsexperimenten. Ihre Hauptfunktion besteht darin, sicherzustellen, dass die aus Simulationen gewonnenen Ergebnisse zur fundierten Entscheidungsfindung herangezogen werden können, denn fehlerhafte Modelle oder ungeeignete Eingangsdaten können zu Fehlentscheidungen führen. (Gutenschwager et al. 2017, 203 f.)

Ein Simulationsmodell ist glaubwürdig, wenn es vom Nutzer als hinreichend genau und zuverlässig akzeptiert wird, um zur Beantwortung einer konkreten Fragestellung herangezogen zu werden (Carson 2002, S. 53). Die Etablierung dieser Glaubwürdigkeit ist anspruchsvoll, da ein einziges negatives Beispiel ausreichen kann, sie zu untergraben. Eine Vielzahl positiver Tests kann lediglich die Wahrscheinlichkeit der Modellgültigkeit erhöhen, jedoch keinen absoluten Nachweis liefern. Zur Bewertung der Glaubwürdigkeit eines Modells nennen RABE et al. neun zentrale Kriterien, die systematisch geprüft werden sollten (Rabe et al. 2008, S. 22 f.):

- **Vollständigkeit:** Berücksichtigung aller relevanten Elemente der Aufgabenstellung
- **Konsistenz:** Nachvollziehbare und dokumentierte Änderungen im Modell
- **Genauigkeit:** Präzise Formulierung der Zielanforderungen
- **Aktualität:** Übereinstimmung der Modellgrundlagen mit dem aktuellen Stand
- **Eignung:** Passende Datenaggregation zur Zielerfüllung
- **Plausibilität:** Erwartungskonformes Verhalten des Modells in Extremsituationen
- **Verständlichkeit:** Nachvollziehbarkeit für alle Beteiligten
- **Machbarkeit:** Umsetzbarkeit angesichts Systemkomplexität und Fragestellung
- **Verfügbarkeit:** Tatsächliches Vorliegen der erforderlichen Daten

Zur Überprüfung dieser Kriterien kommen unterschiedliche Techniken zum Einsatz. BALCI gibt in seiner Arbeit einen umfassenden Überblick über 77 unterschiedliche Techniken zur Verifikation und Validierung (Balci 1998, S. 335 ff.). Zur Auswahl der richtigen Techniken gibt es kein allgemeingültiges Vorgehen und muss in Abhängigkeit des

jeweiligen Simulationsexperiments erfolgen (Sargent 2010, S. 166). Darüber hinaus beschreiben RABE et al. 20 unterschiedliche Techniken zur Verifizierung und Validierung (Rabe et al. 2008, S. 93 ff.), welche in GUTENSCHWAGER et al (Gutenschwager et al. 2017, S. 209 ff.) teilweise konkretisiert werden. Nachfolgend werden die Techniken erläutert, welche im Rahmen dieser Arbeit durchgeführt wurden.

Eine **Sensitivitätsanalyse** dient der experimentellen Untersuchung der Auswirkungen von Änderungen in den Eingangsdaten auf die resultierenden Ausgangsgrößen (Gutenschwager et al. 2017, S. 210). In diesem Kontext wurden unterschiedliche Faktoren und Faktorstufen variiert und das Verhalten auf die Gesamtkosten analysiert. Es konnte beobachtet werden, dass kleine Änderungen der Faktorstufen, auch zu kleinen Änderungen der Gesamtkosten geführt haben. Diese Beobachtungen entsprechen den Erwartungen, sodass von einer stabilen und robusten Modellantwort gesprochen werden kann. Würde ein konträres Systemverhalten beobachtet, so deutet das auf eine mangelnde Validität des Modells hin, schließt diese jedoch nicht aus (Gutenschwager et al. 2017, S. 210).

Im Rahmen der **Trace-Analyse** werden während der Simulationslaufzeit Informationen über das zeitliche Verhalten von beispielsweise Zwischenergebnissen protokolliert. Diese aufgezeichneten Daten ermöglichen eine zeitliche Nachverfolgung von Zustandsänderungen innerhalb des Modells. (Gutenschwager et al. 2017, S. 211) Die Trace-Analyse wurde ebenso im Rahmen der Simulationen durchgeführt, sodass unterschiedliche Zwischenergebnisse gespeichert und analysiert wurden. Gespeicherte Zwischenergebnisse sind beispielsweise Kosten pro Periode, Bestellmengen oder durchschnittlicher Bestand im System. Die Kosten pro Perioden im Zeitverlauf sind exemplarisch in Abbildung 3-9 für eine Faktorkombination ersichtlich. Die Zwischenergebnisse im Zeitverlauf wurden beispielsweise auf Ausreißer analysiert.

Beim **Extreme-Condition Test,** oder auch Grenzwertanalyse genannt, wird das erwartete Modellverhalten mit dem tatsächlichen Modellverhalten, bei unterschiedlichen Faktoren und Faktorstufen verglichen (Gutenschwager et al. 2017, S. 211). Dieses Grenzverhalten wurde untersucht, in dem sehr große oder sehr kleine Faktorstufen als Simulationseingang verwendet und anschließend die Ergebnisse analysiert wurden.

Die **Validierung im Dialog** beschreibt eine Technik, in der im Dialog anderen Personen das Modell erklärt wird und durch die weiteren Personen auf Plausibilität geprüft (Gutenschwager et al. 2017, S. 211). Im Rahmen der vorliegenden Simulationen wurde das Simulationsmodell auf Plausibilität sowie der Programmiercode auf Richtigkeit geprüft.

Beim **Vergleich mit anderen Modellen** werden vorliegende Ergebnisse aus ähnlichen Modellen mit den aktuellen Ergebnissen verglichen (Gutenschwager et al. 2017, S. 212). Wie im Stand der Forschung (vgl. Kapitel 2.4) ersichtlich, existieren bereits einige Vorarbeiten. Diese Vorarbeiten wurden mit den Ergebnissen der Simulationsexperimente verglichen und auf konträre Ergebnisse analysiert.

Für **Tests von Teilmodellen** muss die Simulation modular aufgebaut sein, sodass einzelne Funktionalitäten der Simulation, mit minimierter Komplexität getestet werden können (Gutenschwager et al. 2017, S. 212). In diesem Kontext wurde die Simulation so implementiert, dass beispielsweise die unterschiedlichen Heuristiken, die Generierung der Prognosefehler, das Planungsvorgehen der Bestellmengen modular implementiert wurden. Diese einzelnen Funktionalitäten wurden anschließend isoliert auf Lauffähigkeit und auf Plausibilität der Ergebnisse überprüft.

Abschließend wurde eine **manuelle Berechnung** von Zwischenergebnissen durchgeführt. Dabei wurden die Bestellmengen für einige Faktorkombinationen mit einer kurzen Simulationsdauer berechnet. Gleichzeitig wurden die Bestellmengen manuell ermittelt und anschließend mit den Ergebnissen von den implementierten Heuristiken verglichen.

Zusammenfassend zeigen die zuvor aufgezeigten Techniken keine auffälligen Ergebnisse, sodass anhand der neun zentralen Kriterien nach RABE et al. die Glaubwürdigkeit der Simulationsexperimente bestätigt werden kann.

3.3 Zwischenfazit: Synthetische Daten zur dynamischen Bestellmengenplanung

Das dritte Kapitel fokussiert die Generierung synthetischer Daten als Grundlage für die kostenorientierte Auswahl dynamischer Bestellmengenverfahren und beantwortet damit die zweite Forschungsfrage: „*Wie lassen sich synthetische Daten zur dynamischen Bestellmengenplanung erzeugen?*“. Zur Gewährleistung von Reproduzierbarkeit, Generalisierbarkeit und Abbildungsgüte müssen umfangreiche und vielfältige Daten erhoben werden. Eine derart große und übertragbare Datenbasis lässt sich aus realen Systemen jedoch nur schwer gewinnen. Um diesen Anforderungen gerecht zu werden, werden synthetische Daten mittels Simulationsexperimenten erzeugt.

Die theoretische Verankerung der dritten Forschungsfrage bildet die Systemtheorie. Systeme werden als strukturierte Einheiten verstanden, deren Elemente über Beziehungen interagieren. Für die Simulationsexperimente werden die System- und Modelltheorie miteinander verknüpft und in Form von Systemmodellen dargestellt. Der Aufbau und die Durchführung der Simulationsexperimente folgen dem strukturierten Vorgehensmodell nach RABE et al. (Rabe et al. 2008, S. 4 ff.).

Auf Basis des Systemgedankens sowie der strukturierten Vorgehensweise, wird zunächst das Konzeptmodell vorgestellt, welches das System und seine Elemente beschreibt. Darauf aufbauend erfolgt die Spezifikation des formalen Modells, das den Aufbau und Ablauf der Simulation im Detail beschreibt. Ferner werden die einzelnen Faktoren und Faktorstufen modelliert. Die Faktoren Nachfrage und Prognosefehler werden stochastisch mittels geeigneter Wahrscheinlichkeitsverteilungen abgebildet. Die Nachfrageverläufe werden nach SYNTETOS et al. in vier Typen klassifiziert gleichmäßig, erratisch, sporadisch und klumpig und durch eine geeignete Parametrisierung der Gamma- und Binomialverteilung modelliert. Die Prognosefehler werden differenziert nach systematischen Überschätzungen, Unterschätzungen und Rauschen modelliert. Weitere deterministische Faktoren sind der Servicegrad, die Time Between Order (TBO) als Verhältnis von Bestell- zu Lagerhaltungskosten sowie die Wiederbeschaffungszeit.

Der Experimentplan umfasst 3.888 Faktorkombinationen für die vier Heuristiken. Die Simulationen werden in Python implementiert. Die Länge des Simulationshorizonts (1.150 Perioden) sowie Einschwing- und Abklingphasen werden analytisch und grafisch festgelegt. Die erforderliche Replikationsanzahl wird über die Auswertung von Konfidenzintervallen (99 %) bestimmt, wobei 150 Replikationen als ausreichend präzise identifiziert werden. Dies gewährleistet robuste Simulationsergebnisse. Aus den 3.888 Faktorkombinationen und 150 Replikationen ergeben sich insgesamt 583.200 Simulationsexperimente. Abschließend werden die Modellierung und Simulation mithilfe unterschiedlicher Methoden verifiziert und validiert, um die Zuverlässigkeit der Ergebnisse sicherzustellen.

4 Entwicklung von Entscheidungsregeln zur kostenorientierten Auswahl dynamischer Bestellmengenverfahren

In diesem Kapitel wird die Entwicklung von Entscheidungsregeln zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren fokussiert, damit steht die dritte Forschungsfrage im Mittelpunkt: „*Welche Entscheidungsregeln können zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren definiert werden?*“.

Zunächst werden die synthetischen Daten mittels statistischer Methoden analysiert. Bei der Datenanalyse werden die grundlegenden Wirkzusammenhänge und Eigenschaften der Daten untersucht. Anschließend wird der Einfluss der einzelnen Faktoren und Faktorstufen auf die kostenorientierte Auswahl analysiert und beschrieben. Abschließend werden auf Basis der zuvor gewonnenen Erkenntnisse Entscheidungsregeln abgeleitet, welche eine kostenorientierte Auswahl von dynamischen Bestellmengenverfahren ermöglichen.

Die theoretische Verankerung der dritten Forschungsfrage stellt das **No-Free-Lunch Theorem** (NFL-Theorem) dar. Dabei beschreibt ein Theorem einen Zusammenhang, welcher aus den grundlegenden Axiomen einer Theorie abgeleitet wird. Das NFL-Theorem wurde erstmals formalisiert von David Wolpert und William G. Macready in den 1990er Jahren im Kontext von Optimierungsalgorithmen (Wolpert und Macready 1997, S. 67 ff.). Im Kern besagt das Theorem, dass kein Optimierungsalgorithmus im Durchschnitt besser als ein anderer ist, wenn über alle möglichen Funktionen gemittelt wird. Das bedeutet, dass ein Verfahren, welches für eine bestimmte Problemklasse besonders gut funktioniert, zwangsläufig für andere Probleme schlechter abschneidet. (Ho und Pepyne 2002, S. 549 ff.; Yang et al. 2020, S. 102 f.)

Dieses Theorem trifft auch auf die Auswahl von dynamischen Bestellmengenverfahren zu, sodass unterschiedliche Planungsverfahren für das gleiche Problem zur Verfügung stehen und je nach Faktorkombination die richtige Planungsalternative ausgewählt werden muss. Vergangene Forschungsarbeiten in diesem Bereich zeigten bereits, dass kein Verfahren das beste für alle möglichen Faktorkombinationen ist. Demzufolge muss je nach Faktorkombination das kostengünstigste Verfahren ausgewählt werden.

4.1 Grundlegende Wirkzusammenhänge der dynamischen Bestellmengenplanung

Zunächst werden in der Tabelle 4-1 die Merkmale der synthetischen Daten dargestellt, welche die Grundlage für die Auswahl der statistischen Methoden und die Durchführung der Datenanalyse sind.

Tabelle 4-1: Merkmale der synthetischen Daten

Merkmal	**Skalentyp**	**Beschreibung**
Streuung	Verhältnis	Variationskoeffizient der Gammaverteilung
NNPerioden	Verhältnis	Prozentualer Anteil der Perioden mit Bedarf
Servicegrad	Verhältnis	Maß für die Liefer- oder Leistungsbereitschaft
Fehlertyp	Nominal	Art des Prognosefehlers
Fehlerstärke	Verhältnis	Stärke des Fehlers
TBO	Verhältnis	Time Between Order in Perioden
WBZ	Verhältnis	Wiederbeschaffungszeit in Perioden
Silver Meal	Verhältnis	Gesamtkosten bei Anwendung von SM
Groff	Verhältnis	Gesamtkosten bei Anwendung von Groff
Least Unit Cost	Verhältnis	Gesamtkosten bei Anwendung von LUC
Part Period	Verhältnis	Gesamtkosten bei Anwendung von PP

Um den Kostenvergleich der eingesetzten Heuristiken in den Vordergrund zu stellen und nicht deren absolute Kostenwerte, wurde für jede Heuristik die prozentuale Abweichung zur besten Heuristik innerhalb des jeweiligen Simulationslaufes bzw. Datensatzes berechnet. Die Heuristik mit den geringsten Kosten erhält einen Abweichungswert von 0 %, während für alle anderen Heuristiken die Abweichung gemäß Gleichung (4-1) beispielhaft anhand der Planungsheuristik Part Period (PP) ermittelt wird.

$$PP = \frac{absPP}{absBestHeuristic} - 1 \tag{4-1}$$

mit

$absPP = Gesamtkosten\ Part\ Period$

$absBestHeuristic = Planungsheuristik\ mit\ den\ geringsten\ Gesamtkosten$

Nachfolgend werden mehrere Abbildungen und Diagramme zur statistischen Analyse auf Basis der synthetischen Daten durchgeführt. Diese Abbildungen wurden mithilfe der Python-Bibliotheken pandas (The pandas development team 2025), matplotlib (Thomas A Caswell et al. 2023) und seaborn (Waskom 2021) erstellt.

Die Häufigkeitsverteilung der jeweils kostenoptimalen Heuristik über alle 583.200 Simulationsergebnisse hinweg ist in der Abbildung 4-1 dargestellt.

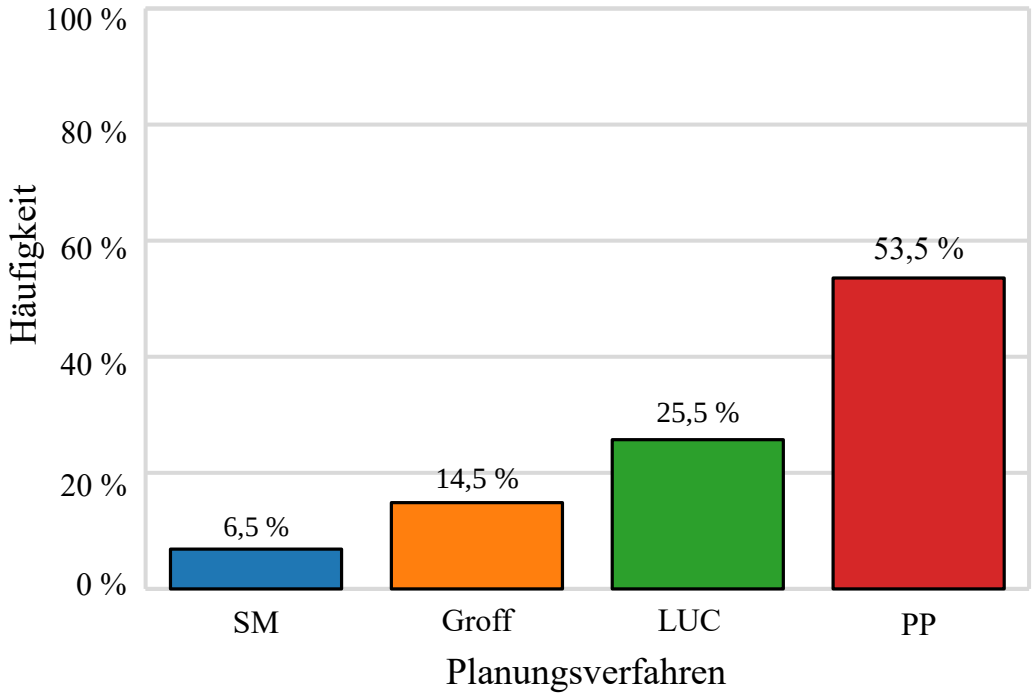

Abbildung 4-1: Häufigkeit der kostenoptimalen Planung nach Planungsheuristik

Mit ca. 53,5 % weist die Heuristik Part Period (PP) am häufigsten die geringsten Gesamtkosten auf. An zweiter Stelle folgt die Least Unit Cost Heuristik (LUC) mit einer Häufigkeit von ca. 25,5 %. Wohingegen die beiden Heuristiken Silver Meal (SM) (ca. 6,5 %) und Groff (ca. 14,5 %) kumuliert nur in ca. 21 % der Fälle die geringsten Gesamtkosten aufweisen. Die Abbildung 4-1 zeigt zwar eine ungleiche Verteilung der Heuristiken, wobei PP dominiert. Jedoch wird auch ersichtlich, dass keine Heuristik über alle Fälle hinweg die beste ist.

Neben der Häufigkeitsverteilung in Abbildung 4-1 werden im nächsten Schritt die Kostenunterschiede und das miteinhergehende Kosteneinsparungspotenzial analysiert. Dazu werden in Abbildung 4-2 die Abweichungen der jeweiligen Heuristiken zur kostenoptimalen Heuristik analysiert. Zur Visualisierung der Abweichungen wurde ein Box-Whisker-Plot verwendet (vgl. Abbildung 4-2). Neben dem Median (horizontale Linie) wurde zusätzlich das arithmetische Mittel durch einen weißen Punkt in der Abbildung markiert.

Aufgrund teilweise stark ausgeprägter Extremwerte wurden diese nicht explizit in den Boxplots visualisiert, um die Interpretation der wesentlichen Datenbereiche zu erleichtern. Um dennoch einen möglichen Informationsverlust zu minimieren, wurde für jede Heuristik das 99,5%-Perzentil dargestellt, während die maximale Abweichung separat in Textform angegeben wurde.

Das 99,5%-Perzentil beschreibt den Wert in einer Datenverteilung, unterhalb dessen 99,5 % aller gemessenen Werte liegen bzw. dass 0,5 % der Werte größer sind als das 99,5%-Perzentil. Die Nutzung dieses Perzentils erfolgt typischerweise, um Extremwerte oder Ausreißer in großen Datensätzen zu analysieren, ohne einzelne extrem hohe oder niedrige Werte explizit darzustellen. Es erlaubt somit eine robustere Interpretation, indem es deutlich macht, wie sich die überwiegende Mehrheit der Daten verhält und gleichzeitig sehr hohe Abweichungen nach oben abschneidet.

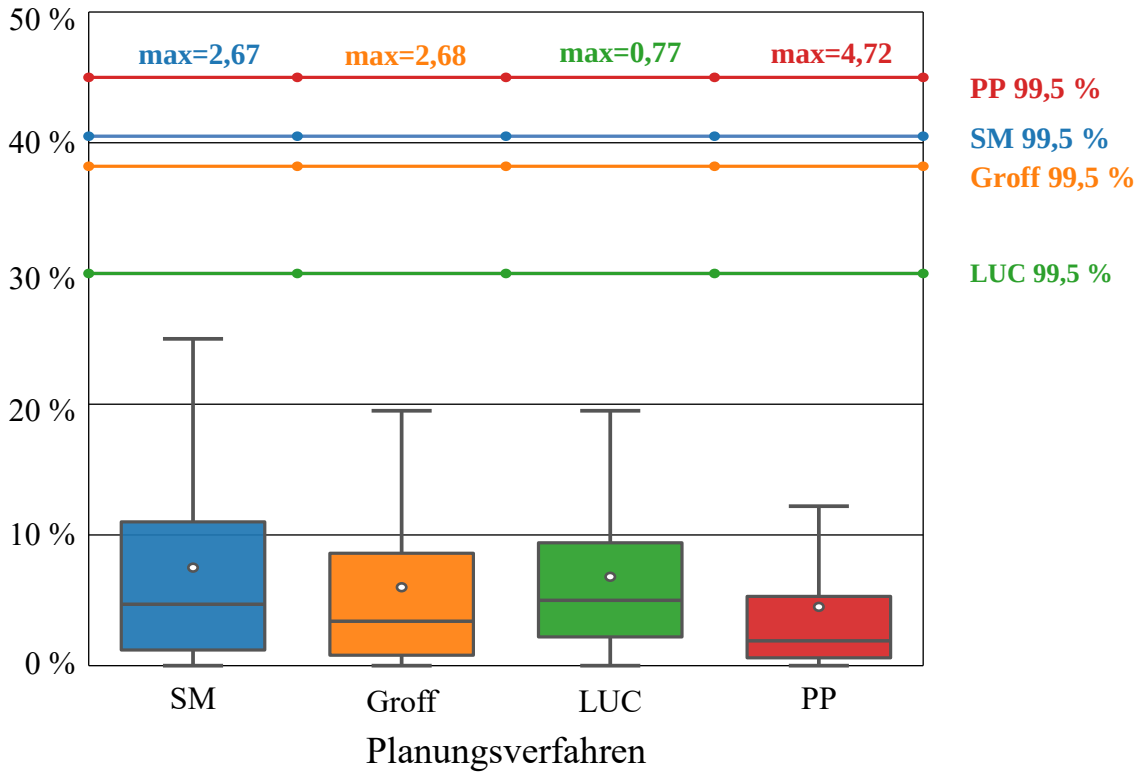

Abbildung 4-2: Abweichung zur kostenoptimalen Heuristik

Die Analyse der dargestellten Box-Plots erlaubt eine Interpretation hinsichtlich der durchschnittlichen Kostenabweichung und Robustheit der vier Heuristiken. Die PP-Heuristik zeigt die geringste Medianabweichung (ca. 1,5 %) sowie den engsten Interquartilsabstand (IQR) und weist damit im Durchschnitt die geringsten Kostenabweichungen zur optimalen Heuristik auf. Im Gegensatz dazu weist die SM-Heuristik sowohl den höchsten Median als auch die größte Varianz auf. Ihre Resultate sind demnach deutlich volatiler. Die

Heuristiken Groff und LUC nehmen eine Mittelstellung ein. Beide zeigen moderate Medianwerte im Bereich von etwa drei bis vier Prozent bei ähnlich ausgeprägter Streuung.

In allen Box-Plots liegt der arithmetische Mittelwert oberhalb des jeweiligen Medians, was auf eine rechtsschiefe Verteilung hinweist. Dies deutet darauf hin, dass die Mehrheit der beobachteten Abweichungen vergleichsweise gering ausfällt, während einzelne Ausreißer den Mittelwert nach oben verschieben.

Besonders auffällig sind die Ausreißer und Grenzwerte. Die zusätzlich dargestellten maximalen Abweichungen (zwischen 0,77 und 4,72) zeigen, dass es bei allen Verfahren zu hohen Fehlplanungen kommen kann. Trotz der insgesamt geringen durchschnittlichen Kostenabweichung vom PP-Verfahren ist die maximale Abweichung mit 4,72 (472 %) die höchste aller vier Heuristiken. Wohingegen die Heuristik LUC ein robustes Verfahren mit geringen Extremwerten (77 %) darstellt.

Aus den Verteilungen und Kostenabweichungen lassen sich bereits erste Erkenntnisse ableiten. Sollten keine weiteren Informationen über die gegebenen Parameter vorliegen, ist in den meisten Fällen die PP-Heuristik zu empfehlen. Dies begründet sich in der hohen Anzahl an kostenoptimalen Ergebnissen sowie der meist geringen Kostenabweichung. Auf Basis der hohen Extremwerte der PP-Heuristik kann die Auswahl der LUC-Heuristik, als ein robustes Verfahren mit etwas höheren durchschnittlichen Kosten, ebenfalls sinnvoll sein.

Im Allgemeinen bestätigt sich das NFL-Theorem, dass es keine Heuristik gibt, die in allen Fällen am besten geeignet ist. Das PP-Verfahren zeigt die besten Ergebnisse auf, jedoch auch die größten Extremwerte in Bezug zur Kostenabweichung, wodurch nochmals die Notwendigkeit einer kostenorientierten Auswahl von dynamischen Bestellmengenverfahren bestätigt wird.

4.2 Einfluss der Systemfaktoren auf die kostenorientierte Auswahl von dynamischen Bestellmengenverfahren

Im folgenden Abschnitt wird der Einfluss der einzelnen Systemfaktoren und Faktorstufen auf die kostenorientierte Heuristikwahl untersucht. Dafür wird zum einen die Stärke des Einflusses auf die Heuristikwahl, als auch das Verhalten der Heuristiken unter den verschiedenen Faktorstufen untersucht.

Für solche Fragestellungen bzw. Analysen werden üblicherweise Korrelationsanalysen eingesetzt, sodass der Zusammenhang zwischen zwei Merkmalen untersucht werden kann. Die Auswahl der richtigen Korrelationsanalyse / -koeffizienten ist abhängig von

den Skalenniveaus (Wirtz und Nachtigall 2012, S. 170). In diesem Kontext gibt die nachfolgende Tabelle 4-2 einen Überblick über die unterschiedlichen Skalenniveaus und die entsprechenden Korrelationskoeffizienten.

Tabelle 4-2: Klassifizierung von Korrelationskoeffizienten

		intervall-skaliert	**ordinal-skaliert**	**nominalskaliert**		
				künstlich dichotom	**natürlich dichotom**	**polytom**
	intervall-skaliert	Produkt-Moment-Korrelation	Spearmans Rangkorrelation Kendal τ polychorische Korrelation*	punktbiseriale Korrelation biseriale Korrelation	punktbiseriale Korrelation	η-Koeffizient
	ordinal-skaliert		Spearmans Rangkorrelation Kendals τ polychorische Korrelation*	biseriale Rangkorrelation polychorische Korrelation*	biseriale Rangkorrelation	Cramér´s Index
nominalskaliert	**künstlich dichotom**			Punkttetrachorische Korrelation (φ-Koeffizient) Tetrachorische Korrelation*	Punkttetrachorische Korrelation (φ-Koeffizient) v-Koeffizient*	Cramér´s Index
	natürlich dichotom				Punkttetrachorische Korrelation (φ-Koeffizient) Yules Y	Cramér´s Index
	polytom					Cramér´s Index Kontingenz-koeffizient CC

Zunächst wird der Einfluss der Systemfaktoren auf die Heuristikwahl untersucht. Für die Analyse wird die Variable kostenoptimale Heuristik eingeführt. Wobei die kostenoptimale Heuristik diejenige Heuristik ist, die in einem Simulationslauf unter gleicher Faktorkombination die geringsten Gesamtkosten aufweist. Die Systemfaktoren (polytom) sowie die Variable kostenoptimale Heuristik (dichotom) sind nominalskaliert. Wobei eine dichotome Variable zwei Ausprägungen besitzt und eine polytome Variable mehr als zwei Ausprägungen hat. Auf Basis der Tabelle 4-2 und den beschriebenen Skalaniveaus wurde der Korrelationskoeffizient Cramér's Index ausgewählt.

Beim **Cramér's Index**, oder auch Cramér's V genannt, handelt es sich um eine Erweiterung des Phi-Koeffizienten. Cramér's Index liefert hierbei ein Maß der Assoziation zwischen zwei nominalen Variablen im Wertebereich zwischen 0 und 1. Dabei deutet eine 0 darauf hin, dass keine Assoziation zwischen den beiden Variablen besteht, während ein Wert von 1 auf eine vollständige Assoziation hindeutet. (Sheskin 2020, S. 678)

Die Analyse wurde für alle Systemfaktoren, ohne Berücksichtigung der Stärke des Prognosefehlers, durchgeführt. Die Fehlerstärke wird aus der ersten allgemeinen Analyse entfernt, da die Stärke des Prognosefehlers vom jeweiligen Fehlertyp abhängt und anders skaliert ist als der Fehlertyp. Für den Einfluss der Fehlerstärke werden im weiteren Verlauf separate Analysen durchgeführt. Zur Interpretation der Ergebnisse von Cramér's V wird w (Cohen's w) verwendet (Cohen 2013, S. 216 ff.). Dieser bestimmt sich aus dem Wert ϕv von Cramér's V und dem Minimum von der Anzahl an Zeilen und Spalten der Kontingenztabelle k durch folgende Formel (Cohen 2013, S. 223; Sheskin 2020, S. 679):

$$w = \phi_v \sqrt{k-1} \tag{4-2}$$

Umso größer der w-index, umso größer ist der Zusammenhang zwischen den Variablen (Cohen 2013, S. 227). Wie in Abbildung 4-3 ersichtlich, besitzt der Prognosefehlertyp einen großen Einfluss auf die kostenorientierte Heuristikwahl. Weiterhin besitzt die TBO den zweitgrößten Einfluss. Wohingegen die übrigen Systemfaktoren einen kleinen oder sehr kleinen Einfluss haben auf die Auswahl.

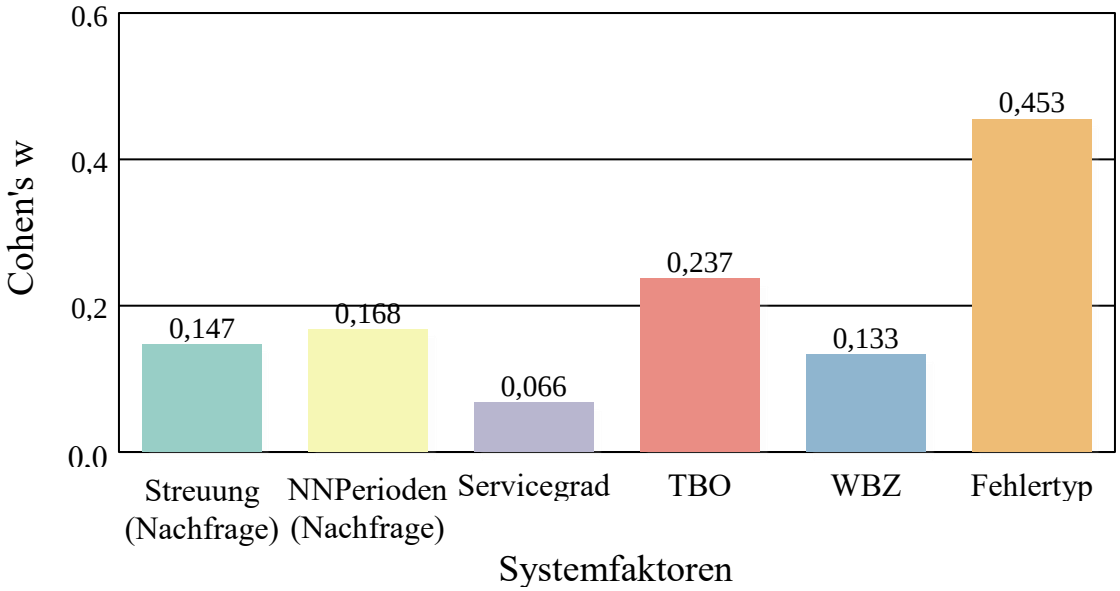

Abbildung 4-3: Einfluss der Systemfaktoren auf die kostenorientierte Heuristikwahl mittels Cohen's w

Da der Fehlertyp des Prognosefehlers den größten Einfluss auf die Auswahl der Heuristiken hat, wurde der Einfluss der einzelnen Faktorstufen für die jeweiligen Fehlertypen tiefergehend untersucht. Hierzu wurde Cohen's w nach Fehlertyp gruppiert analysiert, sodass die Stärke des Prognosefehlers (Faktorstufen) gezielt untersucht werden kann. Die Ergebnisse sind in Abbildung 4-4 dargestellt.

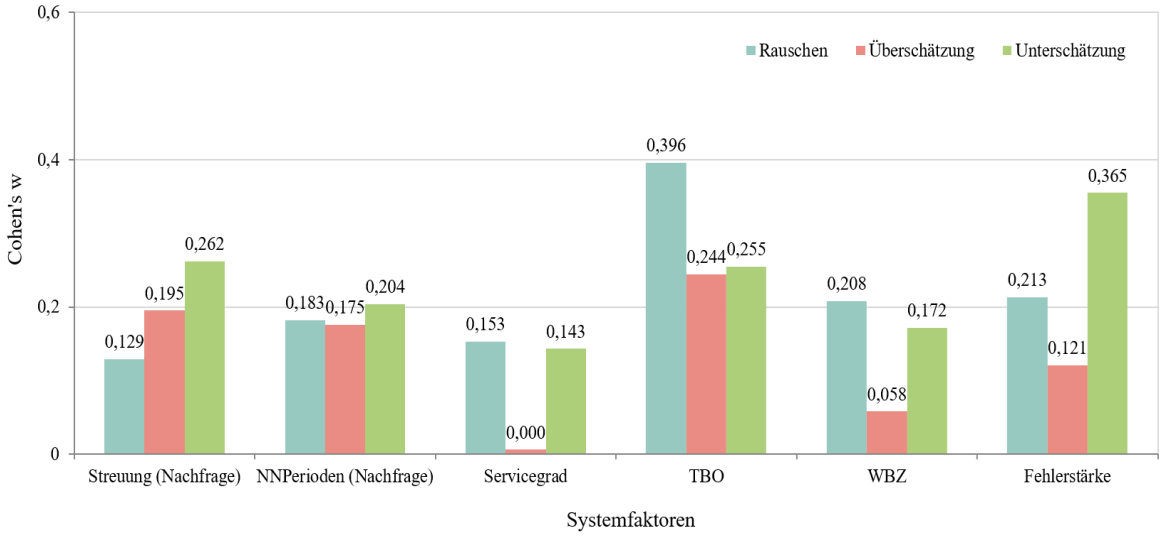

Abbildung 4-4: Einfluss der Systemfaktoren auf die Heuristikwahl gemessen an Cohen's w – unterteilt nach den Fehlertypen

Bei einem vorliegenden Fehlertyp **Rauschen** fällt der große Einfluss des Faktors TBO (0,396) besonders auf. Die Stärke des Rauschens (0,213) hat nach der TBO den zweitgrößten Einfluss auf die Heuristikwahl, gefolgt von der WBZ (0,208).

Im Fall einer systematischen **Überschätzung** treten keine Fehlmengen auf. Folglich müssen keine Sicherheitsbestände in Abhängigkeit vom Servicegrad vorgehalten werden. Dieser Effekt wird auch durch die Korrelationsanalyse bestätigt, sodass der Servicegrad keinen messbaren Einfluss auf die Heuristikwahl hat (0 statt 0,066). Insgesamt weist auch hier die TBO (0,244) den größten Einfluss auf, gefolgt von der Streuung (0,195) und der NNPerioden (0,175). Die WBZ sowie die Höhe des Prognosefehlers haben dagegen nur einen sehr geringen Einfluss auf die Heuristikwahl.

Im Falle einer **Unterschätzung** hat die Stärke der Unterschätzung (0,365) den größten Einfluss auf die Heuristikwahl. Die Streuung (0,262) und die TBO (0,255) weisen einen ähnlich starken Einfluss auf. Insgesamt wird deutlich, dass im Fall einer Unterschätzung höhere Korrelationswerte auftreten als bei den anderen Fehlertypen.

Demzufolge besitzt der Fehlertyp selbst den größten Einfluss auf die kostenorientierte Auswahl der Bestellmengenverfahren, gefolgt von der TBO, während die übrigen Parameter einen geringen oder sehr geringen Einfluss aufweisen. Auch bei der isolierten Betrachtung der einzelnen Fehlertypen zeigt sich, dass die Höhe des Prognosefehlers und

die TBO den größten Einfluss besitzen. Zusammenfassend verdeutlicht die Korrelationsanalyse, dass die Heuristikwahl maßgeblich von den vorliegenden Systemfaktoren und Faktorstufen abhängt und bei der Auswahl auf den jeweiligen Prognosefehler sowie die TBO geachtet werden muss.

4.3 Fehlertyp-spezifischer Einfluss der Systemfaktoren auf die kostenorientierte Auswahl von dynamischen Bestellmengenverfahren

Im folgenden Kapitel wird der Einfluss der einzelnen Systemfaktoren auf die kostenorientierte Auswahl der Heuristiken tiefergehend analysiert. Dabei erfolgt eine differenzierte Analyse der drei systematischen Fehlertypen Rauschen, Überschätzung und Unterschätzung.

Für diese Untersuchung wird zunächst die **Dummy-Variable** der besten Heuristik eingeführt. Bei der Umwandlung in Dummy-Variablen wird für jede Heuristik eine neue binäre Variable erstellt. Ist die zugehörige Heuristik die kostenoptimale für die verwendete Faktorkombination, wird dies mit einer 1 codiert, im Gegenfall mit einer 0 (Bishop 2009, S. 74 f.).

Darüber hinaus wird ein **Schwellenwert** definiert. Eine Heuristik wird als kostenoptimal betrachtet, wenn ihre Abweichung von der kostenoptimalen Heuristik kleiner ist als der definierte Schwellenwert. Somit wird jede Heuristik, deren Abweichung unter dem Schwellenwert liegt, ebenfalls mit einer 1 codiert. Durch die Einführung des Schwellenwerts werden Entscheidungssituationen vereinfacht und der Fokus liegt darauf, eine Heuristik auszuwählen, die die niedrigsten oder nahezu niedrigsten Kosten erzielt, anstatt eine Heuristik zu wählen, die deutlich schlechter abschneidet als andere. Dieser Ansatz erhöht die Robustheit der Analysen und sorgt für eine bessere Übertragbarkeit der Ergebnisse, da kleinere Unterschiede zwischen den Heuristiken die Gesamtergebnisse weniger beeinflussen.

Bei der Auswahl des Schwellenwerts muss zwischen der Robustheit der Ergebnisse und dem Informationsverlust bei einem großen Schwellenwert abgewogen werden. Für die Auswahl des Schwellenwerts wurde der Anteil der Simulationsläufe in Abhängigkeit vom Schwellenwert grafisch dargestellt (vgl. Abbildung 4-5).

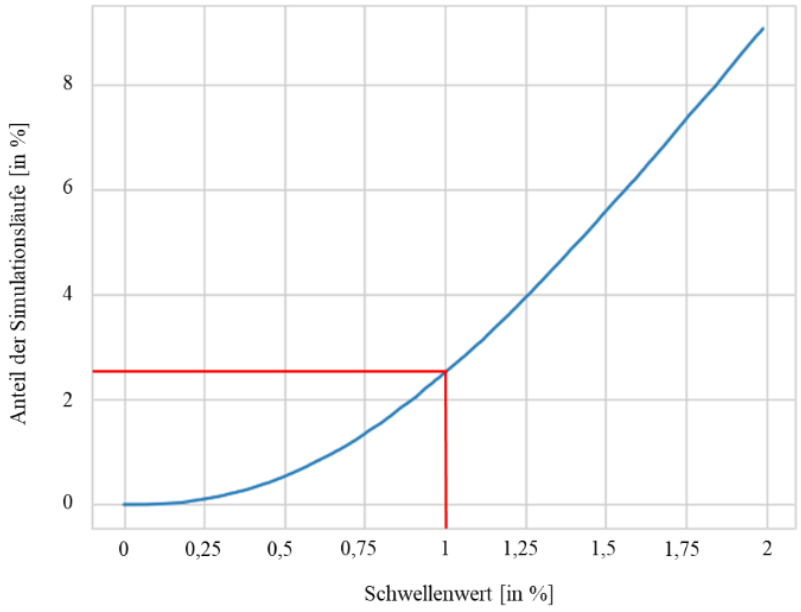

Abbildung 4-5: Anteil der Simulationsläufe, bei denen die relative Kostenabweichung aller vier Heuristiken unterhalb des Schwellwerts liegt

Wie in Abbildung 4-5 ersichtlich, weisen bei ca. 2,5 % aller Simulationsläufe alle vier Heuristiken kostenoptimale Ergebnisse auf, wenn der Schwellenwert bei 1 % liegt (vgl. Abbildung 4-5 rote Markierung). Durch die zunehmende Steigung des Kurvenverlaufs würde eine Erhöhung des Schwellenwerts auf 2 % etwa 9 % der Simulationsläufe betreffen. Wodurch ein hoher Informationsverlust für die kostenorientierte Heuristikwahl mit einhergeht. Demzufolge wird ein Schwellenwert auf 1 % festgelegt.

Für die vertiefende Untersuchung des Einflusses der Systemfaktoren auf die kostenorientierte Auswahl der Heuristiken wurden drei unterschiedliche Analysemethoden ausgewählt:

- *Stapeldiagramme* als visuelle Darstellungsform
- *Korrelationsanalyse* als Messung des Einflusses
- *Boxplots* zur Erfassung der Kostenabweichungen

Die visuelle Aufbereitung zeigt, unter welcher Faktorkombination welche Heuristik kostenoptimal ist, während die Korrelationsanalyse die Stärke des Einflusses auf die kostenorientierte Auswahl untersucht. Abschließend verdeutlichen die Boxplots die tatsächlichen Kostenabweichungen und das damit einhergehende Kosteneinsparungspotenzial. Zusammenfassend bietet die Kombination aus visueller Darstellung, Korrelationsanalyse und Boxplot-Auswertung einen Überblick aus unterschiedlichen Perspektiven über die

Wirkzusammenhänge der untersuchten Einflussfaktoren und ermöglicht fundierte Entscheidungen zur kostenorientierten Heuristikwahl.

Wie bereits zuvor aufgezeigt (vgl. Tabelle 4-2) ist für die Auswahl der richtigen Korrelationsanalyse die Skalierung der Variablen zentral. In diesem Zusammenhang können die Systemfaktoren aufgrund der separaten Analyse nach den drei systematischen Prognosefehlern (Fehlertypen) intervallskaliert dargestellt werden. Auf der anderen Seite lässt sich die Zielvariable kostenoptimale Heuristik durch die Binärkodierung dichotom abbilden.

Daraus ergibt sich, dass für die Korrelationsanalyse die **Point-Biserial-Korrelation** verwendet werden kann (vgl. Tabelle 4-2). Bei der Point-Biserial-Korrelation handelt es sich um einen Sonderfall des Pearson-Korrelationskoeffizienten, der zur Analyse der Korrelation zwischen numerischen und binären Variablen eingesetzt wird. Wie der Pearson-Korrelationskoeffizient nimmt auch die Point-Biserial-Korrelation Werte im Bereich von -1 bis 1 an, wobei Werte nahe -1 bzw. 1 auf eine starke negative bzw. starke positive Korrelation hinweisen und Werte nahe 0 auf eine schwache lineare Korrelation deuten (Sheskin 2020, S. 1324; Cohen et al. 2003, S. 29).

Die Verwendung der Point-Biserial-Korrelation ermöglicht es somit, sowohl die Stärke als auch die Richtung der Korrelation zwischen den Faktoren und Faktorstufen sowie der Wahrscheinlichkeit, dass die betrachtete Heuristik die kostenoptimale ist, zu erfassen. Nachfolgend werden die Faktoren und Faktorstufen tiefergehend analysiert. Dabei werden die Faktoren isoliert nach den drei Fehlertypen untersucht.

4.3.1 Fehlertyp – Rauschen: Einfluss der Systemfaktoren auf die kostenorientierte Auswahl

Im nachfolgenden Kapitel wird der Einfluss der Faktoren und Faktorstufen bei einem vorliegenden Fehlertyp **Rauschen** auf die kostenorientierte Auswahl der Heuristiken untersucht.

Zur ersten Orientierung zeigt Abbildung 4-6 (linke Seite), wie häufig die jeweilige Heuristik bei einem vorliegenden Fehlertyp Rauschen die kostengünstigste Lösung darstellt. Die Häufigkeitsverteilung ähnelt dabei derjenigen der Gesamtpopulation (vgl. Abbildung 4-1). Am häufigsten führt die Planungsheuristik PP (52,3 %) zu den geringsten Kosten, gefolgt von der Heuristik LUC (29,7 %). Die Verfahren SM (6,3 %) und Groff (11,7 %) weisen dagegen deutlich schlechtere Ergebnisse auf.

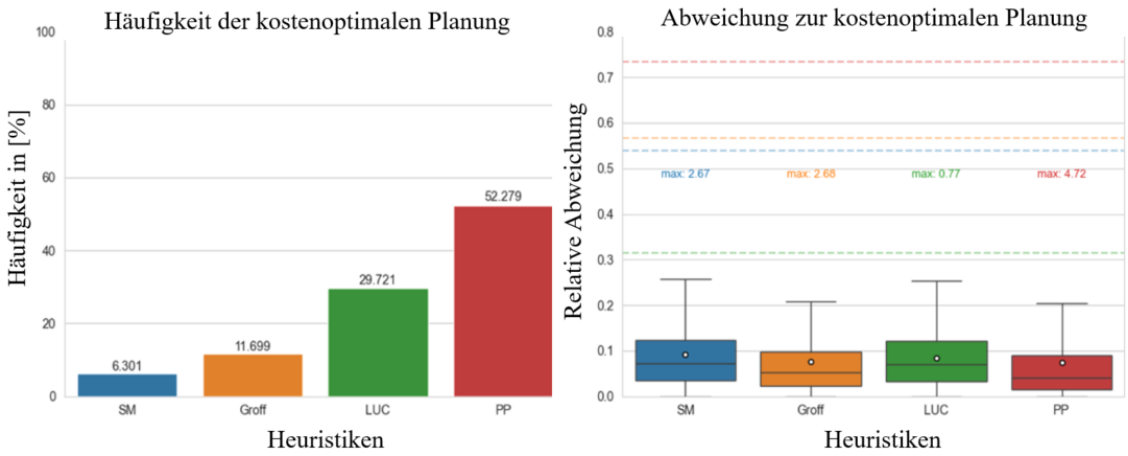

Abbildung 4-6: Rauschen - Häufigkeit der kostenoptimalen Planung nach Heuristik und durchschnittliche Kostenabweichung zur kostenoptimalen Planung

Auf der rechten Seite der Abbildung 4-6 ist die Abweichung zur kostenoptimalen Heuristik als Box-Whisker-Plot dargestellt. Wie bereits bei der Analyse der Gesamtpopulation ist neben dem Median (horizontale Linie) zusätzlich das arithmetische Mittel durch einen weißen Punkt in der Grafik markiert. Aufgrund teilweise stark ausgeprägter Extremwerte wurden diese nicht explizit in den Boxplots visualisiert, um die Interpretation der wesentlichen Datenbereiche zu erleichtern.

Auch hier zeigt sich ein ähnliches Bild wie bei der Gesamtpopulation. Die PP-Heuristik weist mit durchschnittlich die geringste Kostenabweichung auf, jedoch zugleich die höchste maximale Kostenabweichung von rund 472 %. Die durchschnittlichen Kostenabweichungen der drei Heuristiken Groff, LUC und SM liegen dicht beieinander. Die geringste maximale Kostenabweichung zeigt das LUC-Verfahren, was auf eine höhere Robustheit dieses Verfahrens schließen lässt.

Wird die Häufigkeit der kostenoptimalen Planung in Abhängigkeit der unterschiedlichen Systemfaktoren mittels **Stapeldiagrammen** dargestellt, so zeigt sich, dass die PP-Heuristik (rot) dominiert, gefolgt vom LUC-Verfahren (grün) (vgl. Abbildung 4-7). Auffällig ist, dass bei den Verfahren PP und LUC ein gegenläufiges Verhalten zu beobachten ist. Mit zunehmender bzw. abnehmender Faktorstufe erfolgt ein Wechsel des kostenoptimalen Verfahrens von PP zu LUC beziehungsweise von LUC zu PP. Die Verfahren SM (blau) und Groff (orange) sind dagegen nur in wenigen und spezifischen Faktorkombinationen kostenoptimal. Darüber hinaus zeigen die beiden Systemfaktoren Prognosefehler und TBO den stärksten Einfluss auf die kostenoptimale Auswahl, was sich in einem deutlichen Wechsel der kostenoptimalen Heuristik bei Änderung der Faktorstufen widerspiegelt.

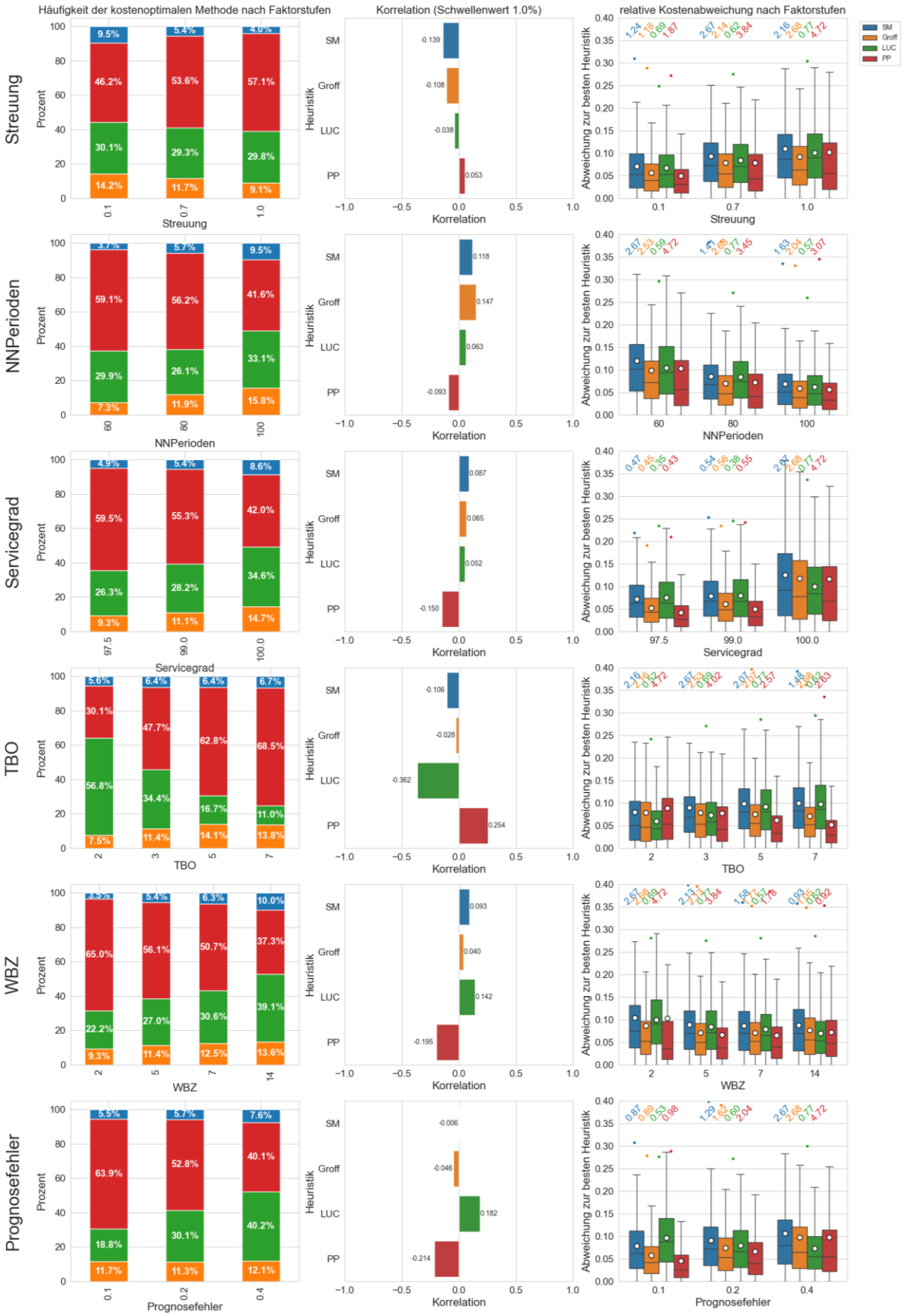

Abbildung 4-7: Rauschen - Wirkzusammenhang der Systemfaktoren auf die kostenorientierte Auswahl der Heuristiken

In der zweiten Spalte von Abbildung 4-7 sind die Ergebnisse der **punktbiserialen Korrelationsanalyse** dargestellt. Die größte Sensitivität der kostenoptimalen Auswahl zeigt sich bei den Faktoren Prognosefehler und TBO. Änderungen der Faktorstufen führen hier zu deutlichen Wechseln in der Auswahl des jeweils optimalen Verfahrens. Dies wird durch die Ergebnisse der **Korrelationsanalyse** bestätigt. Die Faktoren TBO, Prognosefehler und WBZ üben den stärksten Einfluss auf die Heuristikwahl aus, wobei sich für PP und LUC jeweils gegenläufige Korrelationen zeigen. Niedrige TBO-Werte begünstigen LUC, während hohe TBO-Werte PP fördern. Beim Prognosefehler und WBZ wirken sich hohe Werte positiv auf LUC und niedrige Werte positiv auf PP aus. Für SM und Groff besteht hingegen nur ein schwacher Zusammenhang mit den untersuchten Faktoren.

Die Auswertung der **Boxplots** verdeutlicht, dass die Wahl des geeigneten Verfahrens erhebliche Auswirkungen auf die entstehenden Kosten hat. Die mediane Kostenabweichung zwischen den Verfahren liegt in der Regel zwischen 3 % und 10 % und die durchschnittliche Abweichung im Bereich von 4 % bis 13 %. In Einzelfällen können jedoch Abweichungen von über 100 % auftreten. Insbesondere bei hoher Unsicherheit, großen Prognosefehlern oder stark ausgeprägten Schwankungen steigen die Kostenunterschiede an. Das PP-Verfahren erzielt in den meisten Fällen die geringsten Abweichungen und weist die geringste Streuung auf, ist jedoch anfällig für Ausreißer. LUC zeigt hingegen eine vergleichsweise hohe Robustheit gegenüber Ausreißern, während das SM-Verfahren durch große Streuungen und einen erhöhten Median auffällt (vgl. Abbildung 4-7 - Spalte 3).

Zusammenfassend lässt sich festhalten, dass beim vorliegenden Fehlertyp Rauschen das PP-Verfahren in den meisten Fällen das kostenoptimale Verfahren darstellt, mit der geringsten durchschnittlichen und medianen Abweichung. Gleichzeitig ist es jedoch ein sensibles Verfahren, insbesondere bei Unsicherheiten und niedrigen TBO-Werten. Das LUC-Verfahren erweist sich vor allem bei hoher Unsicherheit (großen Prognosefehlern) als robust und liefert stabile Ergebnisse. Die Heuristiken SM und Groff sind hingegen nur selten kostenoptimal, wobei das SM-Verfahren in der Regel die höchste durchschnittliche Kostenabweichung aufweist.

4.3.2 Fehlertyp – Überschätzung: Einfluss der Systemfaktoren auf die kostenorientierte Auswahl

Im nachfolgenden Kapitel wird der Einfluss der Systemfaktoren und ihrer Faktorstufen auf die kostenorientierte Auswahl der Heuristiken bei vorliegenden Fehlertyp Überschätzung analysiert. In der Abbildung 4-8 zeigt die linke Abbildung die Häufigkeit, mit der die einzelnen Heuristiken (SM, Groff, LUC und PP) bei einer **Überschätzung** als

kostenoptimale Lösung identifiziert werden. Die Verteilung verdeutlicht, dass die Heuristik PP mit einem Anteil von 59,46 % mit deutlichem Abstand am häufigsten die geringsten Kosten erzielt. Das Groff-Verfahren folgt mit 28,61 %, während LUC mit 2,56 % und SM mit 9,17 % deutlich seltener kostenoptimal sind.

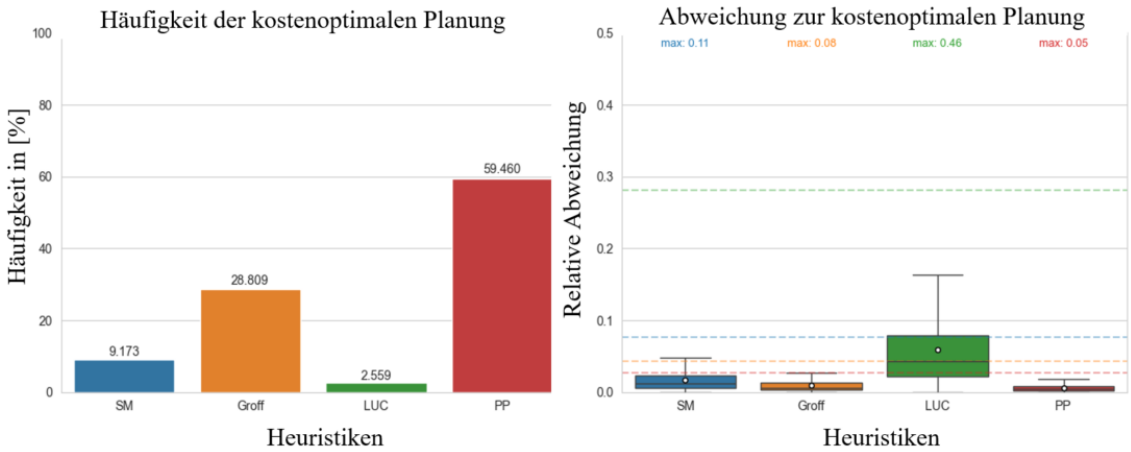

Abbildung 4-8: Überschätzung - Häufigkeit der kostenoptimalen Planung nach Heuristik und durchschnittliche Kostenabweichung zur kostenoptimalen Planung

Die **Häufigkeitsverteilung** unterscheidet sich deutlich von der Gesamtpopulation (vgl. Abbildung 4-1). Das PP-Verfahren dominiert weiterhin, während die Verfahren SM und LUC bei einer Überschätzung deutliche Schwächen zeigen. Gleichzeitig gewinnt das Groff-Verfahren an Relevanz. Die Ergebnisse lassen den Schluss zu, dass das PP-Verfahren bei einer Überschätzung in der Regel die wirtschaftlichste Lösung darstellt.

Die rechte Abbildung zeigt die relativen **Kostenabweichungen** der untersuchten Heuristiken zur jeweils kostenoptimalen Heuristik in Form eines **Box-Whisker-Plots**. Dabei wird deutlich, dass das PP-Verfahren nicht nur am häufigsten kostenoptimal ist, sondern auch die geringste Medianabweichung und die engste Streuung aufweist. Das Groff-Verfahren, gefolgt vom SM-Verfahren, zeigt ebenfalls geringe Streuungen, ist jedoch im Vergleich zum PP-Verfahren seltener kostenoptimal. Insgesamt sind die Kostenabweichungen zwischen den Verfahren PP, Groff und SM bei einer systematischen Überschätzung nur geringfügig. Auffällig ist hingegen die Leistung des LUC-Verfahrens, das im Vergleich zu den anderen drei Heuristiken deutlich größere Abweichungen aufweist.

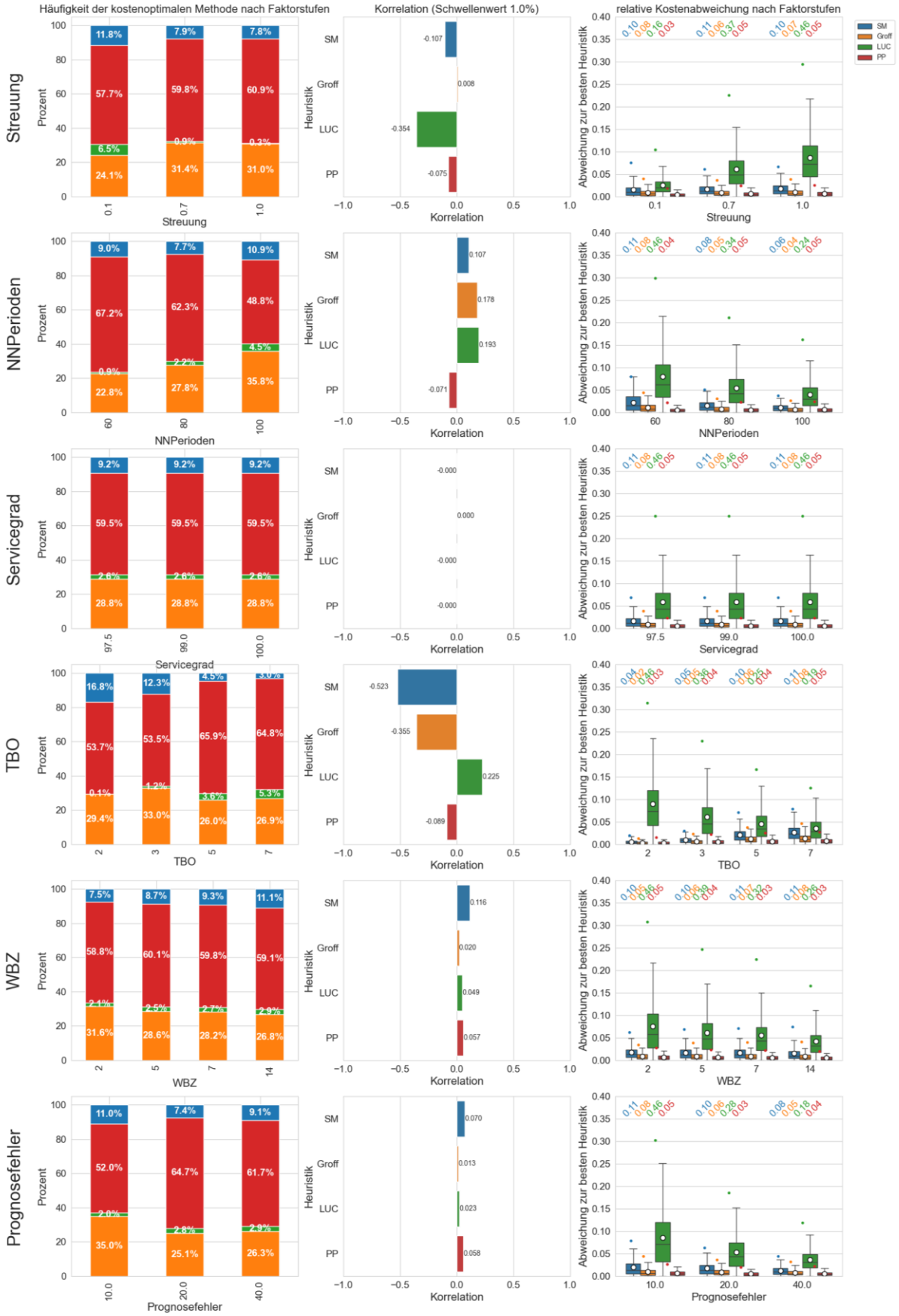

Abbildung 4-9: Überschätzung - Wirkzusammenhang der Systemfaktoren auf die kostenorientierte Auswahl der Heuristiken

Wie in den **Stapeldiagrammen** ersichtlich, dominiert die Heuristik PP (rot), gefolgt vom Groff-Verfahren (orange) (vgl. Abbildung 4-9 – erste Spalte). Die Verfahren SM (blau) und LUC (grün) sind hingegen nur in wenigen Faktorkombinationen kostenoptimal. Den stärksten Einfluss auf die kostenoptimale Auswahl zeigen die Systemfaktoren Streuung (Nachfrage), NNPerioden (Nachfrage) und TBO. Dies wird durch den Wechsel der kostenoptimalen Heuristik bei Änderung der Faktorstufen deutlich. Insgesamt sind die Effekte jedoch deutlich schwächer ausgeprägt als beim Fehlertyp Rauschen.

In der zweiten Spalte von Abbildung 4-9 sind die Ergebnisse der **punktbiserialen Korrelationsanalyse** dargestellt. Die größte Sensitivität der kostenoptimalen Auswahl zeigt sich bei den Systemfaktoren Streuung (Nachfrage), NNPerioden (Nachfrage) und TBO. Allerdings ist die Korrelation bei den beiden dominierenden Verfahren PP und Groff lediglich für den Faktor TBO ersichtlich. In Verbindung mit den Ergebnissen der Stapeldiagramme zeigt sich, dass zwar bei den Verfahren SM und LUC für mehrere Systemfaktoren Korrelationen bestehen, diese Verfahren jedoch nur selten kostenoptimal sind. Somit ist ihr Einfluss auf die kostenorientierte Auswahl von untergeordneter Bedeutung.

Die Auswertung der **Boxplots** verdeutlicht, dass bei einer Überschätzung zwischen den Verfahren SM, Groff und PP im Durchschnitt nur geringe Kosteneinsparungen erzielt werden können. Wird hingegen das LUC-Verfahren gewählt, muss mit deutlich höheren Kosten gerechnet werden, häufig mehr als 5 Prozentpunkte über den übrigen Verfahren (vgl. Abbildung 4-9 – dritte Spalte).

Zusammenfassend lässt sich festhalten, dass beim vorliegenden Fehlertyp Überschätzung das PP-Verfahren in den meisten Fällen die kostenoptimale Lösung darstellt und sowohl die geringste durchschnittliche als auch die niedrigste mediane Kostenabweichung aufweist. Das Kosteneinsparungspotenzial ist jedoch insgesamt geringer als beim Fehlertyp Rauschen, insbesondere beim Vergleich der Verfahren SM, Groff und PP. Bei einer Überschätzung sollte das LUC-Verfahren vermieden werden, da es zu deutlich höheren Kosten führt.

4.3.3 Fehlertyp – Unterschätzung: Einfluss der Systemfaktoren auf die kostenorientierte Auswahl

Im nachfolgenden Kapitel wird der Einfluss der Systemfaktoren und ihrer Faktorstufen auf die kostenorientierte Auswahl der Heuristiken bei vorliegenden Prognosefehlertyp **Unterschätzung** analysiert.

In Abbildung 4-10 zeigt die linke Abbildung die **Häufigkeit**, mit der die einzelnen Heuristiken (SM, Groff, LUC und PP) bei einem vorliegenden Prognosefehler Unterschätzung als kostenoptimale Lösung identifiziert werden. Die Verteilung verdeutlicht, dass die Heuristiken PP (ca. 48 %) und LUC (ca. 45 %) mit deutlichem

Abstand am häufigsten zu den geringsten Kosten führen, während die Verfahren Groff (ca. 4,5 %) und SM (ca. 2,5 %) nur selten eine kostenoptimale Planung ermöglichen.

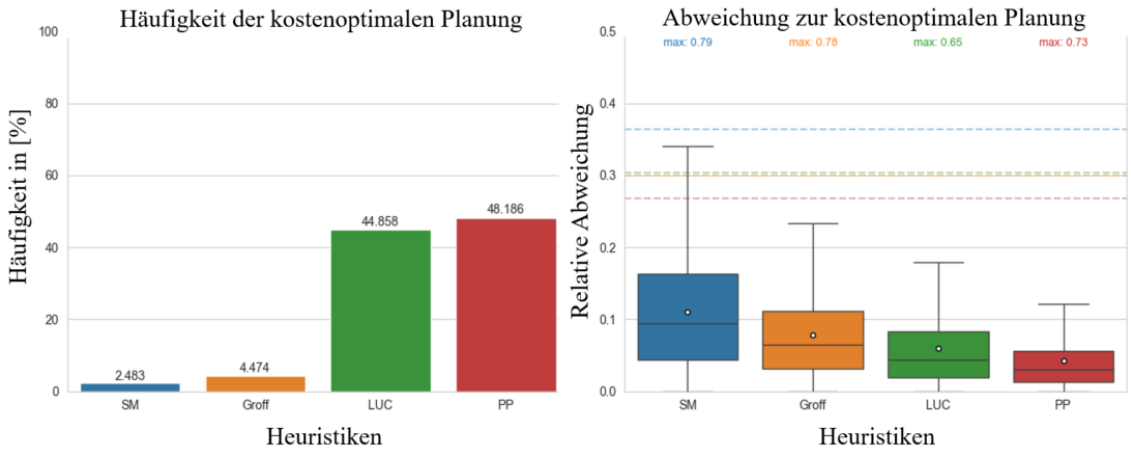

Abbildung 4-10: Unterschätzung - Häufigkeit der kostenoptimalen Planung nach Heuristik und durchschnittliche Kostenabweichung zur kostenoptimalen Planung

Die **Häufigkeitsverteilung** unterscheidet sich von der Gesamtpopulation (vgl. Abbildung 4-1), sodass sich auch hier die Einschätzung bestätigt, dass der Fehlertyp einen erheblichen Einfluss auf die kostenorientierte Auswahl der Heuristiken hat. Liegt eine Unterschätzung vor, sollte entweder das PP-Verfahren oder das LUC-Verfahren gewählt werden.

Die rechte Abbildung zeigt die relativen **Kostenabweichungen** der untersuchten Heuristiken zur jeweils kostenoptimalen Heuristik in Form eines **Box-Whisker-Plots**. Dabei wird deutlich, dass das PP-Verfahren nicht nur am häufigsten kostenoptimal ist, sondern auch die geringste Medianabweichung und die engste Streuung aufweist. Das LUC-Verfahren zeigt ebenfalls geringe Streuungen und weist im Vergleich zum PP-Verfahren nur geringfügig höhere durchschnittliche Kosten auf. Die Verfahren Groff und SM sind dagegen nur selten kostenoptimal, insbesondere das SM-Verfahren weist mit durchschnittlich etwa 10 % die größte Kostenabweichung auf.

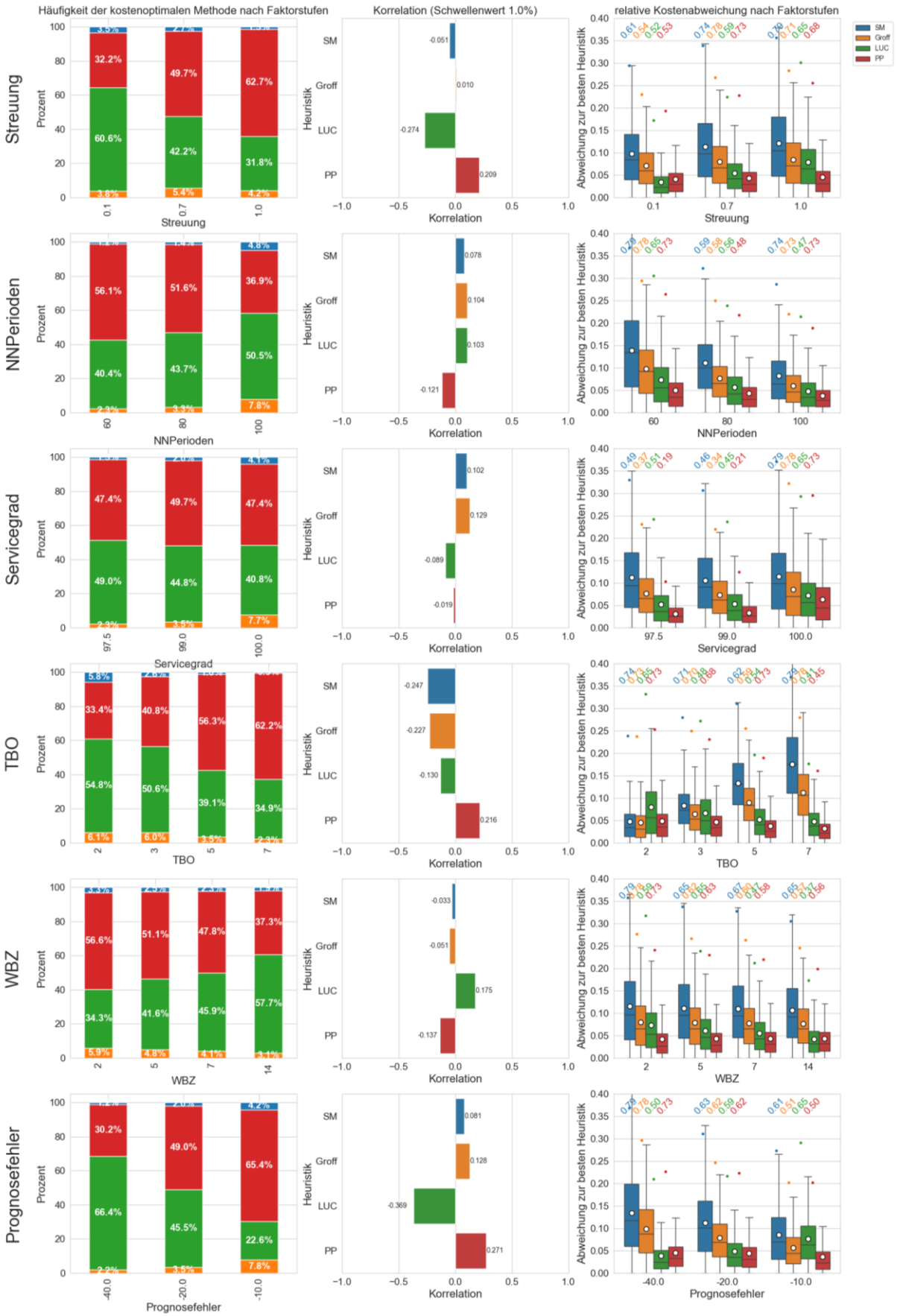

Abbildung 4-11: Unterschätzung - Wirkzusammenhang der Systemfaktoren auf die kostenorientierte Auswahl der Heuristiken

Wird die Häufigkeit der kostenoptimalen Planung in Abhängigkeit der unterschiedlichen Systemfaktoren mittels **Stapeldiagrammen** dargestellt, zeigt sich, dass die Heuristik PP (rot) und das LUC-Verfahren (grün) dominieren (vgl. Abbildung 4-11). Auffällig ist, dass bei den Verfahren PP und LUC ein gegenläufiges Verhalten zu beobachten ist. Mit zunehmender bzw. abnehmender Faktorstufe erfolgt ein Wechsel des kostenoptimalen Verfahrens von PP zu LUC bzw. von LUC zu PP. Die Verfahren SM (blau) und Groff (orange) sind dagegen nur in wenigen und spezifischen Faktorkombinationen kostenoptimal. Den größten Einfluss auf die kostenorientierte Auswahl zeigen die Systemfaktoren Prognosefehler, TBO und Streuung (Nachfrage). Dies wird durch den Wechsel der kostenoptimalen Heuristik bei Änderung der Faktorstufen sichtbar.

In der zweiten Spalte der Abbildung 4-11 sind die Ergebnisse der **punktbiserialen Korrelationsanalyse** dargestellt. Für die beiden dominierenden Verfahren PP und LUC zeigen sich die höchsten Korrelationen bei den Faktoren Prognosefehler, TBO und Streuung (Nachfrage). Änderungen der Faktorstufen führen hier zu Wechseln in der Auswahl des jeweils optimalen Verfahrens, was durch die Ergebnisse der Korrelationsanalyse bestätigt wird. Niedrige Werte bei TBO und Streuung (Nachfrage) begünstigen LUC, während hohe Werte PP begünstigen. Große Prognosefehler wirken sich positiv auf LUC und kleine Prognosefehler positiv auf PP aus. Für die Verfahren SM und Groff besteht hingegen nur eine signifikante Korrelation beim Faktor TBO. Ihre Relevanz ist aufgrund der insgesamt schwachen Leistung jedoch gering.

Die Auswertung der **Boxplots** verdeutlicht, dass auch bei einer vorliegenden Unterschätzung die Wahl des Verfahrens erhebliche Auswirkungen auf die entstehenden Kosten hat. Bei diesem Fehlertyp treten die höchsten Kostenabweichungen auf. Die mediane Kostenabweichung zwischen den Verfahren liegt bei etwa 3 % bis 14 %, die durchschnittliche Abweichung im Bereich von 4 % bis 17 %. Die maximalen Abweichungen (79 %) fallen jedoch geringer aus als bei den anderen beiden Fehlertypen. Besonders große Kostenunterschiede sind bei hohen Prognosefehlern (Unterschätzungen), einer hohen Anzahl an Nullperioden (NNPerioden) oder großen TBO-Werten zu beobachten. Die Verfahren PP und LUC erzielen meist die geringsten Abweichungen und die engste Streuung. Hier ist, abhängig von den relevanten Faktoren, die Auswahl des geeigneten Verfahrens entscheidend (vgl. Abbildung 4-11 – Spalte 3)

Zusammenfassend lässt sich festhalten, dass beim vorliegenden Prognosefehler Unterschätzung die Verfahren PP und LUC in der Regel die kostenoptimalen Verfahren darstellen und sowohl die geringsten durchschnittlichen als auch die niedrigsten medianen Abweichungen aufweisen. Für die Auswahl zwischen PP und LUC sind die Faktoren NNPerioden (Nachfrage), TBO und Prognosefehler von zentraler Bedeutung. Die Heuristiken SM und Groff sind hingegen nur selten kostenoptimal und können weitgehend vernachlässigt werden.

4.4 Ableitung von Entscheidungsregel zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren

Zusammenfassend zeigt sich in den Kapiteln 4.3.1 bis 4.3.3, dass die Heuristik PP in den meisten Fällen das dominierende Verfahren darstellt, gefolgt vom LUC-Verfahren. Nur im Fall einer systematischen Überschätzung erzielen die Verfahren Groff und SM vergleichsweise gute Ergebnisse, übertreffen jedoch auch hier im Durchschnitt nicht das PP-Verfahren. Die Relevanz der beiden Verfahren Groff und SM kann daher insgesamt als gering eingestuft werden. Darüber hinaus treten bei den Fehlertypen Rauschen und Unterschätzung die größten Kostenabweichungen zwischen den vier Heuristiken auf.

Die Systemfaktoren Prognosefehler und TBO besitzen den größten Einfluss auf die kostenorientierte Auswahl der vier Heuristiken. Hinsichtlich der Faktoren und Faktorstufen zeigt sich häufig ein gegenläufiges Verhalten zwischen den Verfahren PP und LUC. Das LUC-Verfahren erweist sich als robust, während das PP-Verfahren empfindlich auf Unsicherheiten reagiert. Dieses Verhalten zeigt sich darin, dass das LUC-Verfahren bei großen Prognosefehlern und kleinen TBO-Werten gute Ergebnisse liefert, wohingegen das PP-Verfahren vor allem bei kleinen Prognosefehlern und hohen TBO-Werten vorteilhaft ist. Die Verfahren SM und Groff sind hingegen nur in wenigen und spezifischen Faktorkombinationen geeignet.

Im Hinblick auf die Gestaltung der Entscheidungsregeln sollen diese niederschwellig und praxisnah anwendbar sein. Da die Entscheidungsregeln einerseits eine einfache Anwendung ermöglichen und andererseits die Faktoren Prognosefehler und TBO den größten Einfluss auf die kostenorientierte Auswahl besitzen, werden die Entscheidungsregeln auf Basis der beiden zentralen Systemfaktoren entwickelt. Eine Lösung, die sämtliche Faktorkombinationen berücksichtigt, wird in Kapitel 5 mittels eines Machine Learning Modells entwickelt.

Neben den Entscheidungsregeln werden als Entscheidungsunterstützung **Scatterplots** für die drei Fehlertypen Rauschen, Überschätzung und Unterschätzung erstellt. In den Scatterplots werden die Ausprägungen der TBO sowie die Höhe des Prognosefehlers als Matrix dargestellt. Für jede Faktorkombination wird die kostenoptimale Heuristik farblich durch einen Kreis dargestellt. Die Größe des Kreises gibt die durchschnittliche prozentuale Kostenabweichung der kostenoptimalen Heuristik im Vergleich zu den anderen drei Verfahren an. Je größer der Kostenvorteil der kostenoptimalen Heuristik, desto größer der Durchmesser des Kreises. Die Abbildung 4-12 bis Abbildung 4-14 zeigen die entsprechenden Scatterplots für die drei Fehlertypen Rauschen, Überschätzung und Unterschätzung.

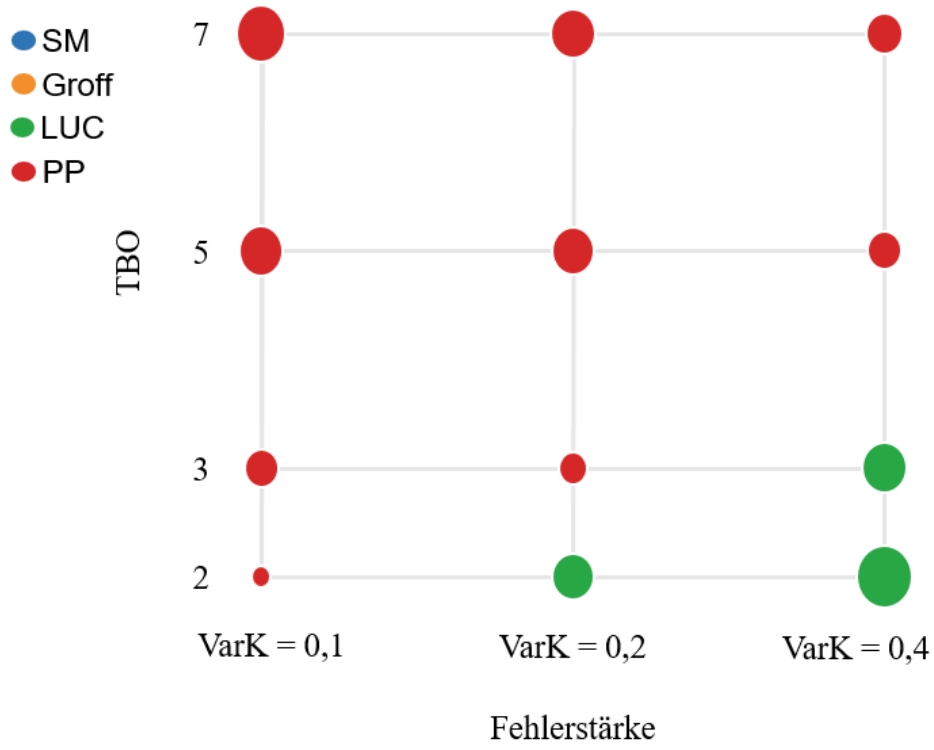

Abbildung 4-12: Scatterplot: Rauschen

Wie in Abbildung 4-12 ersichtlich, führen nur die Verfahren PP und LUC zu kostenoptimalen Ergebnissen, sodass die Auswahl im Falle des Fehlertyps **Rauschen** ausschließlich zwischen diesen beiden Verfahren zu treffen ist. Die Verfahren SM und Groff liefern bei keiner Faktorkombination aus Prognosefehler und TBO kostenoptimale Ergebnisse.

Die Kostenabweichungen zwischen dem kostenoptimalen Verfahren und den durchschnittlichen Kosten der übrigen drei Verfahren liegen im Bereich von etwa 1,1 % bis 9,3 %. Das geringste Kosteneinsparungspotenzial ergibt sich bei kleinen Prognosefehlern und niedrigen TBO-Werten, während das größte Kosteneinsparungspotenzial bei kleinen TBO und großen Prognosefehlern auftritt (vgl. Abbildung 4-12).

Wie bereits in Kapitel 4.3 dargestellt, zeigen die Verfahren LUC und PP ein gegenläufiges Verhalten. Das PP-Verfahren bleibt insgesamt das dominierende Verfahren und eignet sich insbesondere bei hohen TBO-Werten und kleinen Prognosefehlern. Steigt der Prognosefehler an und ist die TBO gering, gewinnt das LUC-Verfahren an Bedeutung.

Für den Fehlertyp *Rauschen* lassen sich konkrete und niederschwellige **Entscheidungsregeln** ableiten, die auf den beiden Systemfaktoren Prognosefehler und TBO sowie auf dem gegenläufigen Verhalten der Heuristiken PP und LUC basieren. Die Analyse zeigt, dass das **LUC-Verfahren** grundsätzlich dann zu bevorzugen ist, wenn der Prognosefehler größer als 0,4 ist. Liegt der Prognosefehler hingegen zwischen 0,1 und 0,4, sollte das LUC-Verfahren nur dann angewendet werden, wenn zusätzlich die TBO kleiner als 2,5

ist. In allen übrigen Fällen führt das **PP-Verfahren** zu den kostengünstigeren Ergebnissen und ist entsprechend auszuwählen.

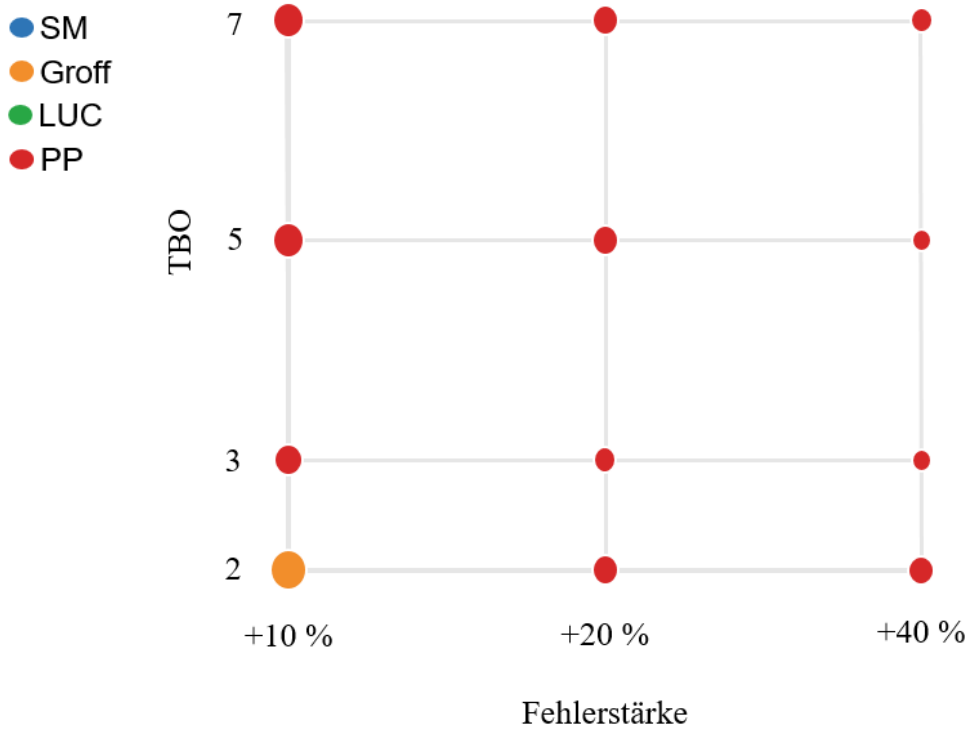

Abbildung 4-13: Scatterplot: Überschätzung

Im Falle einer systematischen **Überschätzung** zeigt sich die Dominanz der Heuristik PP (vgl. Abbildung 4-13). Nur bei kleinen Prognosefehlern (10 %) und einer niedrigen TBO (2) führt das Groff-Verfahren zu den geringsten Kosten. In allen anderen Faktorkombinationen weist das PP-Verfahren die niedrigsten Kosten auf. Insgesamt fallen die Kostenabweichungen zwischen den einzelnen Verfahren geringer aus als beim Fehlertyp Rauschen (vgl. Abbildung 4-12) und Unterschätzung (vgl. Abbildung 4-13). Die Kostenabweichungen reichen von ca. 1,5 % (TBO = 7 und Prognosefehler = +40 %) bis maximal ca. 4 % (TBO = 2 und Prognosefehler = +10%).

Für den Fehlertyp *Überschätzung* ist das Groff-Verfahren ausschließlich bei niedrigen TBO-Werten (TBO < 2) und kleinen Prognosefehlern (Prognosefehler < +10 %) geeignet. In allen übrigen Faktorkombinationen ist das PP-Verfahren zu bevorzugen. Daher lässt sich die vereinfachte **Entscheidungsregel** ableiten, dass bei einer systematischen Überschätzung ausschließlich das **PP-Verfahren** verwendet werden sollte.

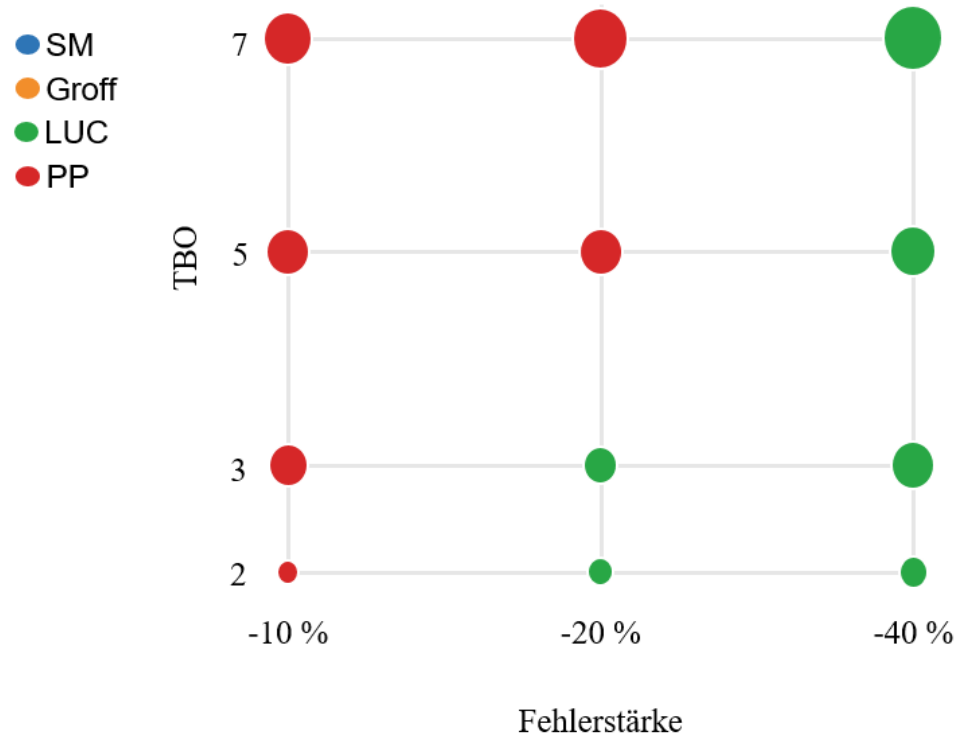

Abbildung 4-14: Scatterplot: Unterschätzung

Liegt ein Prognosefehler in Form einer systematischen Unterschätzung vor, zeigt sich ein ähnliches Bild wie beim Fehlertyp Rauschen, sodass die Verfahren PP und LUC dominieren. Auch das gegenläufige Verhalten dieser beiden Verfahren ist deutlich ersichtlich. Große TBO-Werte und kleine Prognosefehler wirken sich positiv auf das PP-Verfahren aus, während kleine TBO-Werte und große Prognosefehler das LUC-Verfahren begünstigen. Mit zunehmender TBO steigen sowohl die Kostenabweichungen als auch das Kosteneinsparungspotenzial. Die größten Kostenabweichungen (ca. 11,5 %) treten bei einer TBO von sieben und einem Prognosefehler von -40 % auf, während die geringsten Kostenabweichungen (ca. 2 %) bei einer TBO von zwei und einem Prognosefehler von -10 % beobachtet werden.

Für den Fehlertyp *Unterschätzung* lassen sich ebenso konkrete und niederschwellige **Entscheidungsregeln** ableiten, die auf den beiden Systemfaktoren Prognosefehler und TBO sowie auf dem gegenläufigen Verhalten der Heuristiken PP und LUC basieren. Die Analyse zeigt, dass das **LUC-Verfahren** grundsätzlich dann zu bevorzugen ist, wenn der Prognosefehler größer als - 30 % ist. Liegt der Prognosefehler hingegen zwischen -10 % und -30 %, sollte LUC nur dann angewendet werden, wenn zusätzlich die TBO kleiner als 3,5 ist. In allen übrigen Fällen führt das **PP-Verfahren** zu den kostengünstigeren Ergebnissen und ist entsprechend auszuwählen.

Auf Grundlage der definierten Entscheidungsregeln lässt sich das Entscheidungsverhalten mithilfe eines Entscheidungsbaums transparent darstellen. Der Entscheidungsbaum ist in Abbildung 4-15 ersichtlich. Dabei wird hinsichtlich der drei systematischen Prognosefehler, der Höhe des Prognosefehlers sowie der TBO differenziert. Abhängig von der jeweiligen Faktorausprägung kann anschließend kostenorientiert die passende Heuristik ausgewählt werden.

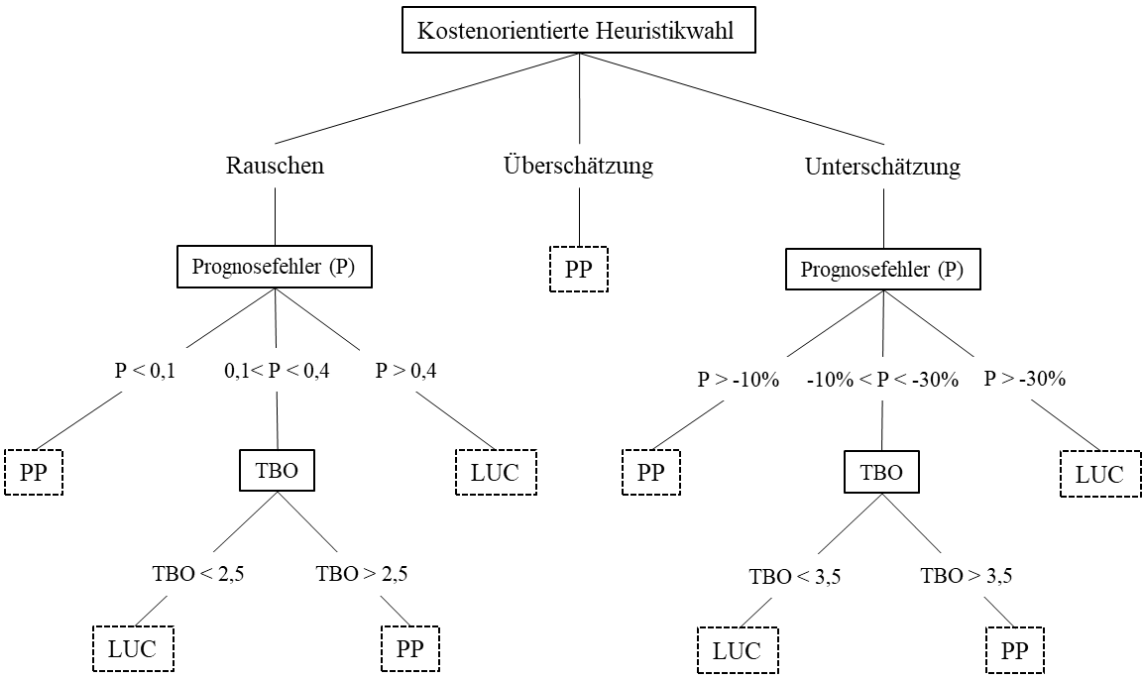

Abbildung 4-15: Entscheidungsbaum - Kostenorientierte Heuristikwahl

Darüber hinaus ist es erforderlich, die getroffene **Heuristikwahl in regelmäßigen Abständen zu überprüfen**. Dabei ist es zu empfehlen, die Systemfaktoren der dynamischen Bestellmengenplanung fortlaufend zu erfassen und zu analysieren. Insbesondere Veränderungen des Prognosefehlers sowie der Kostenstruktur (TBO) können dazu führen, dass die zuvor optimale Heuristik nicht mehr die kosteneffizienteste Wahl darstellt. Anpassungen in der Kostenstruktur lassen sich unmittelbar durch veränderte Lagerhaltungs- oder Bestellkosten identifizieren. Um auch Veränderungen im Prognosefehler systematisch zu erfassen, sollten die Prognosefehler fortlaufend gespeichert, überwacht und analysiert werden.

Von besonderer Bedeutung ist die frühzeitige Erkennung eines möglichen Wechsels des zugrunde liegenden Fehlertyps. Verbesserungen der Prognosegüte können ebenso wie die Wahl der jeweils richtigen Heuristik maßgeblich zur Reduktion der Gesamtkosten beitragen.

4.5 Zwischenfazit: Entscheidungsregeln zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren

Dieses Kapitel widmet sich der dritten Forschungsfrage: „*Welche Entscheidungsregeln können zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren definiert werden?*“. Dabei wird die Entwicklung von Entscheidungsregeln zur kostenorientierten Auswahl fokussiert. Die Basis bildet das No-Free-Lunch Theorem (NFL), das besagt, dass kein einzelnes Verfahren für alle Problemstellungen gute Ergebnisse liefert. Für die Praxis bedeutet dies, dass je nach Faktorkombination situativ das kostengünstigste dynamische Bestellmengenverfahren ausgewählt werden muss.

Als Datengrundlage werden die synthetischen Daten aus dem dritten Kapitel genutzt und deren Eigenschaften und Wirkzusammenhänge analysiert. Zur Vergleichbarkeit der Verfahren werden die prozentualen Kostenabweichungen zum jeweils kostenoptimalen Verfahren berechnet. Die Auswertung zeigt, dass Part Period (PP) mit 53 % am häufigsten die geringsten Kosten erzielt, gefolgt von Least Unit Cost (LUC) mit 26 %. Silver Meal (SM) und Groff sind nur in rund 21 % der Fälle am kostengünstigsten. Die Boxplot-Analysen verdeutlichen, dass PP zwar die niedrigsten Medianabweichungen aufweist, jedoch anfällig für hohe Extremwerte ist. LUC zeigt eine höhere Robustheit, allerdings mit etwas größeren durchschnittlichen Abweichungen.

Mittels Cramér’s V wird der Zusammenhang zwischen den Systemfaktoren und der kostenoptimalen Heuristik untersucht. Der Fehlertyp des Prognosefehlers hat den größten Einfluss, gefolgt von der Time Between Order (TBO). Die anderen Faktoren wie Nachfrage (Streuung und Anzahl Nullperioden), WBZ oder Servicegrad besitzen einen geringeren Einfluss. Wird der Einfluss der Systemfaktoren separat nach den drei Fehlertypen analysiert, so weisen die Fehlertypen Rauschen und Unterschätzung ein ähnliches Verhalten auf. Bei Überschätzungen entfällt der Einfluss des Servicegrades, während die TBO einen starken Einfluss besitzt.

Zur tiefergehenden Untersuchung wurde die Zielvariable „kostenoptimale Heuristik“ binär kodiert und ein Schwellenwert von 1 % festgelegt, um robustere Ergebnisse zu erhalten. Beim Fehlertyp Rauschen und Unterschätzung ist meistens das PP-Verfahren optimal, jedoch zeigt es Schwächen bei kleinen TBOs und großen Prognosefehlern. LUC ist robuster und bei großen Prognosefehlern vorzuziehen. SM und Groff sind selten geeignet. Beim Fehlertyp Überschätzung ist das PP-Verfahren dominierend. Groff ist nur bei kleinen Prognosefehlern relevant. LUC sollte vermieden werden, da es hohe Mehrkosten verursacht.

Auf Basis der Erkenntnis, dass die beiden Systemfaktoren Prognosefehler (Fehlertyp und -stärke) und TBO den größten Einfluss auf die kostenorientierte Auswahl aufweisen, können die Entscheidungsregeln auf Grundlage der beiden Systemfaktoren entwickelt werden. Mit dem Ziel, niederschwellige und praxisnahe Regeln zu erhalten. Die Entscheidungsregeln werden durch Scatterplots, unterteilt nach den drei Fehlertypen unterstützt und mittels eines Entscheidungsbaums transparent visualisiert (vgl. Kapitel 4.4). Damit ermöglichen die entwickelten Entscheidungsregeln eine praxisorientierte, kostenorientierte Auswahl von den dynamischen Bestellmengenverfahren.

5 Entwicklung eines Machine Learning Modells zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren

Die Untersuchung der synthetischen Daten zeigt, dass die betrachteten dynamischen Bestellmengenverfahren je nach Faktorkombination deutliche Kostenunterschiede aufweisen. Die entwickelten Entscheidungsregeln basieren dabei auf den beiden zentralen Systemfaktoren Prognosefehler und TBO, um niederschwellige und praxisnahe anwendbare Ergebnisse zu ermöglichen. Parallel dazu verfolgt dieses Kapitel das Ziel, ein datengetriebenes Machine Learning Modell zur Entscheidungsfindung zu entwickeln, welches alle Faktorkombinationen berücksichtigt. Dementsprechend wird die vierte Forschungsfrage: „*Wie kann ein Machine Learning Modell zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren gestaltet werden?*“ fokussiert.

In diesem Kapitel wird einerseits die kostenorientierte Auswahl dynamischer Bestellmengenverfahren untersucht und andererseits die Verbindung von simulativer Datenerzeugung und Machine Learning. Dabei wird erforscht, inwieweit auf Basis synthetischer Daten ein ML-Modell zur dynamischen Bestellmengenplanung gestaltet werden kann. Im Kern liegt auch diesem Kapitel das No-Free-Lunch-Theorem zugrunde, sodass kontextspezifische Leistungsdifferenzen durch das Entscheidungsmodell erkannt werden sollen, um eine kostenorientierte Auswahl zu ermöglichen. Methodisch wird dies als überwachte Mehrklassen-Klassifikationsaufgabe formuliert, bei der aus mehreren Heuristiken diejenige ausgewählt wird, die bei vorliegender Faktorkombination erwartungsgemäß die geringsten Kosten verursacht.

Zur Beantwortung der vierten Forschungsfrage werden zunächst grundlegende Konzepte des Machine Learnings mit Fokus auf Klassifikationsverfahren eingeführt und der Problemrahmen formalisiert. Darauf aufbauend werden die spezifischen Anforderungen an geeignete lernende Klassifikationsmethoden hergeleitet und die Auswahl der in dieser Arbeit verwendeten Verfahren (Entscheidungsbaum, Random Forest und LightGBM) begründet. Anschließend wird die Implementierung und die Evaluation beschrieben, einschließlich Datenaufbereitung, Validierung, Hyperparameteroptimierung und Auswahl der Bewertungsmetriken. Abschließend werden die Modelle auf den synthetischen Daten der dynamischen Bestellmengenplanung (vgl. Kapitel 3) trainiert, getestet und ihre Ergebnisse hinsichtlich unterschiedlicher Zielgrößen systematisch interpretiert.

5.1 Einführung ins Machine Learning

Machine Learning (ML) ist ein Teilbereich der Künstlichen Intelligenz, der sich mit Algorithmen und Verfahren beschäftigt, welche das Ziel haben ihr Verhalten auf Basis von Erfahrungen bzw. Daten zu verbessern (Mitchell 1997, S. 2). Im Kontext dieser Arbeit kann ML als datengetriebener Prozess verstanden werden, der aus vorhandenen Daten lernt, zugrunde liegende Muster extrahiert und Zusammenhänge identifiziert (Kelleher et al. 2015, S. 3). Klassischerweise wird ML in überwachtes Lernen (supervised learning), unüberwachtes Lernen (unsupervised learning) und Reinforcement Learning unterteilt (Qin 2020, S. 12 ff.). Beim überwachten Lernen werden Modelle anhand gelabelter Trainingsdaten trainiert. Liegt die Zielvariable kategorial vor, wird von Klassifikation gesprochen, bei kontinuierlichen Zielgrößen von Regression. Im unüberwachten Lernen stehen ausschließlich Eingabedaten zur Verfügung, sodass die selbstständige Strukturentdeckung, etwa durch Clustering oder Dichteschätzung das Ziel ist (Bishop 2006, S. 7). Das reinforcement Learning adressiert sequenzielle Entscheidungsprobleme mit begrenzter und meist verzögerter Rückmeldung und ist damit zwischen überwachten und unüberwachten Lernen verortet (Sutton und Barto 2020, S. 6 f.). Im Gegensatz zum überwachten Lernen agiert ein Agent mit der Umgebung, sammelt verzögert bereitgestellte Belohnungen und leitet basierend darauf Entscheidungen ab (Qin 2020, S. 14). Die nachfolgende Abbildung zeigt die beschriebenen Teilbereiche des Machine Learnings.

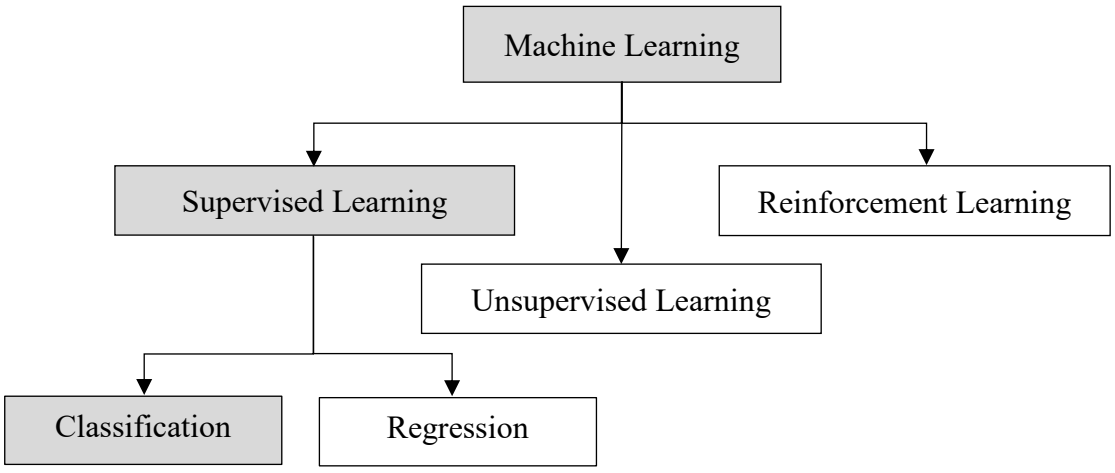

Abbildung 5-1: Teilbereiche des Machine Learnings

Die vorliegende Arbeit fokussiert die Klassifikation, sodass auf Basis gegebener Faktorkombinationen (z. B. Prognosefehler und weiterer Systemfaktoren) die kostenoptimale Heuristik ausgewählt wird. Für diesen Problemtyp wurden unterschiedliche Klassifikationsverfahren im Bereich Machine Learning entwickelt. Die Verfahren unterscheiden sich

z. B. in Funktionsprinzip, Laufzeitverhalten, Prognosegüte, Interpretierbarkeit und Skalierbarkeit. Der Großteil der Forschung im überwachten Lernen adressiert Single-Label-Klassifikation (Tsoumakas et al. 2010, S. 667). Was auch der vorliegenden Zielstellung entspricht, es soll genau eine Heuristik identifiziert werden. In diesem Rahmen können unterschiedliche Machine Learning Modelle eingesetzt werden. Vor allem im Bereich des überwachten Lernens kommen häufig baumbasierenden Methoden für Klassifizierungsaufgaben zum Einsatz. Sie weisen häufig gute Ergebnisse, bei gleichzeitig guter Interpretierbarkeit auf. (Grinsztajn et al. 2022, S. 9; Knuth 2021, S. 364 ff.)

5.2 Auswahl der Klassifizierungsmodelle

Im Rahmen der baumbasierenden Klassifizierungsverfahren wurden eine Vielzahl an unterschiedlichen Verfahren entwickelt. Das beste Verfahren ist dabei von Anwendungsfall zu Anwendungsfall unterschiedlich. So übertrifft das LightGBM-Verfahren in der Vergleichsstudie von WANG et al. die Modelle XGBoost und Random Forest (Wang et al. 2017, S. 7). In anderen Studien war der Random Forest Algorithmus am besten (Lan et al. 2020, S. 2052; Chen et al. 2020, S. 52). Insgesamt weisen die Verfahren Random Forest und LightGBM häufig gute Ergebnisse für Klassifizierungsprobleme auf (Monteiro et al. 2024, S. 5 f.; Indah et al. 2025, S. 17). Demzufolge werden für die vorliegende Arbeit die Verfahren Random Forest und LightGBM ausgewählt. Darüber hinaus wird als grundlegendes Verfahren sowie aufgrund der Interpretierbarkeit auch die klassische Form des Entscheidungsbaumes berücksichtigt. Nachfolgend werden die ausgewählten Klassifikationsverfahren *Entscheidungsbäume*, *Random Forest* und *LightGBM* näher erläutert. Dabei werden die grundlegenden Konzepte, Algorithmen und wichtige Eigenschaften dieser Verfahren beschrieben.

5.2.1 Entscheidungsbäume

Bei Entscheidungsbäumen (Decision Trees) handelt es sich um ein überwachtes Verfahren des Machine Learnings, welches sowohl für die Klassifikation, als auch für die Regression verwendet werden kann (Bishop 2006, S. 654). Entscheidungsbäume sind hierbei, baumähnliche Modelle, bestehend aus Knoten, Blättern und Zweigen.

Die Knoten stellen Entscheidungen, basierend auf dem Wert einer Eingabevariable dar. Die Zweige, die von einem solchen Entscheidungsknoten ausgehen, stellen die möglichen Ausgänge der Entscheidung dar. Am Ende des Baumes bilden die Blätter die Vorhersagen des Modells ab. (Quinlan 1986, S. 86) Im Falle einer Klassifikation können dabei die Blätter entweder eine einzelne Klasse oder eine Wahrscheinlichkeitsverteilung über mehrere Klassen beinhalten (Maimon und Rokach 2005, S. 166). Zur Klassifizierung eines

Objekts werden Entscheidungen von der Wurzel aus durchlaufen, wobei die Variablenwerte die zu verfolgende Verzweigung bestimmen. Das resultierende Blatt veranschaulicht die vom Modell prognostizierte Klasse oder Klassenverteilung (Quinlan 1986, S. 86; Maimon und Rokach 2005, S. 196).

Es gibt mehrere Möglichkeiten für die Implementierung von Entscheidungsbäumen. Dabei werden oft die Methoden CART, ID3 und C4.5 angeführt. ID3 ist ein simpler Algorithmus für Entscheidungsbäume, der erstmals 1986 vorgestellt wurde (Quinlan 1986, S. 81 ff.). Er unterstützt nur kategoriale Attribute und führt kein Pruning durch, was zu Overfitting führen kann (Maimon und Rokach 2005, S. 181).

Entscheidungsbäume haben oft das Problem ein Overfitting an das Trainingsset zu entwickeln, wenn kein Haltekriterium gewählt wird. Dies geschieht, wenn der Baum zu groß wird und sich zu stark an die speziellen Gegebenheiten in den Testdaten anpasst und dort gute Ergebnisse liefert, während die Leistung auf neuen und unbekannten Daten abnimmt. Als Methode gegen ein Overfitting von Entscheidungsbäumen werden Pruning-Methoden verwendet (Maimon und Rokach 2005, S. 175; Breiman et al. 2017, S. 63 ff.). Dabei werden die Entscheidungsbäume erst durch Haltekriterien an die Trainingsdaten overfitted und im Anschluss durch das Entfernen von Teilästen, die nicht zur Generalisierung beitragen, wieder verkleinert.

C4.5 ist eine Weiterentwicklung von ID3, welches zum einen Geschwindigkeitsverbesserungen mit sich bringt und zum anderen Pruning gegen Overfitting einsetzt (Maimon und Rokach 2005, S. 181; Sharma und Kumar 2016, S. 2095). CART (Classification and Regression Trees) ist ein zeitgleich entwickelter Algorithmus, welcher sich hauptsächlich von ID3 und C4.5 darin unterscheidet, dass er binäre Bäume erzeugt, also jeder Entscheidungsknoten immer nur genau zwei Wege hat. Dies vereinfacht die Interpretierung und Visualisierung des Baumes. Weiterhin ist es mit CART möglich, nicht nur eine Klassifizierung, sondern auch eine Regression durchzuführen. CART verwendet außerdem wie C4.5 ein Pruning Verfahren, um Overfitting zu vermeiden (Maimon und Rokach 2005, S. 181). Aufgrund der höheren Interpretierbarkeit von CART und der bereits vorhandenen Implementierung in den verwendeten Frameworks (Pedregosa et al. 2011) wird sich in dieser Arbeit auf die Anwendung von CART konzentriert.

CART erzeugt für gegebene Trainingsdaten einen Entscheidungsbaum, indem der Datensatz rekursiv aufgeteilt wird. Die Aufteilung wird dabei basierend auf den Attributen getroffen, die den höchsten Informationsgewinn versprechen. Dieser Prozess wird so lange wiederholt, bis das Erstellen weiterer Aufteilungen keinen signifikanten Informationsgewinn mehr liefert oder ein Haltekriterium erreicht wird (Maimon und Rokach 2005, S. 168).

Der Informationsgewinn ist ein Maß für die Reduzierung der Unsicherheit durch eine Aufteilung der Daten in zwei unterschiedliche Bereiche. Er bestimmt sich aus dem Unterschied zwischen der Unsicherheit vor und nach dem Aufteilen der Daten. Für den Informationsgewinn können verschiedene Maße verwendet werden, wobei üblicherweise der Gini-Index oder die Cross-Entropy verwendet wird (Bishop 2006, S. 666). Beide fördern die Aufteilung in homogene Bereiche, d. h. Bereiche, in denen möglichst viele Datenpunkte zur gleichen Klasse gehören.

CART verwendet für das Pruning das sogenannte „Minimal Cost-Complexity Pruning". Hierbei wird ein Komplexitätsparameter verwendet, welcher die Pruning Stärke bestimmt. So sorgt ein höherer Komplexitätsparameter für einen höheren Einfluss der Blattanzahl in den Komplexitätswert eines Teilbaums. Demgegenüber führt ein niedriger Komplexitätsparameter zu wenig Pruning und somit einem komplexeren Baum, während höhere Werte zu einem stärkeren Pruning und einem simpleren Baum führen (Breiman et al. 2017).

5.2.2 Ensembleverfahren

Ensembleverfahren sind Verfahren für Machine Learning, welche mehrere einfache Modelle, sogenannte „base learners" (Zhou 2012, S. 15) oder Basislerner verwendet, um durch die Kombination dieser eine verbesserte und robustere Klassifikation zu ermöglichen. Die einzige Voraussetzung an die zu kombinierenden Modelle ist eine jeweils höhere Genauigkeit als zufälliges Raten (Dietterich 2000, S. 1).

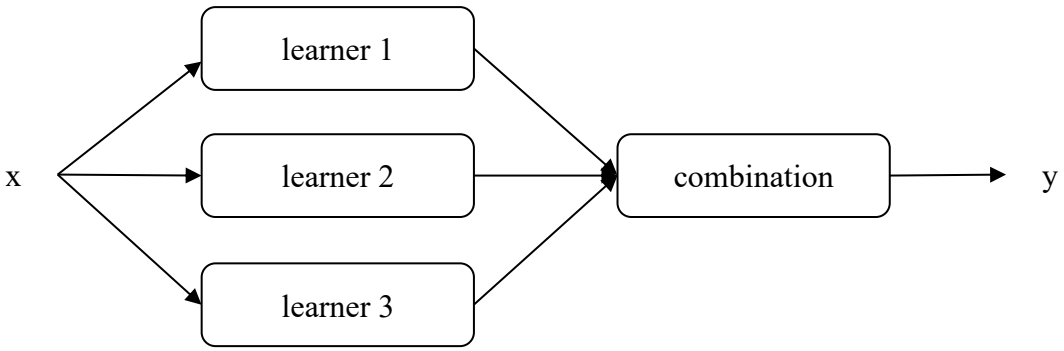

Abbildung 5-2: Aufbau einer Ensemblearchitektur i.A.a. (Zhou 2012, S. 16)

Für die Konstruktionen eines Ensembleverfahrens gibt es verschiedene Techniken. Zum einen die sequentiellen Ensembleverfahren, in denen die Basismethoden sequenziell erstellt werden (Boosting) und parallele Ensembleverfahren, in denen die Basismethoden parallel erstellt werden (Bagging) (Rokach 2010, S. 4 ff.).

Boosting Modelle erstellen ein kombiniertes Modell, indem jeder Basislerner sich stärker auf die richtige Klassifizierung der von vorherigen Basislernern falsch klassifizierten Daten konzentriert (Zhou 2012, S. 24). Bagging (bootstrap aggregating) Modelle hingegen erstellen mehrere Basislerner, indem Teile vom Trainingssatz durch zufällige Auswahl (mit zurücklegen) gezogen werden. Im Anschluss wird auf jeden dieser Teile ein eigener Basislerner trainiert (Rokach 2010, S. 11). Wobei diese Arbeit mit LightGBM ein Boosting und mit Random Forest ein Bagging-Verfahren abdeckt.

LightGBM

LightGBM (Light Gradient Boosting Machine) ist ein Machine Learning Framework, welches auf dem *Gradient-Boosting* Verfahren basiert und von Microsoft entwickelt wurde. Ein Hauptmerkmal von LightGBM ist seine Fähigkeit schnell mit großen Datenmengen umzugehen. Der Algorithmus basiert auf zwei Techniken. Gradientenbasiertes *One-Side Sampling* (GOSS) und *Exklusive Merkmalsbündelung* (EFB). GOSS sorgt dafür, dass ein signifikanter Anteil der Datensatzinstanzen mit kleinen Gradienten ausgeschlossen wird. Dadurch wird nur der Rest verwendet, um den Informationsgewinn zu schätzen. EFB wird verwendet, um die Anzahl der Merkmale zu reduzieren. Dafür bündelt EFB sich gegenseitig ausschließende Merkmale, das sind Merkmale die nur selten gleichzeitig Werte ungleich Null annehmen. (Rufo et al. 2021, S. 7 f.; Korstanje 2021, S. 196)

Random Forest

Random Forest (Breiman 2001) ist ein Ensembleverfahren und basierend auf der *Bagging-Methode*. Es verwendet mehrere Entscheidungsbäume als Basislerner, welche jeweils unabhängig voneinander auf einer zufällig gewählten Stichprobe der Trainingsdaten (mit Zurücklegen) trainiert werden. Beim Erstellen der einzelnen Entscheidungsbäume wird weiterhin bei jedem Entscheidungsknoten eine zufällige Teilmenge aus Merkmalen gewählt, auf denen ein Split durchgeführt werden kann. (Zhou 2012, S. 57 f.)

Die Wahl der finalen Entscheidung des Modells wird anders als bei LightGBM mit einem einfachen Mehrheitsvotum entschieden. Es wird somit die Klasse gewählt, die von den meisten Bäumen vorhergesagt wird (Breiman 2001, S. 6).

5.3 Implementierung und Evaluierung der Klassifizierungsverfahren

In diesem Kapitel wird die Implementierung und Evaluierung von den Klassifizierungsmethoden auf den synthetischen Daten beschrieben. Im ersten Schritt werden die verwendeten Tools vorgestellt. Im Anschluss werden die Datenaufbereitung und Vorverarbeitungsschritte beschrieben, welche notwendig für die weitere Anwendung und Auswertung sind. Nachfolgend werden die relevanten Metriken erläutert, welche zur Bewertung der Klassifizierungsverfahren angewandt werden. Abschließend werden für Klassifizierungsverfahren die Hyperparameter definiert, die Ergebnisse dargestellt sowie interpretiert.

5.3.1 Verwendete Tools und Datenvorverarbeitung

In der vorliegenden Arbeit wurde als Programmiersprache Python verwendet, um die Machine Learning Modelle zu implementieren und zu evaluieren. Python wurde verwendet, da es besonders im Bereich der Datenanalyse eine sehr weitverbreitete Programmiersprache darstellt und eine große Anzahl an Bibliotheken unterstützt, die den Einsatz verschiedener Klassifizierungsverfahren und deren Evaluierung vereinfachen (Robinson 2017; Hao und Ho 2019, S. 352). Um die einzelnen Modelle auf den vorliegenden Daten trainieren zu können, wurden die Daten vorverarbeitet. Die verwendeten Schritte werden im Folgenden erläutert.

Umwandlung des Fehlertyps in Dummy-Variablen

Für die ausgewählten Klassifizierungsverfahren wurden kategoriale Variablen in Dummy-Variablen (auch One-Hot-Encoding genannt) konvertiert. Da die Implementierung von Entscheidungsbäumen mittels der Scikit-learn Bibliothek zum aktuellen Zeitpunkt keine kategorialen Variablen als Eingabe unterstützt (scikit-learn 2025a). Dies wurde für den Systemfaktor Fehlertyp, welcher die vorliegende Prognosefehlerart repräsentiert, entsprechend kodiert.

Aufteilung der Daten in Trainings-, Validierungs- und Testsets

Um die Leistung der Klassifizierungsverfahren beurteilen zu können, werden die Daten in Trainings-, Validierungs- und Testsets aufgeteilt. Das Trainingsset wird verwendet, um die Modelle zu trainieren, während das Validierungsset dazu dient, die Hyperparameter der Modelle zu optimieren, indem die Leistung unabhängig von den Trainingsdaten überprüft werden kann. Letztendlich wird das Testset dazu verwendet, die Leistungsfähigkeit

der trainierten Modelle bei der Vorhersage mithilfe der optimierten Hyperparameter unabhängig zu bewerten. (Witten et al. 2017, S. 222)

Die Optimierung der Modellparameter erfolgt in der vorliegenden Arbeit unter Verwendung einer **k-Fold Cross-Validation**. Es wird hierbei ein 75:25 Split verwendet, bei dem 75 % der Daten für das Training und 25 % für die Validierung angewendet werden. Bei der k-Fold Cross-Validation wird das Datenset in k gleich große Teile („Folds“) aufgeteilt. Innerhalb von k Iterationen wird jeweils ein Fold als Validierungsset und die restlichen $k - 1$ Folds als Trainingsset verwendet. Die einzelnen Ergebnisse werden dann am Schluss gemittelt. Die Verwendung von k-Fold Cross-Validation ermöglicht hierbei eine robustere Schätzung der Modellleistung, die weniger stark von der Verteilung der zufälligen gewählten Daten abhängt. (Witten et al. 2017, S. 167) Die nachfolgende Abbildung 5-3 visualisiert das Vorgehen bei einer k-Fold Cross-Validation.

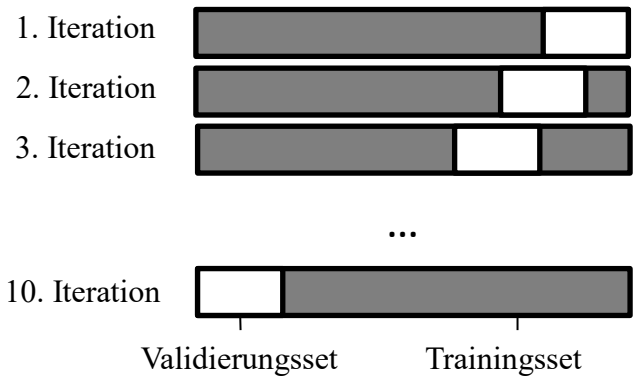

Abbildung 5-3: k-Fold Cross-Validation

Sowohl bei der initialen Aufteilung in Trainings- und Validierungsset, als auch bei der k-Fold Cross-Validation wird *Stratifizierung* eingesetzt. Dieses Vorgehen sorgt dafür, dass das Verhältnis der Heuristiken als beste Heuristik zwischen dem Trainings- und Validierungsset jeweils gleichbleibt.

Nach WITTEN et al. ist die Verwendung einer stratifizierten 10-fold Cross-Validation optimal (Witten et al. 2017, S. 168). Für die finale Auswertung wird weiterhin ein separates Testset bereitgehalten, welches weder in das Training, noch in die Suche der Hyperparameter einfließt.

5.3.2 Bewertungskriterien und Metriken

In diesem Kapitel werden die zentralen Metriken vorgestellt, welche zur Bewertung der Klassifizierungsverfahren verwendet werden. Dabei werden die Metriken beschrieben,

welche häufig in der Literatur für Mehrklassenklassifikation verwendet werden, sowie individuelle Metriken, die speziell für das untersuchte Problem entwickelt wurden.

Die **Genauigkeit** misst den Anteil der korrekt klassifizierten Beispiele an der Gesamtzahl der Beispiele. Sie ist eine grundlegende und am häufigsten verwendete Metriken für die Klassifizierung in der Literatur (Sokolova und Lapalme 2009, S. 428). Jedoch ist die Genauigkeit bei unausgeglichenen Klassen, wie im vorliegenden Fall der Auswahl von Heuristiken häufig irreführend (Gu et al. 2009, S. 463; Japkowicz 2011, S. 11), wodurch in der vorliegenden Arbeit von der Verwendung abgesehen wurde.

Häufig verwendete Metriken für die Klassifizierung bei unbalancierten Datensets sind die Präzision (Precision), die Sensitivität (Recall) und der F1-Score (Gu et al. 2009, S. 469). Diese Metriken basieren auf der **binären Konfusionsmatrix** und werden für jede Heuristik einzeln bestimmt. Eine beispielhafte Konfusionsmatrix zeigt die Abbildung 5-4. Diese stellt in den Zeilen die Vorhersage und in den Spalten die tatsächliche Klasse dar. Für jede der untersuchten Heuristiken wird individuell eine Konfusionsmatrix ermittelt, wobei *Positiv* die richtige Auswahl der jeweiligen Heuristik und *Negativ* die Auswahl einer alternativen Heuristik repräsentiert. (Gu et al. 2009, S. 462).

		Klassifikation: Positiv	Klassifikation: Negativ	
Tatsächlich	Positiv	Richtig Positiv: RP	Falsch Negativ: FN	Tats. Anzahl gut (RP+FN)
	Negativ	Falsch Positiv: FP	Richtig Negativ: RN	Tats. Anzahl schlecht (FP+RN)
		Vorhersage Anzahl gut (RP+FP)	Vorhersage schlecht (FN+RN)	

Abbildung 5-4: Binären Konfusionsmatrix i.A.a. (Schulz et al. 2022, S. 36)

Aus der Konfusionsmatrix lassen sich nun die Präzision und die Sensitivität herleiten. Die **Präzision** misst, wie gut das Modell positive Fälle identifiziert. Indem der Anteil der korrekt klassifizierten positiven Fälle (RP) zu allen vom Modell als positiv identifizierten Fällen (RP + FP) dargestellt wird (Maimon und Rokach 2005, S. 840). Somit ist die Präzision ein Maß dafür, wie korrekt das Modell ist, wenn es einen Datensatz als positiv einstuft.

$$Präzision = \frac{RP}{(RP + FP)} \tag{5-1}$$

Bei der **Sensitivität** wird nicht der Anteil an allen als positiv identifizierten Fällen, sondern der Anteil an den tatsächlich positiven Fällen (TP + FN) verwendet (Maimon und Rokach 2005, S. 840). Dies kann als Maß dafür angesehen werden, wie gut das Modell tatsächlich positive Fälle auch als solche erkennt.

$$Sensitivität = \frac{RP}{(RP + FN)} \tag{5-2}$$

Die Ziele der Präzision und der Sensitivität stehen meistens im Konflikt miteinander (Maimon und Rokach 2005, S. 840). Um ein Gleichgewicht zwischen beiden Metriken zu bestimmen, wird meist der **F1-Score** verwendet. Dieser stellt das harmonische Mittel von Präzision und Sensitivität dar und ermöglicht somit einen Kompromiss zwischen beiden Metriken (Gu et al. 2009, S. 469).

$$F1 = \frac{2 * \text{Präzision} * Sensivität}{Präzision + \text{Sensitivität}} \tag{5-3}$$

Cohens Kappa (Cohen 1960, S. 37 ff.) wurde ursprünglich entwickelt, um den Grad der Übereinstimmung zwischen zwei Beobachtern zu messen, die das gleiche Phänomen wahrnehmen (Tallón-Ballesteros und Riquelme 2014, S. 418). Cohens Kappa kann allerdings auch als Metrik für die Klassifizierung verwendet werden, indem die Vorhersage des Modells als der erste Beobachter und die tatsächlichen Werte als der zweite Beobachter gesetzt werden. Als Metrik ist Cohens Kappa gut geeignet, da es auch den Einfluss der rein zufälligen Wahl der richtigen Heuristik einbezieht (Tallón-Ballesteros und Riquelme 2014, S. 418). Cohens Kappa ist somit eine Metrik, die die Übereinstimmung zwischen zwei Schätzern (oder in diesem Fall, zwischen der Vorhersage des Modells und den tatsächlichen Werten) misst und dabei den Zufallseinfluss berücksichtigt. Berechnet wird Cohens Kappa nach der nachfolgenden Gleichung.

$$\kappa = \frac{P(A) - P(E)}{1 - P(E)} \tag{5-4}$$

Hierbei handelt es sich bei $P(A)$ um den Anteil der Fälle, in denen die Vorhersagen des Modells mit den tatsächlichen Werten übereinstimmen und bei $P(E)$ um den bei Zufall erwarteten Anteil. Genauere Details zur Berechnung lassen sich beispielsweise in (Tallón-Ballesteros und Riquelme 2014, S. 419) finden. Der Kappa-Wert variiert dabei zwischen -1 und 1, wobei eine 1 einer perfekten Übereinstimmung und eine 0 einer Übereinstimmung aufgrund von Zufall entspricht. Wohingegen negative Werte eine Übereinstimmung schlechter als der Zufall entsprechen (Carletta und Hirschberg 1996, S. 252). Für die Werte zwischen 0 und 1 gilt folgende Unterteilung hinsichtlich des Zusammenhangs (Grandini et al. 2020, S. 14):

- leichte Übereinstimmung: 0,1 bis 0,2
- faire Übereinstimmung: 0,21 bis 0,4
- moderate Übereinstimmung: 0,41 bis 0,6
- substanzielle Übereinstimmung: 0,61 bis 0,8
- perfekte Übereinstimmung: 0.81 bis 0.99

Als Nächstes wird die **durchschnittliche Abweichung** der jeweils vorhergesagten Heuristik zur besten Heuristik sowie die **Standardabweichung** der Abweichungen ausgewählt. Für den Fall, dass die kostenminimale Heuristik vorausgesagt wurde, wird der Wert 0 als Abweichung definiert. Damit soll die Metrik die Gesamteffektivität der Vorhersage berücksichtigen und nicht durch eine hohe Genauigkeit, aber dafür wenige Ausnahmen mit größeren Abweichungen, beeinflusst werden. Eine niedrigere durchschnittliche Abweichung weist darauf hin, dass das Modell effektiver darin ist, Heuristiken zu identifizieren, die nahe an der besten Heuristik liegen. Die Standardabweichung der Abweichungen gibt Aufschluss über die Streuung der Abweichungen und damit über die Konsistenz der Modellvorhersagen. Eine niedrigere Standardabweichung deutet auf eine höhere Konsistenz der Vorhersagen hin, da die Abweichungen enger um den Durchschnitt liegen.

Diese Metrik spiegelt die tatsächlichen Kostenabweichungen der Modellentscheidungen wider. Ferner trägt die Einbeziehung der Standardabweichung dazu bei, Modelle zu identifizieren, die nicht nur im Durchschnitt gute Vorhersagen liefern, sondern auch konsistent in ihrer Leistung sind. Dies ist in dem Anwendungsfall dieser Arbeit von besonderer Bedeutung, da die Zuverlässigkeit der Vorhersagen eine entscheidende Rolle für die Akzeptanz des Modells spielt.

Die Metrik **Anteil Top Heuristik** berechnet den Anteil der richtig vorhergesagten Heuristiken, deren Kostenabweichung zur kostenoptimalen Heuristik innerhalb des definierten Prozentsatzes (Schwellwert) liegen. Die Heuristiken, die innerhalb dieses Schwellenwertes liegen, werden als *Top* betrachtet. Diese Metrik ermöglicht es, die Fähigkeit des

Modells zu bewerten, Heuristiken zu identifizieren, die nahe an der optimalen Lösung liegen. Als Schwellenwert wurde hierbei 1 % gewählt. Dies entspricht dem gewählten Schwellenwert aus Kapitel 4.3.

5.3.3 Methodik und Implementierung

Viele Klassifizierungsalgorithmen haben Parameter, welche vor dem Trainingsprozess definiert werden müssen. Diese Parameter werden auch Hyperparameter genannt. Die Hyperparameter müssen so gewählt werden, um eine bestimmte Zielmetrik zu maximieren. Dieser Prozess wird als Hyperparameteroptimierung bezeichnet. (Bergstra und Bengio 2012, S. 282). Für die Hyperparameteroptimierung werden häufig die Verfahren Grid Search oder Random Search verwendet.

Grid Search ist eine systematische Methode zur Hyperparameteroptimierung, bei der ein Suchraum mit expliziten Werten für jeden Hyperparameter definiert wird. Im Anschluss werden alle möglichen Kombinationen der Hyperparameterwerte innerhalb dieses Suchraums ausprobiert, um die beste Kombination in Bezug auf die Modellleistung zu finden. (Bergstra und Bengio 2012, S. 282)

Demgegenüber ist die Methode **Random Search** eine alternative Methode zur Hyperparameteroptimierung, bei der zufällige Kombinationen von Hyperparameterwerten innerhalb eines vorgegebenen Suchraums ausprobiert werden. Im Gegensatz zu Grid Search, bei der alle möglichen Kombinationen ausprobiert werden, werden bei Random Search nur eine begrenzte Anzahl von zufällig ausgewählten Kombinationen getestet. BERGSTRA und BENGIO zeigt auf, dass Random Search in vielen Fällen eine effizientere und schnellere Methode zur Hyperparameteroptimierung darstellt als ein Grid Search. (Bergstra und Bengio 2012, S. 281 ff.) Auf der Grundlage der guten Ergebnisse sowie des geringen Rechenaufwands, wurde für diese Arbeit die Methode Random Search als Hyperparameteroptimierung verwendet.

Zur Implementierung von Random Search wird die Bibliothek Scikit-learn verwendet. Als Metrik für die Hyperparameteroptimierung wird die individuelle Metrik Anteil der Top-Heuristiken (ATH) als Zielgröße verwendet. Demzufolge ist es das Ziel, eine Heuristik zu finden, welche eine maximale Kostenabweichung von 1 % (Schwellwert) zur kostenoptimalen Heuristik hat. Damit soll erreicht werden, dass das ML-Modell vor allem die Faktorkombinationen erkennt und lernt, in denen große Kostenunterschiede zwischen den vier Heuristiken vorhandenen sind. Die weiteren beschriebenen Metriken (vgl. Kapitel 5.3.2) werden später beim Vergleich der drei ML-Modelle (vgl. Kapitel 5.4) berücksichtigt, um eine tiefergehende Leistungsanalyse durchzuführen.

Nachdem die besten Hyperparameter für jedes Verfahren gefunden wurden, wird die Leistung jedes Klassifikators auf einem separaten Testset validiert, welches nicht für die

Hyperparameterwahl und das Trainieren des ML-Modells verwendet wurde (vgl. Kapitel 5.3.1). Auf dem separaten Testset werden die im Kapitel 5.3.2 jeweils beschriebenen Metriken ausgewertet. Weiterhin wird das Testset nach Fehlertyp aufgeteilt und die Metriken auf diesen Teilmengen auch ausgewertet. Dies ermöglicht die Interpretation der Leistung der Klassifizierungsmodelle bei verschiedenen Fehlertypen.

Hyperparameter: Entscheidungsbaum

Für den Entscheidungsbaum wurde die Klasse DecisionTreeClassifier der Bibliothek Scikitlearn verwendet. Dabei handelt es sich um eine optimierte Version des CART-Algorithmus (vgl. Kapitel 5.2.1). Die Leistung des Entscheidungsbaumes kann durch die Optimierung verschiedener Hyperparameter beeinflusst werden. Die wichtigsten Hyperparameter für Entscheidungsbäume sind in Scikit-learn (scikit-learn 2025a):

- criterion: Die Funktion, welche die Qualität einer Aufteilung bewertet. Mögliche Werte sind „gini“ für den Gini-Index und „entropy“ für den Informationsgewinn.
- max_depth: Die maximale Tiefe des Baums. Tiefergehende Bäume können komplexere Strukturen abbilden, sind aber anfälliger für Overfitting.
- min_samples_split: Die minimale Anzahl von Datenpunkten, die erforderlich sind, um einen internen Knoten weiter aufzuteilen.
- min_samples_leaf: Die minimale Anzahl von Datenpunkten, die in einem Blattknoten enthalten sein müssen.
- ccp_alpha: Komplexitätsparameter für das Cost-Complexity Pruning.

In Tabelle 5-1 sind die Hyperparameter für den Entscheidungsbaum ersichtlich, welche sich durch die Verwendung der Methode Random Search ergeben haben.

Tabelle 5-1: Hyperparameter - Entscheidungsbaum

Parameter	**Wert**
criterion	entropy
max_depth	19
min_samples_split	3
min_samples_leaf	15
ccp_alpha	$2.4043733078091174\ e\text{-}05^{-5}$

Hyperparameter: LightGBM

Für den LightGBM wurde die Klasse LGBMClassifier verwendet. Auch die Leistung des LightGBM Verfahrens hängt von den definierten Hyperparametern ab. Nachfolgend werden die zentralen Hyperparameter erläutert (scikit-learn 2025b):

- learning_rate: Die Lernrate, die den Einfluss der schwachen Lernmodelle in der endgültigen Kombination steuert. Ein niedriger Wert führt zu einer langsameren Anpassung, was möglicherweise zu einem besseren Modell führt, aber mehr Iterationen erfordert
- max_depth: Die maximale Tiefe des Baums. Tiefergehende Bäume können komplexere Strukturen abbilden, sind aber anfälliger für Overfitting
- n_estimators: Die Anzahl der Basis-Klassifikatoren, die im Ensemble verwendet werden. Eine höhere Anzahl von Klassifikatoren kann zu einer besseren Leistung führen, aber auch das Risiko von Overfitting erhöhen und die Trainingszeit verlängern.

In Tabelle 5-2 sind die Hyperparameter für das Verfahren LightGBM ersichtlich, welche sich durch die Verwendung der Methode Random Search ergeben haben.

Tabelle 5-2: Hyperparameter - LightGBM

Parameter	**Wert**
learning_rate	0.10310149462998894
max_depth	14
n_estimators	243

Hyperparameter: Random Forest

Für den Random Forest wurde die Klasse RandomForestClassifier der Bibliothek Scikit-learn verwendet. Auch Random Forest besitzt Hyperparameter, welche die Leistung des Modells beeinflussen. Betrachtet wurden hierbei folgende Hyperparameter (scikit-learn 2025b):

- n_estimators: Die Anzahl der Bäume. Eine größere Anzahl von Bäumen führt in der Regel zu einer besseren Leistung, kann aber auch die Trainingszeit verlängern und das Risiko von Overfitting erhöhen.
- max_depth: Die maximale Tiefe der Bäume. Tiefergehende Bäume können komplexere Strukturen abbilden, sind aber anfälliger für Overfitting.

- bootstrap: Gibt an, ob beim Erstellen der Bäume Bootstrap-Stichproben (mit Zurücklegen) verwendet werden sollen oder ob das gesamte Datenset für jeden Baum verwendet wird.

Aus der Hyperparameteroptimierung haben sich für den Random Forest die Hyperparameter in Tabelle 5-3 ergeben.

Tabelle 5-3: Hyperparameter - Random Forest

Parameter	Wert
n_estimators	82
max_depth	11
bootstrap	False

5.4 Vergleich der Klassifizierungsmethoden

In diesem Kapitel erfolgt der Vergleich der drei Klassifizierungsverfahren, auf Basis des Testsets. Dabei werden die zuvor ermittelten Hyperparameter (vgl. 5.3.3) und beschriebenen Metriken (vgl. 5.3.2) verwendet. Die Ergebnisse der drei Verfahren sind nach den unterschiedlichen Metriken in Tabelle 5-4 dargestellt. Um die Lesbarkeit der Tabelle zu verbessern, werden in den folgenden Tabellen Abkürzungen für die verschiedenen Metriken verwendet Precision (P), Recall (R), F1-Score (F1), Cohens Kappa (CK), durchschnittliche Abweichung (avgA), Standardabweichung (stdA) und Anteil Top-Heuristiken (ATH).

Tabelle 5-4: Gesamtpopulation - Auswertung der Klassifizierungsmodelle mittels des Testsets

Methode	P	R	F1	CK	avgA	stdA	ATH
Entscheidungsbaum	0,645	0,666	0,627	0,414	0,0106	0,035	81,86%
LightGBM	0,645	0,667	0,628	0,415	0,0103	0,0326	81,94%
Random Forest	0,645	0,667	0,625	0,411	0,0105	0,036	81,93%

Im Allgemeinen lässt sich feststellen, dass alle drei ML-Modelle LightGBM, Random Forest und Entscheidungsbaum ähnlich gute Ergebnisse erzielen. Hinsichtlich aller Metriken weist der LightGBM die beste Leistung auf. Mit einem Cohens Kappa von 0,415

lässt sich ein Zusammenhang zwischen den Vorhersagen des Modells und den tatsächlichen Klassifikationen feststellen. Die durchschnittliche Abweichung von ca. 1,01 % und die Standardabweichung von 3,26 % weisen weiterhin darauf hin, dass das LightGBM-Modell im Allgemeinen zuverlässig die kostenoptimale oder eine nahezu kostenoptimale Heuristik auswählt. Das spiegelt sich auch in dem Anteil der Top-Heuristiken (ATH = 81,94 %) mit einen 1 % Schwellenwert wider. Größere Abweichungen zwischen den ML-Modellen sind im Bereich der Standardabweichung zu erkennen. Hier zeigt das LightGBM-Modell eine um 0,3 Prozentpunkte bessere Leistung gegenüber den anderen beiden Verfahren.

Um die Leistung bzw. Vorteilhaftigkeit des LightGBMs genauer beurteilen zu können, wird ein Leistungsvergleich zwischen der Heuristikwahl mittels *LightGBM, der durchschnittlich besten Heuristik* (vgl. Kapitel 4), dem PP-Verfahren und einer *zufälligen Auswahl* der vier Heuristiken durchgeführt. Bei der zufälligen Auswahl der Heuristiken werden die Abweichungen durch den Mittelwert aller vier Heuristiken berechnet. Der Vergleich erfolgt anhand der Zielgrößen Anteil der Top-Heuristiken, durchschnittliche Abweichung und Standardabweichung. Die Ergebnisse der drei Handlungsalternativen sind nach den drei Zielgrößen in Tabelle 5-5 dargestellt.

Tabelle 5-5: Gesamtpopulation - Leistungsvergleich unterschiedlicher Handlungsalternativen gegenüber der Heuristikwahl mittels LightGBM

Methodik	**avgA**	**stdA**	**ATH**
Zufällige Auswahl	0,047	0,068	27,50%
Nur PP	0,020	0,060	70,49%
LightGBM	0,010	0,033	81,94%

Es zeigt sich, dass das Modell LightGBM bei allen drei Zielgrößen bessere Ergebnisse liefert, als bei der dauerhaften Nutzung von PP sowie bei einer zufälligen Heuristikwahl. Wie in Tabelle 5-5 ersichtlich, können durchschnittlich ca. 3,7 Prozentpunkte der Bestellmengenkosten durch die Verwendung des LightGBM Verfahrens eingespart werden. Im Vergleich zur konsequenten Nutzung des PP-Verfahrens kann 1 Prozentpunkt der Kosten durch die Verwendung des LightGBM Verfahren eingespart werden. Ein ähnliches Bild zeigt die Standardabweichung der drei Handlungsalternativen auf, zufällige Auswahl (6,8 %), nur PP (6 %) und LightGBM (3,3 %). Auch hier ist die Genauigkeit der Heuristikwahl mittels des Verfahrens LightGBM vorteilhaft. Dabei wird die Zielgröße ATH im Falle einer zufälligen Auswahl durch die Wahrscheinlichkeit bestimmt, dass eine der vier Heu-

ristiken die kostengünstigste ist (25 %). Zusätzlich werden jene Simulationsläufe berücksichtigt, in denen alle vier Heuristiken eine Kostenabweichung von weniger als 1 % (Schwellwert) aufweisen (ATH = 27,5 %).

Zusammenfassend kann gesagt werden, dass auf Basis der synthetischen Daten (Gesamtpopulation) die Heuristikwahl mittels dem LightGBM deutliche Kosteneinsparungen gegenüber einer zufälligen Heuristikwahl oder auch gegenüber der konsequenten Nutzung des PP-Verfahrens erzielt werden können. In den folgenden Abschnitten werden die Leistungen der Modelle tiefgehender untersucht, indem die Modelle für jeden Fehlertyp einzeln ausgewertet werden.

Rauschen

Die Tabelle 5-6 zeigt die Ergebnisse der drei Klassifizierungsverfahren in Abhängigkeit der Zielgrößen bei vorliegenden Fehlertyp Rauschen.

Tabelle 5-6: Rauschen - Auswertung der Klassifizierungsmodelle mittels des Testsets

Methode	P	R	F1	CK	avgA	stdA	ATH
Entscheidungsbaum	0,544	0,637	0,578	0,349	0,0195	0,0519	71,14%
LightGBM	0,575	0,639	0,583	0,359	0,019	0,047	71,24%
Random Forest	0,572	0,638	0,579	0,350	0,0194	0,0539	71,31%

Im Allgemeinen sind die Ergebnisse beim vorliegenden Fehlertyp Rauschen schlechter als bei der Gesamtpopulation. Darüber hinaus bestehen nur marginale Leistungsunterschiede zwischen den drei Verfahren. Die Rangfolge bleibt jedoch weiterhin bestehen, sodass unter Berücksichtigung aller Zielgrößen das LightGBM Verfahren die besten Ergebnisse aufweist. Nur im Bereich der Zielgröße ATH weist das Verfahren LightGBM (ATH = 71,24 %) leicht schlechtere Ergebnisse als das Verfahren Random Forest (ATH = 71,31 %) auf.

Mit einem Cohens Kappa von unterhalb 0,4 lässt sich eine faire Übereinstimmung zwischen der Modellvorhersage und der tatsächlich kostengünstigsten Heuristik erkennen. Ebenso ist ersichtlich, dass die Standardabweichung mit ca. 5 % stärkere Abweichung aufweist als in der Gesamtpopulation. Insgesamt kann jedoch gesagt werden, dass mit allen drei Klassifizierungsverfahren in den meisten Fällen die kostenoptimale oder nahezu kostenoptimale Heuristik vorhergesagt werden kann (ATH = ca. 71 %).

Für die nachfolgenden Analysen wird nur das LightGBM Verfahren berücksichtigt, da mit diesem Verfahren unter Beachtung aller Zielgrößen die besten Ergebnisse erzielt werden. Wie bereits bei der Gesamtpopulation werden nachfolgend die drei **Handlungsalternativen** *zufällige Auswahl, konsequente Nutzung der PP-Heuristik* und *Auswahl mittels des LightGBM-Verfahrens* vergleichen.

Tabelle 5-7: Rauschen - Leistungsvergleich unterschiedlicher Handlungsalternativen gegenüber der Heuristikwahl mittels LightGBM

Methodik	**avgA**	**stdA**	**ATH**
Zufällige Auswahl	0,062	0,086	26,32%
Nur PP	0,036	0,094	60,81%
LightGBM	0,019	0,047	71,24%

Der Vergleich der drei Handlungsalternativen ist in Tabelle 5-7 dargestellt. Das LightGBM-Verfahren liefert deutlich bessere Ergebnisse als sowohl die dauerhafte Nutzung des PP-Verfahrens als auch die zufällige Heuristikwahl. Durch die Anwendung des LightGBM zur Heuristikwahl können durchschnittlich 1,7 Prozentpunkte gegenüber der Verwendung des PP-Verfahrens und 4,3 Prozentpunkte gegenüber der zufälligen Auswahl eingespart werden. Gleichzeitig sinkt die Standardabweichung von 8,6 % (zufällige Auswahl) bzw. 9,4 % (nur PP-Verfahren) auf 4,7 % (LightGBM), was auf robuste Entscheidungen hinweist. Diese Ergebnisse werden ebenso durch die Zielgröße ATH bestätigt, sodass mit einer Genauigkeit ca. 71 % die kostenoptimale bzw. nahezu kostenoptimale Heuristik ausgewählt wird. Somit stellt LightGBM eine klar überlegene Methode zur Heuristikwahl dar.

Unterschätzung

Die Tabelle 5-8 zeigt die Ergebnisse der drei Klassifizierungsmethoden in Abhängigkeit der Zielgrößen bei vorliegenden Fehlertyp Unterschätzung.

Tabelle 5-8: Unterschätzung - Auswertung der Klassifizierungsmodelle mittels des Testsets

Methode	P	R	F1	CK	avgA	stdA	ATH
Entschei-dungsbaum	0,709	0,719	0,701	0,481	0,0102	0,0300	80,12%
LightGBM	0,710	0,720	0,702	0,483	0,010	0,0287	80,27%
Random Forest	0,708	0,720	0,700	0,480	0,0101	0,0288	80,17%

Auch beim Fehlertyp Unterschätzung weisen alle drei Klassifizierungsverfahren nur geringfügige Unterschiede auf. Über alle Metriken hinweg zeigt das LightGBM-Verfahren die besten Resultate und erreicht mit einem Cohens Kappa von 0,483 eine moderate Korrelation zwischen der Modellvorhersage und der tatsächlich kostenoptimalen Heuristik. Die durchschnittliche Abweichung (avgA = 1,0 %) und die Standardabweichung (stdA = 2,87 %) zeigen, dass das LightGBM in der Lage ist, in den meisten Fällen die kostenoptimale oder nahezu optimale Heuristik korrekt vorherzusagen. Dies wird durch den hohen Anteil der Top-Heuristiken (ATH = 80,27 %) bestätigt. Im Vergleich zum Fehlertypen Rauschen ist es bei vorliegender Unterschätzung möglich, präziser und besser die kostenoptimale Heuristik vorherzusagen.

Wie bereits bei der Gesamtpopulation und dem Fehlertyp Rauschen wird für die nachfolgenden Analysen nur das LightGBM Verfahren berücksichtigt, da dieses Verfahren die besten Ergebnisse erzielt. Nachfolgend werden die drei **Handlungsalternativen** *zufällige Auswahl*, *konsequente Nutzung der PP-Heuristik* und *Auswahl mittels des LightGBM-Verfahrens* vergleichen.

Tabelle 5-9: Unterschätzung - Leistungsvergleich unterschiedlicher Handlungsalternativen gegenüber der Heuristikwahl mittels LightGBM

Methodik	avgA	stdA	ATH
Zufällige Auswahl	0,060	0,059	26,20%
Nur PP	0,022	0,039	58,92%
LightGBM	0.010	0,029	80,27%

Die Tabelle 5-9 zeigt, dass das LightGBM-Verfahren deutlich bessere Ergebnisse liefert als die anderen beiden Handlungsalternativen. Durch die Anwendung des LightGBM zur Heuristikwahl können durchschnittlich 1,2 Prozentpunkte gegenüber der Verwendung des PP-Verfahrens und 5 Prozentpunkte gegenüber der zufälligen Auswahl eingespart

werden. Gleichzeitig sinkt die Standardabweichung von 5,9 % (zufällige Auswahl) bzw. 3,9 % (nur PP-Verfahren) auf 2,9 % (LightGBM), was auf robuste Entscheidungen hinweist. Diese Ergebnisse werden ebenso durch die hohen Werte der Zielgröße ATH bestätigt, sodass mit einer Genauigkeit von über 80 % die kostenoptimale bzw. nahezu kostenoptimale Heuristik ausgewählt wird. Somit stellt LightGBM eine klar überlegene Methode zur Heuristikwahl dar.

Überschätzung

Die Tabelle 5-10 zeigt die Ergebnisse der drei Klassifizierungsmethoden in Abhängigkeit der Zielgrößen bei vorliegenden Fehlertyp Überschätzung.

Tabelle 5-10: Überschätzung - Auswertung der Klassifizierungsmodelle mittels des Testsets

Methode	P	R	F1	CK	avgA	stdA	ATH
Entscheidungsbaum	0,606	0,643	0,579	0,211	0,0019	0,0039	94,37%
LightGBM	0,609	0,643	0,576	0,207	0,0019	0,0039	94,35%
Random Forest	0,608	0,642	0,574	0,204	0,0019	0,0039	94,34%

Alle drei Klassifizierungsverfahren LightGBM, Random Forest und Entscheidungsbaum weisen kaum Leistungsunterschiede auf. Vor allem im Bereich der Standardabweichung (ca. 0,04 %) und durchschnittlichen Abweichung (0,2 %) sind keine Leistungsunterschiede festzustellen. Die Korrelation zwischen der Modellvorhersage und der tatsächlich kostenoptimalen Heuristik mit einem Cohens Kappa von ca. 0,2 entspricht nur einer leichten Korrelation und ist deutlich geringer als bei den anderen Fehlertypen. Dennoch wird mit einem ATH über 94 % die kostenoptimale bzw. nahezu kostenoptimale Heuristik vorhergesagt. Auf Basis der geringen Korrelation und durchschnittlichen Abweichung in Kombination mit der guten Vorhersagbarkeit lässt sich erkennen, dass die Heuristiken bei vorliegender systematischer Überschätzung sehr ähnliche Werte aufweisen. Eine situative Auswahl in diesem Bereich birgt nur geringe Einsparungspotenziale.

Wie bereits bei den anderen Fehlertypen wird nachfolgend nur das LightGBM Verfahren berücksichtigt. Für die tiefergehende Analyse werden auch bei diesem Fehlertyp die drei **Handlungsalternativen** *zufällige Auswahl, konsequente Nutzung der PP-Heuristik* und *Auswahl mittels des LightGBM-Verfahrens* verglichen.

Tabelle 5-11: Überschätzung - Leistungsvergleich unterschiedlicher Handlungsalternativen gegenüber der Heuristikwahl mittels LightGBM

Methodik	**avgA**	**stdA**	**ATH**
Zufällige Auswahl	0,0202	0,0201	30,00%
Nur PP	0,0023	0,0044	92,58%
LightGBM	0,0019	0,0039	94,35%

Der Vergleich der Handlungsalternativen zeigt, dass die konsequente Nutzung des PP-Verfahrens bereits sehr gute Ergebnisse liefert (ATH = 92,58 %, avgA = 0,23 %). Dennoch kann das LightGBM-Verfahren die Ergebnisse nochmals geringfügig verbessern und erreicht nahezu perfekte Vorhersagen. Eine zufällige Auswahl der Heuristik ist dagegen, deutlich unterlegen, sowohl bei der Präzision der Vorhersage wie auch bei der durchschnittlichen Kostenabweichung (ca. 2 %). Insgesamt lassen sich die gute Prognostizierbarkeit sowie eine geringe durchschnittliche Kostenabweichung auf die geringen Leistungsunterschiede beim vorliegenden Fehlertyp Überschätzung zwischen den vier Heuristiken zurückführen.

Die Analyse der verschiedenen Klassifizierungsverfahren zeigt, dass alle drei untersuchten Verfahren Entscheidungsbaum, Random Forest und LightGBM, vergleichbare Ergebnisse liefern. Unter Berücksichtigung aller Zielgrößen erzielt das LightGBM-Verfahren die besten Ergebnisse. Im Hinblick auf die Gesamtpopulation kann LightGBM die Auswahl der kostenoptimalen oder nahezu optimalen Heuristik mit hoher Präzision vorhersagen. Dies führt zu einer durchschnittlichen Kostenreduktion von 3,7 Prozentpunkte im Vergleich zu einer zufälligen Heuristikwahl und von 1 Prozentpunkten gegenüber der konsequenten Nutzung der PP-Heuristik. Zudem weist LightGBM eine geringere Standardabweichung auf, was auf eine höhere Robustheit der Entscheidungen hinweist.

Die detaillierte Betrachtung der unterschiedlichen Fehlertypen verdeutlicht, dass die Vorhersagequalität und das Kosteneinsparungspotenzial stark von der Art des Prognosefehlers abhängen. Bei Unterschätzungen erreicht LightGBM eine hohe Genauigkeit mit einem Anteil der Top-Heuristiken von über 80 %, was auf eine stabile und verlässliche Klassifikation hinweist. Im Fall von Rauschen sinkt die Vorhersagegenauigkeit leicht auf ca. 71 %. Das Kosteneinsparungspotenzial ist ähnlich wie bei Fehlertyp Unterschätzung und liegt durchschnittlich zwischen 1 % und 5 % der gesamten Bestellmengenkosten. Bei Überschätzungen zeigt sich hingegen, dass die Heuristiken insgesamt sehr ähnliche Kosten aufweisen, wodurch die potenziellen Einsparungen infolge einer kontextspezifischen Heuristikwahl begrenzt sind. Hier liefert bereits die dauerhafte Nutzung der PP-Heuristik sehr gute Ergebnisse, die durch LightGBM nur geringfügig verbessert werden können.

Zusammenfassend lässt sich feststellen, dass der Einsatz von LightGBM einen deutlichen Mehrwert gegenüber einer zufälligen Auswahl oder gegenüber einer konsequenten Nutzung des PP-Verfahrens bei der kostenorientierten Heuristikwahl bietet, insbesondere bei den Fehlertypen *Unterschätzung* und *Rauschen*. Die Methode ermöglicht nicht nur eine signifikante Reduktion der Bestellmengenkosten, sondern sorgt zugleich für eine höhere Robustheit der Entscheidungsprozesse. Beim Fehlertypen *Überschätzung* ist der Nutzen einer ML-basierten Auswahl hingegen begrenzt, aufgrund der geringen Leistungsunterschiede zwischen den Heuristiken. Insgesamt zeigt diese Analyse, dass Verfahren des Machine Learnings ein effektives Instrument, zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren sein können.

Darüber hinaus ist es erforderlich, die getroffene **Heuristikwahl in regelmäßigen Abständen zu überprüfen**. Dabei ist es zu empfehlen, die Systemfaktoren der dynamischen Bestellmengenplanung fortlaufend zu erfassen und zu analysieren. Es ist bei Änderung einzelner Systemfaktoren zu empfehlen, die Heuristikwahl mithilfe des ML-Modells zu überprüfen. Von besonderer Bedeutung ist die frühzeitige Erkennung eines möglichen Wechsels des zugrunde liegenden Fehlertyps. Verbesserungen der Prognosegüte können ebenso wie die Wahl der jeweils richtigen Heuristik maßgeblich zur Reduktion der Gesamtkosten beitragen.

5.5 Zwischenfazit: Machine Learning Modell zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren

Das Kapitel 5 behandelt die Entwicklung eines datengetriebenen Machine Learning Modells zur kostenorientierten Auswahl dynamischer Bestellmengenverfahren. Ziel ist es, die kontextspezifischen Leistungsdifferenzen der unterschiedlichen Heuristiken systematisch zu erfassen, um eine kontextspezifische Auswahl der Heuristiken zu ermöglichen. In diesem Rahmen stellt sich die vierte Forschungsfrage: „*Wie kann ein Machine Learning Modell zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren gestaltet werden?*“

Zur Beantwortung dieser Forschungsfrage werden nach einer Einführung in die Grundlagen des Machine Learnings, die relevanten Machine Learning Modelle identifiziert. Dabei werden die drei baumbasierten Verfahren Entscheidungsbaum, Random Forest sowie LightGBM ausgewählt. Die Implementierung erfolgt in Python unter Verwendung der Bibliothek scikit-learn. Die Daten werden durch One-Hot-Encoding aufbereitet und in Trainings-, Validierungs- und Testsets aufgeteilt. Zur robusten Schätzung der Modellleistung wird eine stratifizierte 10-fold Cross-Validation eingesetzt. Als Optimierungsverfahren für die Hyperparameter wird die Methode Random Search genutzt, wobei die Ma-

ximierung von Anteil-Top-Heuristik (ATH) als Zielgröße dient. Weitere Bewertungsmetriken wie Präzision, Recall, F1-Score, Cohens Kappa, durchschnittliche Abweichung und Standardabweichung werden zur detaillierten Leistungsbewertung herangezogen.

Die Ergebnisse zeigen, dass alle drei Verfahren vergleichbare Vorhersageleistungen erzielen, wobei LightGBM in nahezu allen Metriken leicht überlegen ist. In der Gesamtpopulation können durch den Einsatz von LightGBM im Vergleich zu einer zufälligen Heuristikwahl durchschnittlich 3,7 Prozentpunkten der relevanten Kosten eingespart werden. Auch im Vergleich zur permanenten Nutzung der durchschnittlich besten Heuristik (PP) lassen sich Einsparungen von rund 1 Prozentpunkten erzielen. Bei den Fehlertypen Rauschen und Unterschätzungen können besonders hohe Kosteneinsparungen erzielt werden. Im Fall von Überschätzungen sind die Kostenunterschiede zwischen den Heuristiken gering, sodass die ML-basierte Auswahl nur einen marginalen Zusatznutzen gegenüber der konsequenten Nutzung der PP-Heuristik bietet.

Insgesamt zeigt das Kapitel, dass Machine Learning Modelle, insbesondere LightGBM, eine kostenorientierte Auswahl von dynamischen Bestellmengenverfahren ermöglichen können. In Kombination mit einer guten Datenbasis können deutliche Kosteneinsparung erzielt werden.

6 Validierung der Entscheidungsregel sowie des ML-Modells zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren

Dieses Kapitel widmet sich der Validierung der entwickelten Entscheidungsregeln sowie des trainierten ML-Modells zur kostenorientierten Auswahl dynamischer Bestellmengenverfahren. Zu diesem Zweck werden Simulationsexperimente auf Basis realer Unternehmensdaten durchgeführt. Für die kostenorientierte Auswahl werden die drei Handlungsalternativen *zufällige Auswahl*, *Auswahl mittels Entscheidungsregeln* und *Auswahl mittels des ML-Modells* miteinander verglichen und hinsichtlich der Anwendbarkeit und der Gesamtkosten bewertet.

Um die Übertragbarkeit der Forschungsergebnisse zusätzlich zu untersuchen, werden Datensätze aus zwei unterschiedlichen Unternehmen aus unterschiedlichen Branchen herangezogen. Beide Unternehmen bieten verschiedene Produkte an, die sich über unterschiedliche Faktorstufen der sechs untersuchten Systemfaktoren erstrecken. Dadurch kann eine breite Varianz relevanter Einflussgrößen abgebildet werden.

Bei den zwei Datensätzen handelt es sich teilweise um vertrauliche und unternehmensspezifische Informationen, daher wurden beide Datensätze vor der Analyse vollständig anonymisiert. Das erste Unternehmen handelt mit industriellen Gütern (Anwendungsfall 1). Wohingegen das zweite Unternehmen im Bereich des Handels von Gastronomiebedarfen (Anwendungsfall 2) tätig ist.

6.1 Simulationsexperimente: Anwendungsfall 1 – Handel von industriellen Gütern

Zunächst werden die Daten des ersten Anwendungsfalls hinsichtlich der sechs Systemfaktoren und deren Faktorstufen analysiert und beschrieben (vgl. Kapitel 6.1.1). Anschließend werden im Kapitel 6.1.2 die Simulationsexperimente erläutert sowie die verschiedenen Handlungsalternativen zur Heuristikwahl auf Basis der entstehenden Gesamtkosten analysiert und bewertet.

6.1.1 Vorstellung der Datengrundlage: Anwendungsfall 1 – Handel von industriellen Gütern

Der erste Anwendungsfall umfasst 121 unterschiedliche Produkte, für welche die Bestellmengen und -zeitpunkte geplant werden. Die verschiedenen Systemfaktoren und ihre jeweiligen Faktorstufen werden nachfolgend näher beschrieben.

Die historischen **Nachfragen** zu den 121 Produkten weisen eine Dauer zwischen 1.424 Perioden und 1.559 Perioden auf, sodass eine umfangreiche Nachfragehistorie vorliegt. Die unterschiedlichen Nachfrageverläufe werden hinsichtlich der zwei Eigenschaften Streuung des Bedarfs (*VarK²*) und Anzahl der Nullperioden (*NNPerioden*) in Abbildung 6-1 klassifiziert.

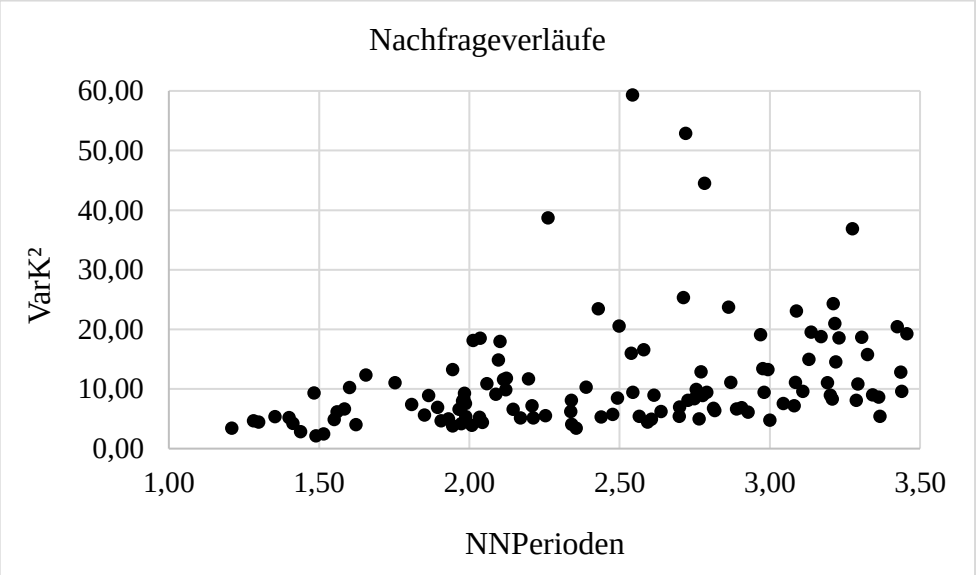

Abbildung 6-1: Anwendungsfall 1 – Nachfrage

Die Abbildung zeigt, dass die Produkte über ein breites Spektrum an VarK² (2,13 bis 59,29) und NNPerioden (1,21 und 3,46) verfügen. Die hohen Ausprägungen dieser beiden Nachfrage-Eigenschaften weisen auf volatile und sporadische Nachfrageverläufe hin. Da einzelne Werte außerhalb des in den synthetischen Daten abgedeckten Bereichs liegen, ermöglicht der Datensatz die Prüfung der Übertragbarkeit der Ergebnisse auf besonders volatile und unregelmäßige Nachfrageverläufe.

Vom Unternehmen des ersten Anwendungsfalls konnten keine historischen Prognosefehler bereitgestellt werden, aus diesem Grund wurden zunächst grundlegende Prognoseverfahren auf Basis der vorliegenden Nachfrageverläufe angewandt. Allerdings zeigte sich, dass die grundlegenden Prognosemethoden wie der gleitende Mittelwert, Croston-Verfahren (Armadani et al. 2025, S. 13307 f.) oder ARIMA-Modelle (Armadani et al. 2025, S. 13306 f.), für die spezifischen Charakteristika der betrachteten Nachfrageprozesse nicht gut geeignet sind. Der relative Prognosefehler lag bei allen drei Prognosemethoden über 100 % für sämtliche Produkte, wodurch eine belastbare Bestellmengenplanung nicht gewährleistet werden kann. Es ist davon auszugehen, dass im Unternehmen im Rahmen

des Anwendungsfalls 1 erfahrungsbasiertes Wissen oder andere Prognoseverfahren genutzt werden, die zu deutlich geringeren Prognosefehlern führen. Konkrete Informationen hierüber liegen jedoch nicht vor.

Vor diesem Hintergrund werden für die Validierungsexperimente synthetische Prognosefehler auf Basis einer Normalverteilung erzeugt (vgl. Kapitel 3.2.3). Die Parametrisierung erfolgt über den Erwartungswert und die Varianz. Zur Untersuchung von unterschiedlich großen Prognosefehlern werden drei relative Prognosefehler von 5 %, 20 % und 50 % synthetisch erzeugt. Die Prognosefehler entsprechen somit bewusst nicht genau den in der synthetischen Datenerzeugung verwendeten Fehlern, um die Übertragbarkeit der Ergebnisse zu überprüfen. Negative Nachfragewerte, welche durch die Normalverteilung entstehen können, werden auf null gesetzt. Insgesamt ergeben sich aus den 121 Produkten mit jeweils drei unterschiedlichen Prognosefehlern 363 Validierungsexperimente für den ersten Anwendungsfall.

Im Folgenden werden die Verteilungen der synthetisch erzeugten Prognosefehler dargestellt. Dabei erfolgt eine differenzierte Analyse hinsichtlich Fehlertyp und Fehlerausprägung. Die Abbildung 6-2 zeigt wie häufig die drei **Fehlertypen** *Rauschen, Überschätzung* und *Unterschätzung* auftreten.

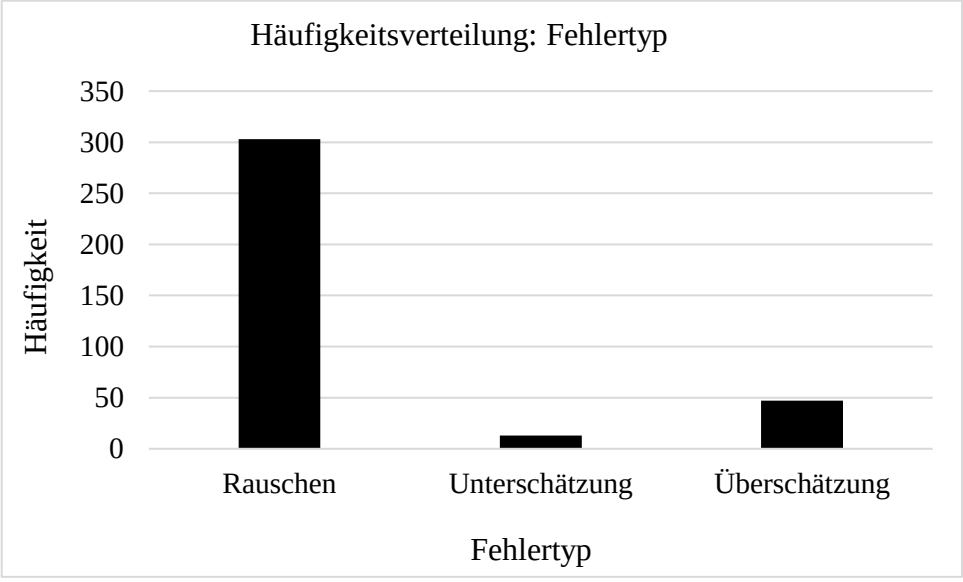

Abbildung 6-2: Anwendungsfall 1 – Fehlertyp

Am häufigsten tritt der Fehlertyp Rauschen (Häufigkeit = 303) auf. Dabei wird vom Fehlertyp Rauschen gesprochen, wenn die Summe aller einzelnen Prognosefehler in den jeweiligen Perioden zwischen -5 % und +5 % liegt. Wohingegen eine Überschätzung durch

eine Abweichung von größer als +5 % und eine Unterschätzung durch eine Abweichung von kleiner als -5 % charakterisiert ist.

Die **Größe des Prognosefehlers** ist in Abbildung 6-3 dargestellt. Sie berechnet sich aus der Summe der Beträge der einzelnen Prognosefehler je Periode. Insgesamt erstrecken sich die Prognosefehler im Datensatz von rund 4 % bis etwa 54 %. Somit entsprechen die Prognosefehler weitestgehend den drei definierten synthetischen Prognosefehlern. Die Abweichungen von den drei definierten Prognosefehlern begründen sich einerseits aus der Verwendung einer Wahrscheinlichkeitsverteilung und andererseits daraus, dass negative Nachfragewerte per Modellannahme ausgeschlossen werden.

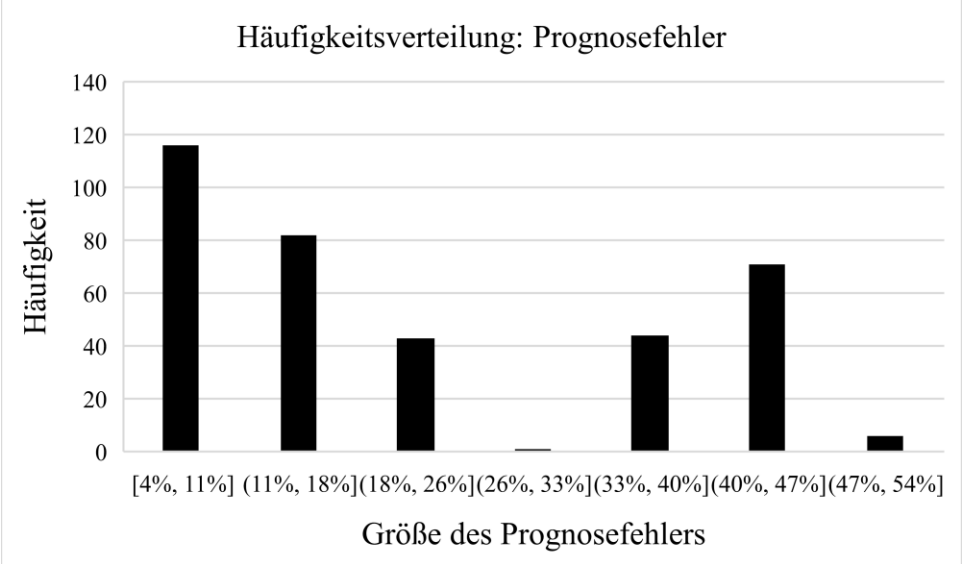

Abbildung 6-3: Anwendungsfall 1 – Prognosefehler

Die **Wiederbeschaffungszeiten** (WBZ) erstrecken sich von zwei Perioden bis 18 Perioden (vgl. Abbildung 6-4). Der Großteil der Produkte weist eine WBZ von 14 Perioden auf, was einer Wiederbeschaffungszeit von etwa zwei Wochen entspricht.

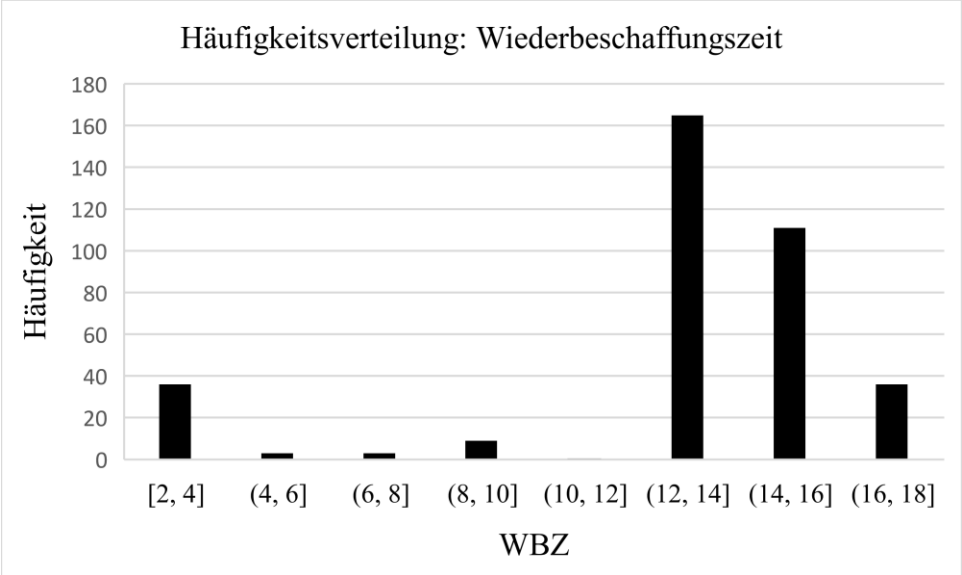

Abbildung 6-4: Anwendungsfall 1 – WBZ

Zur Berechnung der **Time Between Order** (TBO) wird die Formel (3-10) herangezogen. Da vom Unternehmen keine *Bestellkosten* zur Verfügung gestellt werden konnten, wird im Rahmen dieser Arbeit auf branchenübliche Durchschnittswerte zurückgegriffen. CAPS Research erhebt hierzu jährlich die durchschnittlichen Bestellkosten, differenziert nach verschiedenen Branchen. Für den Bereich Industrial Manufacturing liegen diese durchschnittlichen Kosten bei ca. 413 € pro Bestellung und umfassen sämtliche im Bestellprozess anfallenden Aufwände, von der Bedarfsidentifikation über den Wareneingang bis hin zur Zahlung (CAPS RESEARCH 2021). Demgegenüber ergeben sich die *Lagerhaltungskosten* aus dem Produktwert, multipliziert mit dem Lagerhaltungssatz von 7 %.

Die verschiedenen Produkte weisen eine TBO zwischen zwei und 13 Perioden auf. Dies bedeutet, dass eine kostenorientierte Bestellung im Durchschnitt alle zwei bis 13 Perioden erfolgt. Die Häufigkeitsverteilung der TBO ist in Abbildung 6-5 dargestellt.

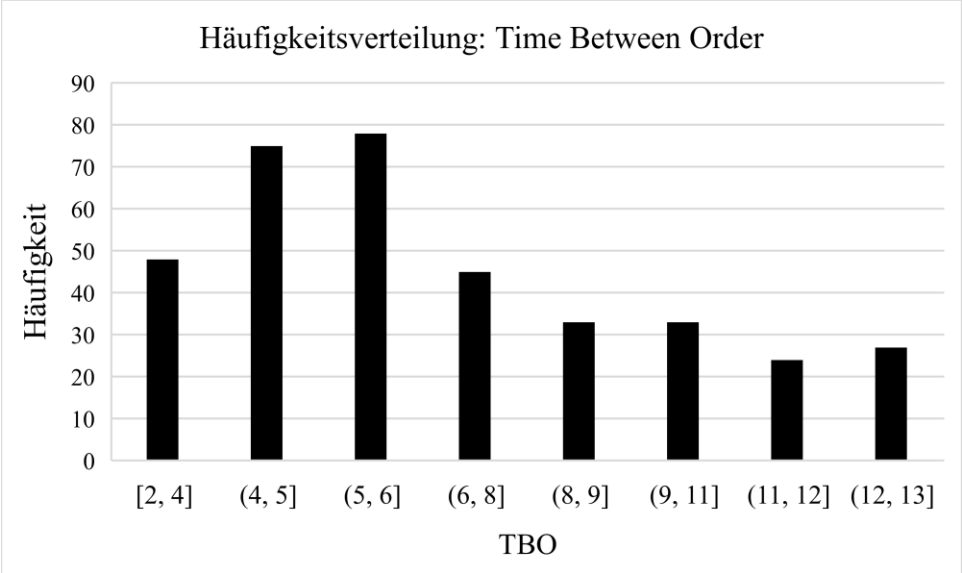

Abbildung 6-5: Anwendungsfall 1 – TBO

Abschließend wird der **Lieferservicegrad** im Anwendungsfall 1 unternehmensintern mit 99 % festgelegt. Dabei kommt der α-Lieferservicegrad zum Einsatz, sodass 99 % der eingehenden Nachfragen ohne Lieferverzug vollständig bedient werden können.

Zusammenfassend weist der erste Anwendungsfall hinsichtlich der Systemfaktoren und Faktorstufen ähnliche Ausprägungen wie die synthetischen Daten auf. Größere Abweichungen zeigen sich im Bereich der Nachfrageverläufe. Diese Abweichungen unterstützen jedoch die Aussagekraft der Validierung, da sie die Übertragbarkeit der entwickelten Ergebnisse über die synthetisch abgedeckten Faktorstufen hinaus gewährleisten. Im nachfolgenden Kapitel werden die Handlungsalternativen zur Heuristikwahl auf Basis des zuvor beschriebenen Datensatzes des ersten Anwendungsfalls analysiert und bewertet.

6.1.2 Ergebnisse der Validierungsexperimente: Anwendungsfall 1 – Handel von industriellen Gütern

In diesem Kapitel erfolgt die Validierung der entwickelten *Entscheidungsregeln* sowie des *ML-Modells* zur kostenorientierten Auswahl geeigneter Heuristiken. Hierzu werden Simulationsexperimente auf Basis der Unternehmensdaten aus dem ersten Anwendungsfall durchgeführt.

Für die Experimente wird ein Datensplit vorgenommen, bei dem in etwa die erste Hälfte des Datensatzes (entsprechend 750 Perioden) zur Ermittlung der Faktorstufen der relevanten Systemfaktoren herangezogen wird. Auf Grundlage dieser Faktorstufen erfolgt die kostenorientierte Auswahl der jeweils geeigneten Heuristik sowohl unter Anwendung der Entscheidungsregeln (vgl. Kapitel 4) als auch des trainierten ML-Modells (vgl. Kapitel

5). Die zweite Datenhälfte (zwischen 674 und 824 Perioden) wird anschließend in das in Kapitel 3 entwickelte Simulationsmodell eingespeist, um die resultierenden Gesamtkosten der vier betrachteten Heuristiken zu berechnen.

Für jedes Produkt werden mit dem Simulationsmodell, in Abhängigkeit der spezifischen Faktorstufen, ex post die Gesamtkosten der vier Heuristiken bestimmt. Da für die 121 Produkte unterschiedliche Zeithorizonte vorliegen, werden die Ergebnisse auf eine einjährige Laufzeit normalisiert. Im Anschluss werden die Planungsalternativen *zufällige Auswahl*, *Entscheidungsregeln* und *ML-Modell* systematisch gegenübergestellt und im Hinblick auf die ex post resultierenden Gesamtkosten bewertet. Die Ergebnisse der Simulationsexperimente werden nachfolgend erläutert.

Die Abbildung 6-6 zeigt die prozentualen Abweichungen der vier Heuristiken sowie die durchschnittliche Abweichung gegenüber der kostenoptimalen Heuristik über alle 363 Simulationsexperimente. Es zeigt sich, dass mit dem PP-Verfahren, mit einer durchschnittlichen Abweichung von etwa 2,81 % die geringsten Gesamtkosten erzielt werden, gefolgt vom Groff-Verfahren mit rund 3,91 %. Die durchschnittliche Abweichung über alle vier Verfahren beträgt 4,00 %.

Die Ergebnisse bestätigen damit die zuvor anhand synthetischer Daten gewonnenen Erkenntnisse. Die Gesamtkosten der vier Heuristiken liegen insgesamt eng beieinander, wobei das PP-Verfahren im Durchschnitt die besten Ergebnisse erzielt. Das Groff-Verfahren weist ebenfalls gute Ergebnisse auf, erweist sich jedoch nur selten als kostenoptimale Wahl. Damit wird auch das No-Free-Lunch-Theorem (vgl. Kapitel 4) bestätigt. Es existiert keine Heuristik, die für alle Faktorkombinationen optimale Ergebnisse liefert. Vielmehr ist eine kontextspezifische Auswahl erforderlich, bei der die jeweils geeignete Heuristik abhängig von den spezifischen Systemfaktoren zu bestimmen ist.

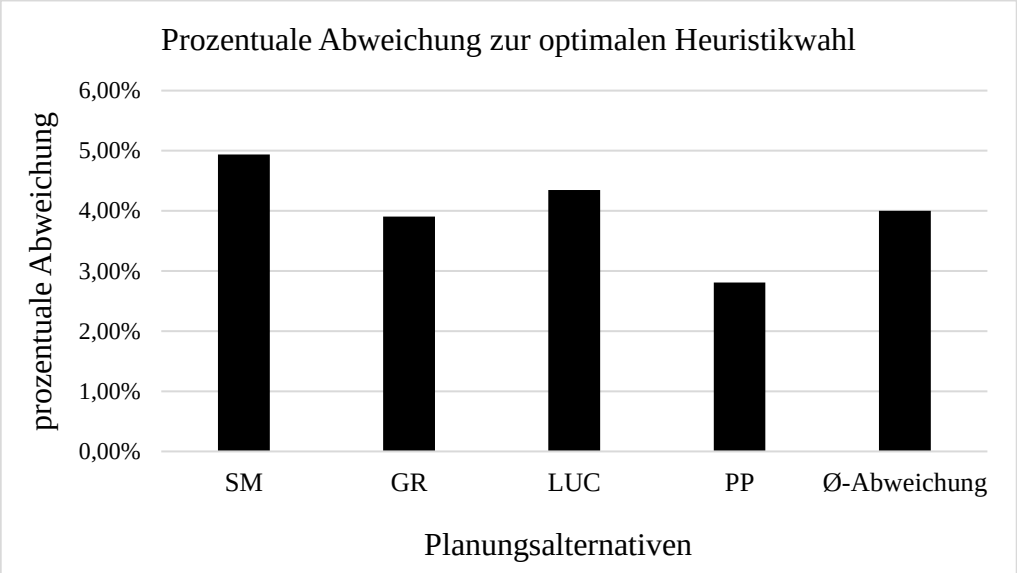

Abbildung 6-6: Anwendungsfall 1 – Prozentuale Abweichung zur optimalen Heuristikwahl

Im nächsten Schritt werden die drei Handlungsalternativen *zufällige Auswahl*, *Entscheidungsregeln* und *ML-Modell* systematisch untersucht. Bei der zufälligen Auswahl werden keine unterstützenden Methoden zur kostenorientierten Heuristikwahl eingesetzt. Diese Handlungsalternative wird daher durch die durchschnittliche Abweichung der vier Heuristiken von der kostenoptimalen Wahl (4,00 %) repräsentiert.

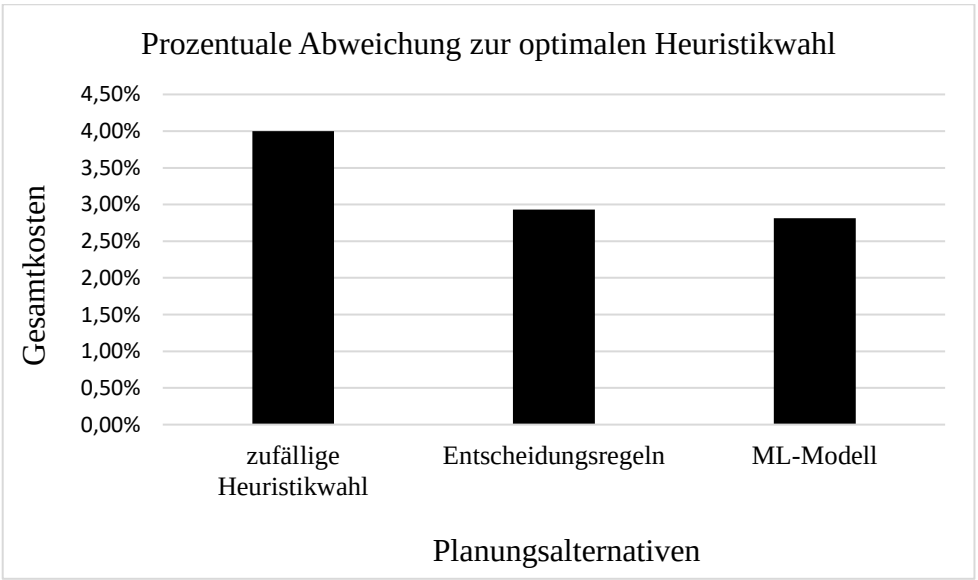

Abbildung 6-7: Anwendungsfall 1 – Prozentuale Abweichung zwischen den unterschiedlichen Handlungsalternativen

In Abbildung 6-7 sind die drei Handlungsalternativen hinsichtlich ihrer prozentualen Abweichung von der kostenoptimalen Heuristikwahl gegenübergestellt. Das trainierte ML-Modell erzielt über alle 363 Simulationsexperimente hinweg die geringsten Gesamtkosten und weist eine durchschnittliche Abweichung von lediglich 2,81 % auf. Die Entscheidungsregeln folgen mit einer Abweichung von 2,93 %. Beide methodischen Vorgehensweisen ermöglichen somit eine Reduktion der Gesamtkosten im Bezug zur dynamischen Bestellmengenplanung um mehr als 1 Prozentpunkt im Vergleich zur zufälligen Heuristikwahl. Diese Gesamtkosten umfassen einen wesentlichen Anteil der unternehmensrelevanten Kosten, insbesondere die Bestell- und Lagerhaltungskosten.

Wird die zufällige Heuristikwahl als derzeit vorherrschende Vorgehensweise interpretiert, ergibt sich über alle 363 Simulationsexperimente ein maximales theoretisches Einsparpotenzial aus der Differenz zwischen zufälliger und optimaler Heuristikwahl. Für den ersten Anwendungsfall beträgt dieses theoretische Maximum rund 1.290.000 € pro Jahr. Aufgrund von Änderungen der Einflüsse im Zeitverlauf, etwa von Prognosefehlern, kann dieses Einsparpotenzial jedoch nicht vollständig realisiert werden, es stellt vielmehr eine obere Grenze unter Verwendung der vier untersuchten Heuristiken dar.

Werden statt der zufälligen Auswahl die Entscheidungsregeln eingesetzt, können jährlich etwa 344.000 € für die 363 Simulationsexperimente eingespart werden. Der Einsatz des ML-Modells führt zu einem Einsparpotenzial von rund 383.000 € pro Jahr. Die resultierenden Einsparungen sind in Abbildung 6-8 dargestellt.

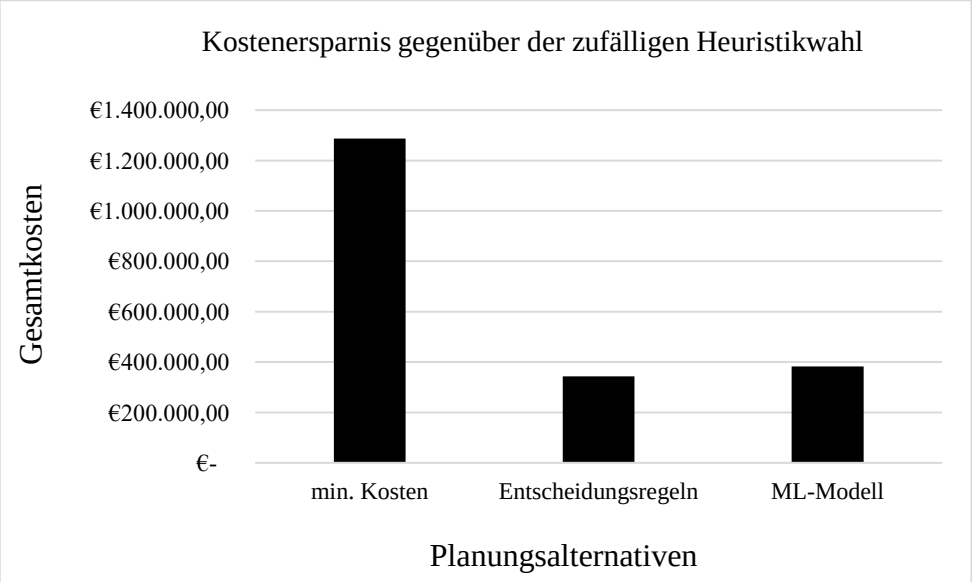

Abbildung 6-8: Anwendungsfall 1 – Kostenersparnis gegenüber der zufälligen Heuristikwahl

Die Analyse zeigt, dass durch eine kostenorientierte Auswahl dynamischer Bestellmengenverfahren sowohl mittels der entwickelten Entscheidungsregeln als auch durch das trainierte ML-Modell signifikante Kosteneinsparungen erzielt werden können. Für den ersten Anwendungsfall erweist sich insbesondere das ML-Modell als vorteilhaft. Gegenüber der Auswahl mittels Entscheidungsregeln lassen sich zusätzliche Einsparungen in Höhe von rund 39.000 € pro Jahr für die betrachteten 363 Simulationsexperimente mit dem ML-Modell realisieren. Bei den 121 Produkten bzw. 363 Simulationsexperimente handelt es sich um einen Ausschnitt des Produktportfolios, sodass das Einsparungspotenzial für alle Produkte des Unternehmens noch größer sein könnte.

6.2 Simulationsexperimente: Anwendungsfall 2 – Handel von Gastronomieprodukten

Im nachfolgenden Kapitel werden die Daten des zweiten Anwendungsfalls hinsichtlich der sechs Systemfaktoren und deren Faktorstufen analysiert und beschrieben (vgl. Kapitel 6.2.1). Anschließend werden im Kapitel 6.2.2 die Simulationsexperimente erläutert sowie die verschiedenen Handlungsalternativen zur Heuristikwahl auf Basis der entstehenden Gesamtkosten analysiert und bewertet.

6.2.1 Vorstellung der Datengrundlage: Anwendungsfall 2 – Handel von Gastronomieprodukten

Der zweite Anwendungsfall aus dem Bereich des Gastronomiebedarfs umfasst 26 unterschiedliche Produkte, für welche die Bestellmengen und -zeitpunkte geplant werden. Für alle Produkte wurden die sechs relevanten Systemfaktoren erhoben. Die verschiedenen Systemfaktoren und ihre jeweiligen Faktorstufen werden nachfolgend näher beschrieben.

Die historischen **Nachfragen** der 26 Produkte weisen eine Dauer zwischen 155 und 836 Perioden auf. Die unterschiedlichen Nachfrageverläufe werden wie zuvor hinsichtlich der zwei Eigenschaften Streuung des Bedarfs (*VarK²*) und Anzahl der Nullperioden (*NNPerioden*) klassifiziert und in Abbildung 6-9 visualisiert.

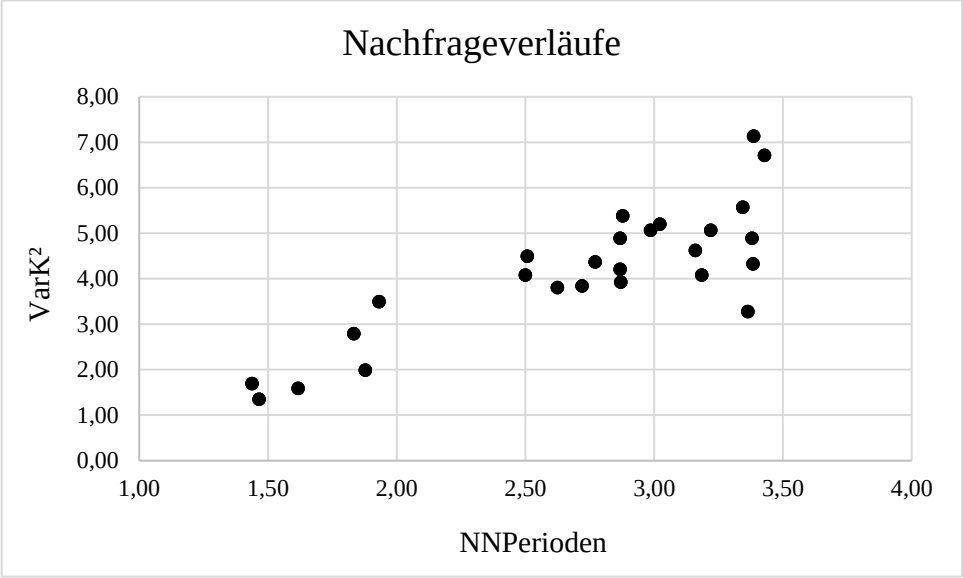

Abbildung 6-9: Anwendungsfall 2 – Nachfrage

Die Abbildung 6-9 zeigt, dass die Produkte über ein breites Spektrum an VarK² (0,98 bis 7,13) und NNPerioden (1,29 und 3,43) verfügen. Auch wenn die Streuung deutlich niedriger ist als im ersten Anwendungsfall, deuten die hohen Ausprägungen von Streuung und NNPerioden dennoch auf volatile und sporadische Nachfrageverläufe hin.

Auch vom Unternehmen des zweiten Anwendungsfalls konnten keine historischen Prognosefehler bereitgestellt werden. Ebenso führen auch die grundlegende Prognoseverfahren (gleitende Mittelwert, Croston- und ARIMA-Verfahren) zu großen relativen Prognosefehlern. Der relative Prognosefehler lag bei allen drei Prognosemethoden über 100 % für sämtliche Produkte, wodurch eine belastbare Bestellmengenplanung nicht gewährleistet werden kann. Es ist davon auszugehen, dass im Unternehmen im Rahmen des Anwendungsfalls 2 erfahrungsbasiertes Wissen oder andere Prognoseverfahren genutzt werden, die zu deutlich geringeren Prognosefehlern führen. Konkrete Informationen hierüber liegen jedoch nicht vor.

Vor diesem Hintergrund werden wie im ersten Anwendungsfall auch hier drei synthetische Prognosefehler (5 %, 20 % und 50 %) auf Basis einer Normalverteilung erzeugt (vgl. Kapitel 3.2.3). Das Vorgehen zur Erzeugung der Prognosefehler gleicht dem Vorgehen des ersten Anwendungsfalls (vgl. Kapitel 6.1.1). Insgesamt ergeben sich aus den 26 Produkten mit jeweils drei unterschiedlichen Prognosefehlern 78 Validierungsexperimente für den zweiten Anwendungsfall.

Im Folgenden werden die Verteilungen der synthetisch erzeugten Prognosefehler dargestellt. Dabei erfolgt eine differenzierte Analyse hinsichtlich Fehlertyp und Fehlerausprägung. Die Abbildung 6-10 zeigt wie häufig die drei Fehlertypen *Rauschen, Überschätzung* und *Unterschätzung* auftreten.

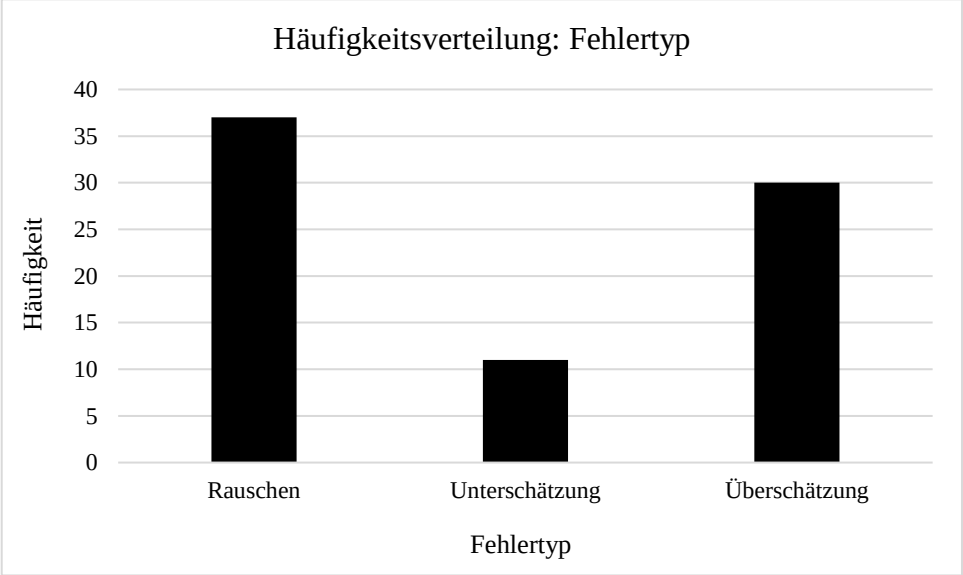

Abbildung 6-10: Anwendungsfall 2 – Fehlertyp

Am häufigsten tritt der Fehlertyp Rauschen (Häufigkeit = 37), gefolgt von Überschätzungen (Häufigkeit = 30) und Unterschätzungen (Häufigkeit = 11) auf. Die Größe des Prognosefehlers ist in Abbildung 6-11 dargestellt. Sie berechnet sich aus der Summe der Beträge der einzelnen Prognosefehler je Periode. Insgesamt erstrecken sich die Prognosefehler im Datensatz von rund 5 % bis etwa 74 %. Die Abweichungen von den drei definierten Prognosefehlern begründen sich einerseits aus der Verwendung einer Wahrscheinlichkeitsverteilung und andererseits daraus, dass negative Nachfragewerte per Modellannahme ausgeschlossen werden.

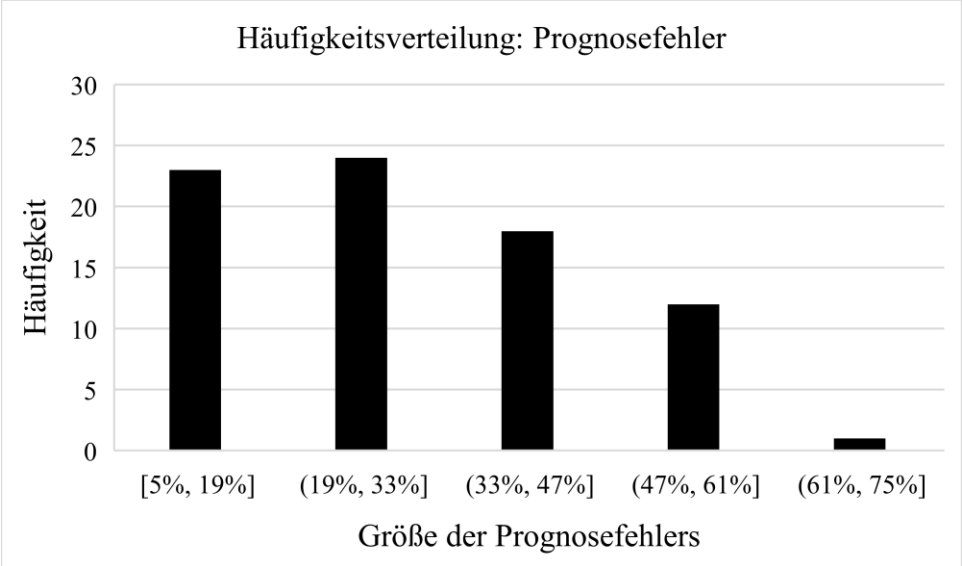

Abbildung 6-11: Anwendungsfall 2 – Prognosefehler

Die **Wiederbeschaffungszeiten** (WBZ) erstrecken sich von drei Perioden bis 15 Perioden (vgl. Abbildung 6-12). Der Großteil der Produkte weist eine WBZ zwischen drei und fünf Perioden auf. Dementsprechend treten häufig kleinere WBZ als im Anwendungsfall 1 auf.

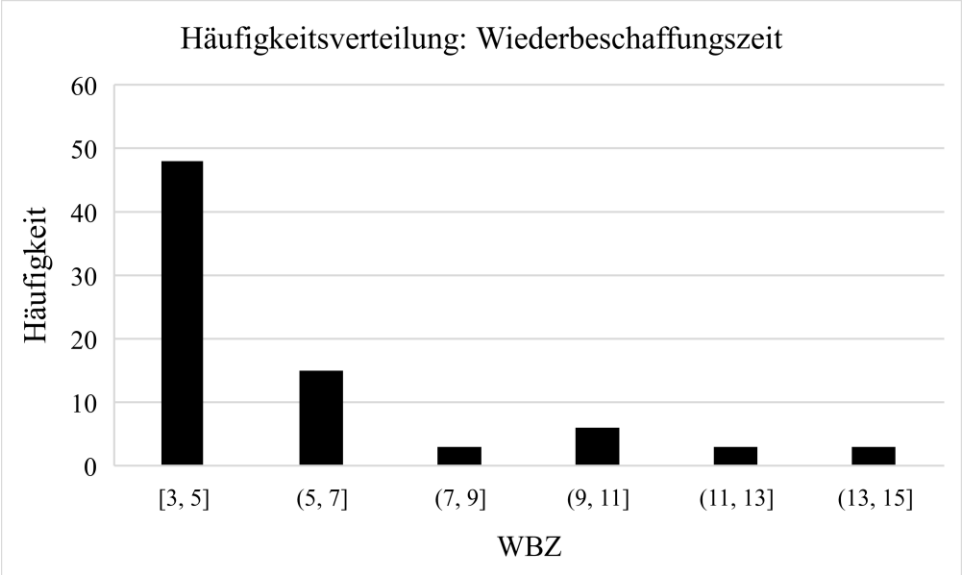

Abbildung 6-12: Anwendungsfall 2 – WBZ

Zur Berechnung der **Time Between Order** (TBO) wird die Formel (3-10) herangezogen. Die Bestellkosten sind mit 33,95 € pro Bestellung definiert und umfassen sämtliche Kosten, die im Rahmen eines Bestellvorgangs entstehen, von der Bedarfsermittlung bis zum

Wareneingang. Die Lagerhaltungskosten ergeben sich aus dem Produktwert, multipliziert mit dem Lagerhaltungssatz von 5 %.

Die verschiedenen Produkte weisen eine TBO zwischen drei und 13 Perioden auf. Dies bedeutet, dass eine Bestellung im Durchschnitt alle drei bis 13 Perioden erfolgt. Die Häufigkeitsverteilung der TBO ist in Abbildung 6-13 dargestellt.

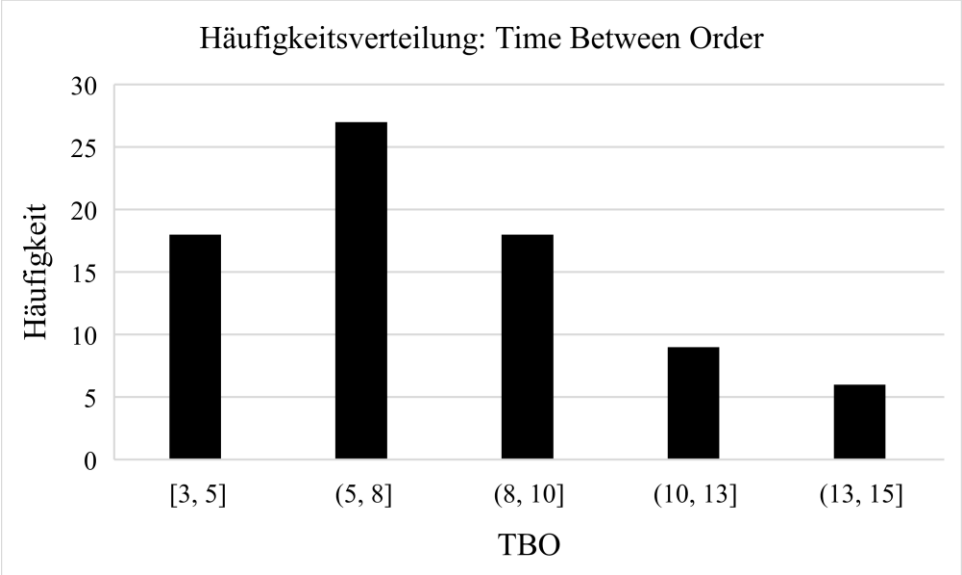

Abbildung 6-13: Anwendungsfall 2 – TBO

Abschließend wird der Lieferservicegrad im Anwendungsfall 2 unternehmensintern mit 97,5 % festgelegt. Dabei kommt der α-Lieferservicegrad zum Einsatz, sodass 97,5 % der eingehenden Nachfragen ohne Lieferverzug vollständig bedient werden können.

Zusammenfassend weist der zweite Anwendungsfall hinsichtlich der Systemfaktoren und Faktorstufen ähnliche Ausprägungen wie die synthetischen Daten auf. Größere Abweichungen zeigen sich lediglich im Bereich der Nachfrageverläufe. Diese Abweichungen erhöhen die Aussagekraft der Validierung, da die Abweichung eine Überprüfung der Übertragbarkeit der entwickelten Ergebnisse über die in den synthetischen Daten hinaus berücksichtigten Faktorkombinationen ermöglichen. Ebenso ergänzt Anwendungsfall 2 den ersten Anwendungsfall, indem weitere Faktorstufen berücksichtigt werden können. Dies wird insbesondere im Bereich der (WBZ) und der Kostenstruktur deutlich. Hier treten geringere WBZ auf und es werden zudem kostengünstigere Produkte mit niedrigeren Bestellkosten einbezogen.

Im nachfolgenden Kapitel werden die Handlungsalternativen zur Heuristikwahl auf Basis des zuvor beschriebenen Datensatzes des zweiten Anwendungsfalls analysiert und bewertet.

6.2.2 Validierungsexperimente: Anwendungsfall 2 – Handel von Gastronomieprodukten

In diesem Kapitel erfolgt die Validierung der entwickelten *Entscheidungsregeln* sowie des *ML-Modells* auf Basis des zweiten Anwendungsfalls. Das Vorgehen gleicht der Validierung auf Basis des ersten Anwendungsfalls (vgl. Kapitel 6.1.2).

Für die Experimente wird ebenso ein Datensplit vorgenommen, bei dem die erste Hälfte des Datensatzes (zwischen 77 und 418 Perioden) zur Ermittlung der Faktorstufen der Systemfaktoren herangezogen wird. Auf Grundlage dieser Faktorstufen erfolgt die kostenorientierte Auswahl der jeweils geeigneten Heuristik sowohl unter Anwendung der Entscheidungsregeln (vgl. Kapitel 4) als auch des trainierten ML-Modells (vgl. Kapitel 5). Die zweite Datenhälfte wird anschließend in das in Kapitel 3 entwickelte Simulationsmodell eingespeist, um ex post die resultierenden Gesamtkosten der vier betrachteten Heuristiken zu berechnen.

Für jedes Produkt werden mit dem Simulationsmodell, in Abhängigkeit der spezifischen Faktorstufen, die Gesamtkosten der vier Heuristiken bestimmt. Da für die 26 Produkte unterschiedliche Zeithorizonte vorliegen, werden sämtliche Ergebnisse auf eine einjährige Laufzeit normalisiert. Im Anschluss werden die Planungsalternativen *zufällige Auswahl*, *Entscheidungsregeln* und *ML-Modell* systematisch gegenübergestellt und im Hinblick auf die resultierenden Gesamtkosten ex post bewertet. Die Ergebnisse der Simulationsexperimente werden nachfolgend erläutert.

Die Abbildung 6-14 zeigt die prozentualen Abweichungen der vier Heuristiken sowie die durchschnittliche Abweichung gegenüber der kostenoptimalen Heuristik über alle 78 Simulationsexperimente. Es zeigt sich, dass mit dem PP-Verfahren, mit einer durchschnittlichen Abweichung von etwa 4,47 % die geringsten Gesamtkosten erzielt werden, gefolgt vom LUC-Verfahren mit rund 9,25 %. Die durchschnittliche Abweichung über alle vier Verfahren beträgt 10,39 %.

In diesem Anwendungsfall sind die Leistungsunterschiede der vier Heuristiken deutlich größer als im ersten Anwendungsfall, wobei das PP-Verfahren im Durchschnitt weiterhin die besten Ergebnisse erzielt. Die deutlichen Leistungsunterschiede lassen sich vor allem auf die Branchenunterschiede mit anderen Produkten und die miteinhergehenden Faktorkombinationen zurückführen, z. B. Verteilung der systematischen Prognosefehler. Insgesamt wird auch hier das No-Free-Lunch-Theorem (vgl. Kapitel 4) bestätigt. Es existiert keine Heuristik, die für alle Faktorkombinationen optimale Ergebnisse liefert. Vielmehr ist eine kontextspezifische Auswahl erforderlich, bei der die jeweils geeignete Heuristik abhängig von den spezifischen Systemfaktoren zu bestimmen ist.

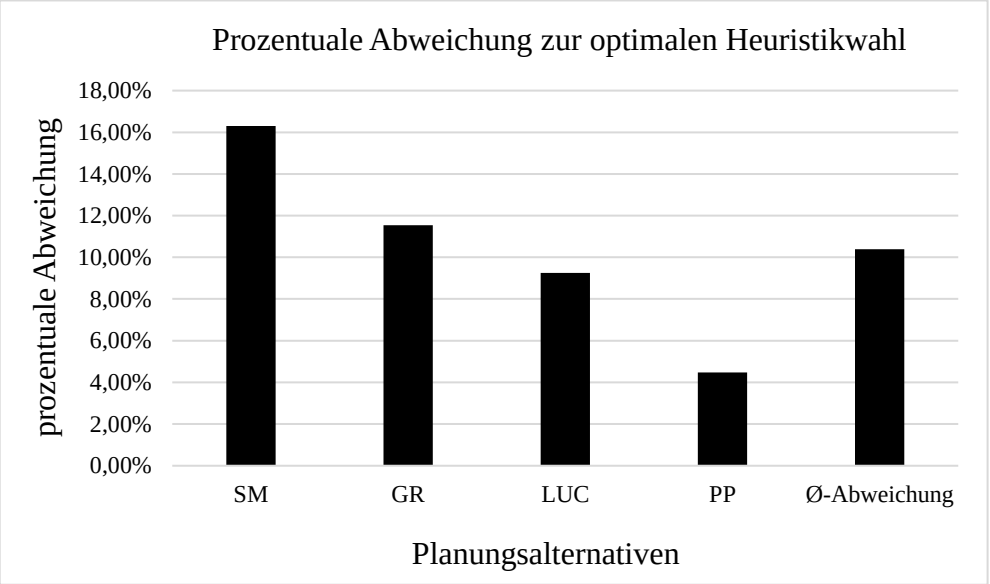

Abbildung 6-14: Anwendungsfall 2 – Prozentuale Abweichung zur optimalen Heuristikwahl

Im nächsten Schritt werden die drei Handlungsalternativen *zufällige Auswahl*, *Entscheidungsregeln* und *ML-Modell* systematisch untersucht. Bei der zufälligen Auswahl werden keine unterstützenden Methoden zur kostenorientierten Heuristikwahl eingesetzt. Diese Handlungsalternative wird daher durch die durchschnittliche Abweichung der vier Heuristiken von der kostenoptimalen Wahl (10,39 %) repräsentiert.

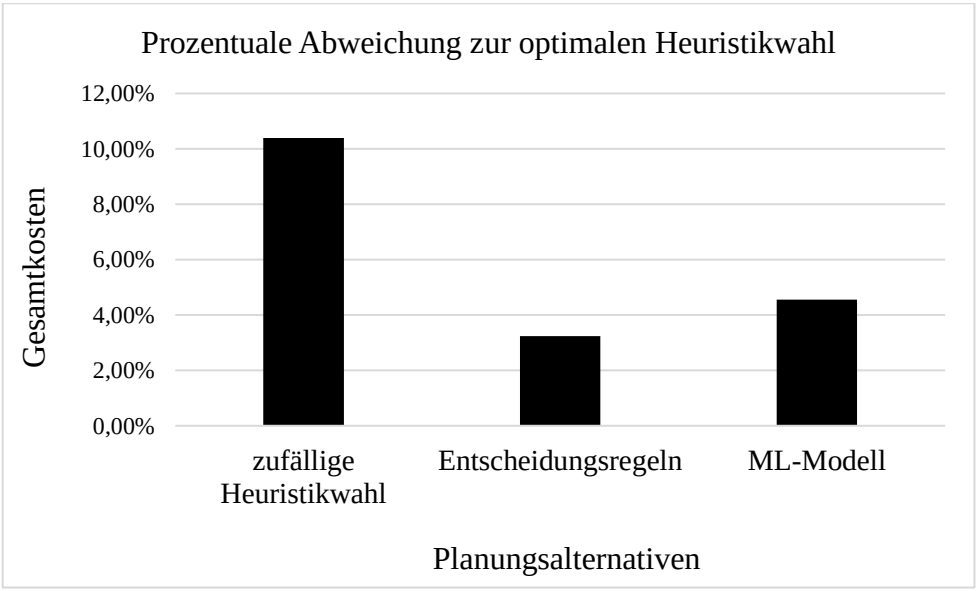

Abbildung 6-15: Anwendungsfall 2 – Prozentuale Abweichung zwischen den unterschiedlichen Handlungsalternativen

In Abbildung 6-15 sind die drei Handlungsalternativen hinsichtlich ihrer prozentualen Abweichung von der kostenoptimalen Heuristikwahl gegenübergestellt. Die Entscheidungsregeln erzielen über alle 78 Simulationsexperimente hinweg die geringsten Gesamtkosten und weisen eine durchschnittliche Abweichung von lediglich 3,24 % auf. Das trainierte ML-Modell folgt mit einer Abweichung von 4,56 %. Beide Vorgehensweisen ermöglichen somit eine Reduktion der Gesamtkosten zwischen 5,83 Prozentpunkte und 7,15 Prozentpunkte im Vergleich zur zufälligen Heuristikwahl.

Wird die zufällige Heuristikwahl als derzeit vorherrschende Vorgehensweise interpretiert, ergibt sich für die 78 Simulationsexperimente ein maximales theoretisches Einsparpotenzial aus der Differenz zwischen zufälliger und optimaler Heuristikwahl. Für den zweiten Anwendungsfall beträgt dieses theoretische Maximum rund 32.700 € pro Jahr. Aufgrund von Änderungen der Einflüsse im Zeitverlauf, etwa von Prognosefehlern, kann dieses Einsparpotenzial jedoch nicht vollständig realisiert werden, es stellt auch in diesem Fall vielmehr eine obere Grenze unter Verwendung der vier untersuchten Heuristiken dar.

Werden die Entscheidungsregeln zur Heuristikwahl eingesetzt, können jährlich etwa 22.500 € für die 78 Simulationsexperimente eingespart werden. Der Einsatz des ML-Modells führt zu einem Einsparpotenzial von rund 18.400 € pro Jahr. Die resultierenden Einsparungen sind in Abbildung 6-16 visualisiert.

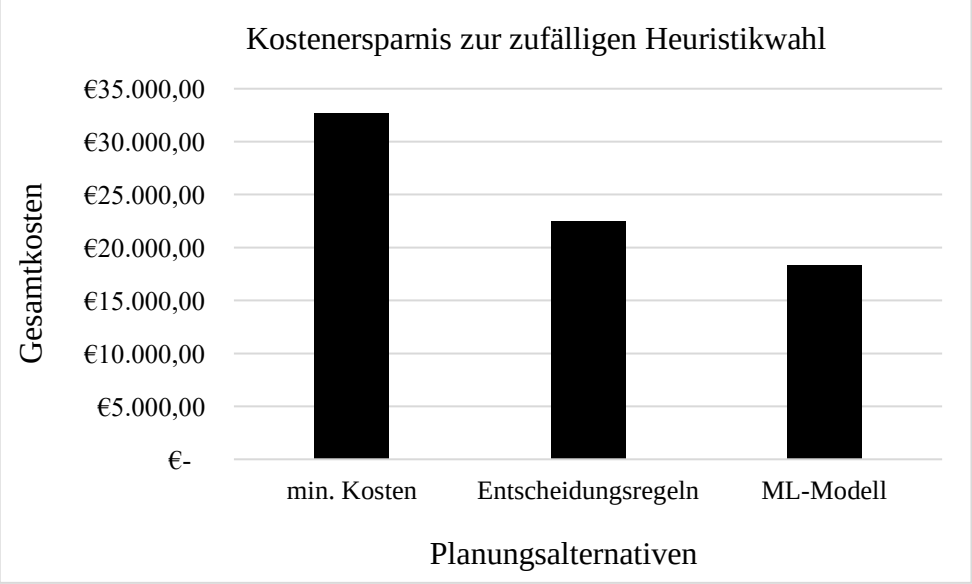

Abbildung 6-16: Anwendungsfall 2 – Kostenersparnis gegenüber der zufälligen Heuristikwahl

Die Ergebnisse der Analyse belegen, dass eine kontextspezifische kostenorientierte Auswahl dynamischer Bestellmengenverfahren sowohl durch die entwickelten Entscheidungsregeln als auch durch das trainierte ML-Modell zu deutlichen Kostensenkungen führt. Im zweiten Anwendungsfall sind die Entscheidungsregeln gegenüber dem ML-Modell leicht im Vorteil, sodass zusätzliche Einsparungen von rund 4.000 € pro Jahr für die 78 Simulationsexperimente realisiert werden. Gleichzeitig zeigt die Analyse, dass es sich im zweiten Anwendungsfall um deutlich günstigere Produkte mit niedrigeren Bestellkosten handelt, sodass die Einsparungspotenziale in absoluten Kosten geringer ausfallen als im ersten Anwendungsfall. Bei den 78 Simulationsexperimenten handelt es sich um einen Ausschnitt des Produktportfolios, sodass das Einsparungspotenzial für alle Produkte des Unternehmens ebenfalls größer sein könnte.

6.3 Zwischenfazit: Validierungsexperimente

Im Rahmen der Validierung werden die entwickelten Entscheidungsregeln sowie das ML-Modell zur kostenorientierten Auswahl dynamischer Bestellmengenverfahren anhand realer Unternehmensdaten erfolgreich validiert. Dafür werden insgesamt 441 Simulationsexperimente auf Basis von zwei Anwendungsfällen durchgeführt. Ziel ist es, die entstehenden Gesamtkosten der unterschiedlichen Handlungsalternativen systematisch zu bewerten: *zufällige Heuristikwahl*, *Entscheidungsregeln* sowie *ML-Modell*.

Die Ergebnisse beider Anwendungsfälle bestätigen, dass keine der betrachteten Heuristiken in allen Situationen eine kostenoptimale Lösung bietet. Damit bestätigt sich das No-Free-Lunch-Theorem, sodass eine situative Auswahl hinsichtlich der Faktorkombinationen erforderlich ist, um Kostenvorteile realisieren zu können.

Anwendungsfall 1 (363 Simulationsexperimente) zeigt im Bereich der Nachfrageverläufe eine hohe Volatilität der Nachfrage auf. Die übrigen Faktorstufen weisen eine ähnliche Struktur wie die synthetischen Daten auf. Hier erzielt das trainierte ML-Modell die geringsten Gesamtkosten mit einer durchschnittlichen Abweichung von 2,81 % gegenüber der optimalen Heuristikwahl. Die Entscheidungsregeln erreichen ein ähnliches Niveau (2,93 %). Beide Ansätze reduzieren die Gesamtkosten um mehr als 1 Prozentpunkt im Vergleich zu einer zufälligen Auswahl. Das jährliche Einsparpotenzial beträgt rund 344.000 € (Entscheidungsregeln) bzw. 383.000 € (ML-Modell), bezogen auf 363 Simulationsexperimente.

Anwendungsfall 2 (78 Simulationsexperimente) weist günstigere Produkte mit niedrigen Bestellkosten, kürzere Wiederbeschaffungszeiten und häufige Über- oder Unterschätzungen im Bereich der Prognose auf. In diesem Fall schneiden die Entscheidungsregeln am besten ab, mit einer durchschnittlichen Abweichung von 3,24 % gegenüber der optimalen Heuristikwahl. Das ML-Modell erreicht 4,56 % und die zufällige Auswahl liegt bei 10,39

%. Die absoluten Einsparungen liegen bei den Entscheidungsregeln bei 22.500 € und beim ML-Modell bei 18.400 €.

Über beide Anwendungsfälle hinweg lässt sich nachweisen, dass sowohl die Entscheidungsregeln als auch das ML-Modell zu einer deutlichen Kostenreduktion führen. Die Ergebnisse sollten bei der praktischen Anwendung hinsichtlich der jeweiligen Unternehmensbedingungen reflektiert und hinsichtlich weiterer Anwendungsfälle überprüft werden.

7 Fazit und Ausblick

In diesem Kapitel werden zunächst die zentralen Ergebnisse zusammengefasst und die vier Forschungsfragen beantwortet (vgl. Kapitel 7.1). Anschließend werden der Forschungsprozess und -methoden sowie die entwickelten Ergebnisse kritisch reflektiert (vgl. Kapitel 7.2). Abschließend wird ein Ausblick auf zukünftige Forschungsbedarfe gegeben, wobei weitere forschungsrelevante Fragestellungen aufgezeigt werden (vgl. Kapitel 7.3).

7.1 Zusammenfassung der Ergebnisse

Im Mittelpunkt der vorliegenden Arbeit steht die Verbindung von synthetischer Datenerzeugung und Machine Learning zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren. In der betrieblichen Praxis stehen dafür in den gängigen ERP-Systemen vier Heuristiken (SM, Groff, LUC und PP) zur Verfügung. Die Auswahl der Verfahren erfolgt häufig zufällig. Durch bestehende Forschungsarbeiten in diesem Bereich wird aufgezeigt, dass deutliche Kostenunterschiede zwischen den vier Verfahren je nach Planungsumgebung auftreten können. Vor diesem Hintergrund birgt die Kombination aus synthetischer Datenerzeugung, statistischer Datenanalyse und Machine Learning ein großes Potenzial, um unter gegebenen System- und Umweltbedingungen das jeweils kostenminimale Verfahren zu identifizieren. Hierzu wurden insgesamt vier Forschungsfragen definiert und beantwortet, die den Untersuchungsrahmen strukturieren.

Im zweiten Kapitel wurde die erste Forschungsfrage: „*FF1: Welche Anforderungen an die Bestellmengenplanung charakterisieren das unternehmerische Umfeld?*“ adressiert. Im Rahmen einer systematischen Aufarbeitung des Untersuchungsbereichs wurden die in der Praxis relevanten dynamischen Bestellmengenverfahren identifiziert und die zentralen Systemfaktoren herausgearbeitet. Die Analyse des Forschungsstands zeigt, dass bestehende Arbeiten die Systemfaktoren häufig nur eingeschränkt berücksichtigen und es kein übertragbares und praxisorientiertes Entscheidungsmodell zur Heuristikwahl gibt. Daraus ergibt sich zugleich der Forschungsbedarf und -lücke für die Erhebung einer umfassenden und allgemeingültigen Datengrundlage sowie weiterführend für die Entwicklung von Entscheidungsmodellen zur Heuristikwahl.

Da reale Unternehmensdaten nur einen begrenzten Ausschnitt möglicher Faktorkombinationen abbilden und häufig nicht in ausreichender Menge und Qualität vorliegen, wurde ein simulativer Ansatz gewählt, um synthetische Daten der dynamischen Bestellmengenplanung zu erzeugen (vgl. Kapitel 3). Zur Beantwortung der zweiten Forschungsfrage: *„FF. 2: Wie lassen sich synthetische Daten zur dynamischen Bestellmengenplanung erzeugen?“* wurde die dynamische Bestellmengenplanung systemtheoretisch beschrieben

und modelliert. Im Rahmen der Simulationsexperimente wurden die relevanten Systemfaktoren und ihre Faktorstufen modelliert. Auf dieser Basis wurde ein umfangreicher Versuchsplan mit 3.888 Faktorkombinationen und mit 150 Replikationen je Faktorkombination für die vier dynamischen Bestellmengenverfahren erstellt. Die Modellierung und Simulation wurde nach einem etablierten Vorgehensmodell verifiziert und validiert, sodass die Glaubwürdigkeit der Ergebnisse nachgewiesen wurde. Damit wurde eine umfassende, reproduzierbare und generalisierbare Datenbasis zur Beantwortung der zweiten Forschungsfrage geschaffen.

Für die dritte Forschungsfrage *„FF. 3: Welche Entscheidungsregeln können zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren definiert werden?“* (vgl. Kapitel 4) wurden aufbauend auf den synthetischen Daten zunächst die grundlegenden Wirkzusammenhänge analysiert. Es zeigte sich, dass das PP-Verfahren zwar am häufigsten die geringsten Kosten verursacht, gleichzeitig aber die höchsten Extremwerte bei der Kostenabweichung aufweist. Das LUC-Verfahren erwies sich demgegenüber als robuster, mit geringeren maximalen Abweichungen, jedoch im Mittel mit höheren Kosten. Die Verfahren SM und Groff sind nur in wenigen, spezifischen Faktorkombinationen kostenoptimal. Bereits diese Analyse bestätigte das No-Free-Lunch-Theorem, sodass kein Verfahren existiert, das in allen Situationen überlegen ist. Vielmehr hängt die kostenoptimale Heuristikwahl von der jeweiligen Faktorkombination ab.

Zur Ableitung konkreter Entscheidungsregeln wurde anschließend der Einfluss der Systemfaktoren detailliert untersucht. Dabei zeigte sich, dass der Prognosefehler und die TBO den größten Einfluss auf die kostenorientierte Auswahl besitzen sowie, dass das PP- und LUC-Verfahren ein gegenläufiges Verhalten hinsichtlich ihres Kostenvorteils aufweisen. Vor diesem Hintergrund wurden praxisorientierte Entscheidungsregeln definiert, durch Scatterplots ergänzt sowie mittels eines Entscheidungsbaums visualisiert.

Parallel zu den Entscheidungsregeln adressiert die vierte Forschungsfrage: *„FF. 4: Wie kann ein Machine Learning Modell zur kostenorientierten Auswahl von dynamischen Bestellmengenverfahren gestaltet werden?“* die Entwicklung eines Machine Learning Modells (vgl. Kapitel 5). Nach einer Einführung in grundlegende Konzepte des Machine Learnings wurden als geeignete Klassifikationsverfahren ein Entscheidungsbaum, Random Forest und LightGBM ausgewählt. Als zentrale Zielgröße dienten Anteil-Top-Heuristik (ATH), ergänzt um Präzision, Recall, F1-Score, Cohens Kappa sowie durchschnittliche Kostenabweichung und deren Standardabweichung.

Die Ergebnisse zeigen, dass alle drei ML-Verfahren vergleichbare Vorhersageleistungen erzielten, das LightGBM-Verfahren jedoch in nahezu allen Bewertungsmetriken leicht überlegen war. Besonders deutlich waren die Kostenunterschiede zwischen den vier Heuristiken bei den Fehlertypen Rauschen und Unterschätzung, sodass ein großes Kosteneine-

sparungspotenzial durch die richtige Heuristikwahl besteht. Bei systematischen Überschätzungen waren die Leistungsunterschiede zwischen den Heuristiken hingegen gering, sodass der Zusatznutzen eines ML-basierten Ansatzes begrenzt ist. Insgesamt bestätigt sich, dass ML-Modelle, insbesondere LightGBM, ein leistungsfähiges Instrument zur kostenorientierten Auswahl dynamischer Bestellmengenverfahren darstellen, sofern eine geeignete Datenbasis vorliegt.

Zur Validierung der entwickelten Entscheidungsregeln und des ML-Modells wurden Simulationsexperimente auf Basis realer Unternehmensdaten durchgeführt. In zwei Anwendungsfällen mit insgesamt 441 Simulationsexperimenten wurden die Handlungsalternativen zufällige Heuristikwahl, Entscheidungsregeln und ML-Modell hinsichtlich der entstehenden Gesamtkosten ex post miteinander verglichen. Im Anwendungsfall 1 erzielte das ML-Modell die geringsten Gesamtkosten mit einer durchschnittlichen Abweichung von 2,81 % gegenüber der optimalen Heuristikwahl, während die Entscheidungsregeln mit 2,93 % ein ähnliches Niveau erreichten. Beide Ansätze senkten die Gesamtkosten um mehr als 1 Prozentpunkt im Vergleich zur zufälligen Auswahl, was jährlichen Einsparungen von rund 344.000 € (Entscheidungsregeln) bzw. 383.000 € (ML-Modell) entspricht.

Im Anwendungsfall 2 zeigten sich größere Leistungsunterschiede zwischen den Handlungsalternativen. Die Entscheidungsregeln erwiesen sich als vorteilhaft, sodass eine durchschnittliche Abweichung von 3,24 % gegenüber der optimalen Heuristikwahl erreicht wurde, während das ML-Modell bei 4,56 % und die zufällige Auswahl bei 10,39 % lagen. Gegenüber der zufälligen Auswahl konnten mit den Entscheidungsregeln ca. 22.500 € eingespart werden. Die Ergebnisse beider Anwendungsfälle bestätigen somit einerseits, dass das No-Free-Lunch-Theorem auch unter realen Bedingungen zutrifft, sodass keine Heuristik bei allen Faktorkombinationen überlegen war und unterstreichen andererseits die Notwendigkeit einer kontextspezifischen Auswahl in Abhängigkeit von den Systemfaktoren.

Neben der Beantwortung der vier Forschungsfragen leistet die vorliegende Arbeit für die Wissenschaft einen wichtigen Beitrag, indem sie ein Vorgehen zur methodischen Verbindung von simulativer, synthetischer Datenerzeugung und Machine Learning in der Bestellmengenplanung aufzeigt und validiert. Die erzeugte synthetische Datenbasis, die abgeleiteten Entscheidungsregeln und das entwickelte ML-Modell öffnen einen Bezugsrahmen für die weiterführende Forschung innerhalb der Bestandsplanung. Für die betriebliche Praxis wurden konkrete, niederschwellige Regeln sowie ein trainiertes ML-Modell bereitgestellt, die eine kostenorientierte Auswahl dynamischer Bestellmengenverfahren ermöglichen.

7.2 Kritische Würdigung und Limitationen

Zur Sicherstellung eines nachvollziehbaren und wissenschaftlich fundierten Forschungsprozesses ist eine kritische Reflexion mit den gewählten Methoden und den erzielten Ergebnissen hinsichtlich der bestehenden Limitationen erforderlich.

Die systemtheoretische Modellierung erforderte eine explizite Festlegung von Systemgrenzen und relevanten Einflussfaktoren. Die Arbeit fokussiert sich bewusst auf sechs allgemeingültige Systemfaktoren und grenzt sich von weiteren unternehmensspezifischen Faktoren wie beispielsweise Kapazitätsrestriktionen, Mindestbestellmengen, Rabattstrukturen, Lieferantenrestriktionen oder organisatorischen Einflüssen in der Planung ab. Diese Reduktion ist notwendig, um den Problemraum systematisch zu erfassen und analysieren zu können. Darüber hinaus wurden im Rahmen der vorliegenden Arbeit die vier praxisrelevanten Heuristiken betrachtet, da diese in den gängigen ERP-Systemen zur Verfügung stehen. Neben diesen Planungsmöglichkeiten können unternehmensindividuell noch andere Verfahren zur Bestellmengenplanung relevant sein, welche nicht berücksichtigt wurden.

Bei der synthetischen Datenerzeugung konnten durch eine systematische Variation eine Vielzahl von Faktorkombinationen berücksichtigt werden und damit eine breitere Datengrundlage geschaffen werden, als durch reale Unternehmensdaten erhoben werden könnte. Durch die Wahl etablierter Verteilungen für Nachfrageprozesse, die Modellierung unterschiedlicher Prognosefehlertypen sowie eine Verifikation und Validierung des Simulationsmodells wurde eine fundierte Datenbasis erzeugt.

Dem stehen jedoch mehrere Limitationen gegenüber. Die Nachfrageverläufe beruhen auf definierten stochastischen Modellen und Parametern. Bei der Erzeugung der Nachfrageverläufe wurde keine zeitliche Abhängigkeit innerhalb der Zeitreihe berücksichtigt, sodass z. B. Trends, Saisonalitäten, Lebenszykluseffekte oder Strukturbrüche nicht berücksichtigt wurden. Über alle Simulationsexperimente hinweg ist der Simulationshorizont mit 1150 Perioden gleichbleibend und wurde nicht variiert. Auch die Modellierung von Servicegraden, Wiederbeschaffungszeiten und Kostenverhältnissen erfolgte über diskrete Faktorstufen. Dabei wurden typische Faktorkombinationen berücksichtigt, jedoch wurde damit nur ein Ausschnitt möglicher Kombinationen dargestellt. In diesem Kontext konnte bereits bei der Validierung gezeigt werden, dass auch für Faktorkombinationen außerhalb der synthetischen Daten die Entscheidungsregeln und das ML-Modell vorteilhaft sind.

Die Entscheidungsregeln wurden aus der synthetischen Datenbasis abgeleitet und bewusst auf Interpretierbarkeit und Niedrigschwelligkeit ausgelegt. Dies erforderte eine Komprimierung der komplexen Zusammenhänge zwischen Systemfaktoren und kostenorientierter Heuristikwahl, wodurch Interaktionseffekte nicht im Detail abgebildet wurden.

Die Validierung anhand der zwei realen Anwendungsfälle stärkt die Praxisnähe der Arbeit, bleibt jedoch hinsichtlich Breite und Tiefe limitiert. In beiden Anwendungsfällen konnten keine Prognosefehler durch die Unternehmen bereitgestellt werden, sodass hier auf synthetische Prognosefehler zurückgegriffen werden musste. Darüber hinaus basieren die Bestellkosten im ersten Anwendungsfall auf branchenüblichen Durchschnittskosten. Ebenso repräsentieren die betrachteten Unternehmen spezifische Branchen- und Produktkonstellationen, andere Einflussfaktoren, etwa saisonale Konsumgüter oder hochpreisige Investitionsgüter, wurden nur begrenzt erfasst. Zudem erfolgte keine langfristige Feldstudie, in der die Einführung der Entscheidungsregeln und des ML-Modells im operativen Planungsalltag begleitet wurde und Effekte auf Prozesse, IT-Systeme, Kennzahlen und Akzeptanz systematisch untersucht wurden.

Insgesamt ist die Arbeit als theoretisch fundierter und methodischer Beitrag zur kostenorientierten Auswahl dynamischer Bestellmengenverfahren zu verstehen. Die Ergebnisse ermöglichen eine kontextspezifische Heuristikwahl und zeigen deutliche Einsparpotenziale auf. Die Aussagekraft der Ergebnisse ist jedoch an die getroffenen Modellannahmen, den definierten Experimentplan, die eingesetzten statistischen Methoden sowie die spezifischen Anwendungsfälle gebunden. Die entwickelten Entscheidungsregeln und das trainierte ML-Modell sollten daher bei der praktischen Anwendung hinsichtlich der jeweiligen Rahmenbedingungen eines Unternehmens kritisch reflektiert werden.

7.3 Ausblick: zukünftige Forschungsbedarfe

Die vorliegende Arbeit zeigt, dass die Leistungsfähigkeit dynamischer Bestellmengenverfahren maßgeblich von Systemfaktoren und unternehmensspezifischen Rahmenbedingungen beeinflusst wird. Die entwickelten Entscheidungsregeln sowie das Machine Learning Modell ermöglichen eine kostenorientierte Heuristikwahl und leisten damit einen wissenschaftlich fundierten und praxisrelevanten Beitrag zur integrierten Betrachtung synthetisch erzeugter Daten und datenbasierter Entscheidungsfindung. Gleichzeitig bilden die erzielten Ergebnisse einen wichtigen Ausgangspunkt für weiterführende Forschung.

Ein zentraler zukünftiger Forschungsstrang betrifft die Weiterentwicklung der systemtheoretischen Modellierung der dynamischen Bestellmengenplanung. Die vorliegende Arbeit fokussiert sich auf sechs allgemeingültige Systemfaktoren und abstrahiert bewusst weitere Einflussgrößen wie Kapazitätsrestriktionen, Mindestbestellmengen, Staffelpreislogiken oder Unsicherheiten in den Wiederbeschaffungszeiten. Zukünftige Untersuchungen sollten diese Systemgrenzen erweitern und analysieren, wie zusätzliche Faktoren in die Entscheidungslogik integriert werden können. Ebenso erscheint die Berücksichtigung

weiterer Zielgrößen neben den Gesamtkosten vielversprechend, beispielsweise in Umgebungen mit limitierten Lagerkapazitäten durch bauliche Restriktionen, in denen Bestandskennzahlen wie durchschnittlicher oder maximaler Bestand im Vordergrund stehen können.

Ebenso kann eine weitergehende Variation relevanter Simulationsparameter, etwa des Simulations- bzw. Planungshorizonts, sowie eine zusätzliche Variation der Faktorstufen der sechs Systemfaktoren die Generalisierbarkeit und Übertragbarkeit der Ergebnisse erhöhen. Darüber hinaus sollten zukünftige Untersuchungen die Nachfrageverläufe um zeitliche Abhängigkeiten erweitern, sodass beispielsweise Trends und Saisonalitäten berücksichtigt werden können.

Darüber hinaus ist die Einbeziehung weiterer dynamischer Bestellmengenverfahren über die bislang betrachteten praxisrelevanten Heuristiken hinaus vielversprechend, um deren Leistungsfähigkeit unter unterschiedlichen Rahmenbedingungen vergleichen zu können. Auch die Untersuchung alternativer Dispositionsstrategien, die über die dynamische Bestellmengenplanung hinausgehen, stellt einen relevanten Ansatzpunkt für zukünftige Arbeiten dar.

Die hohe Bedeutung des Prognosefehlers für die Heuristikwahl verdeutlicht zudem die Notwendigkeit einer engeren Kopplung von Prognose- und Dispositionsentscheidungen. Während in dieser Arbeit die Wahl des Bestellmengenverfahrens in Abhängigkeit von Prognosefehlern untersucht wurde, besteht ein wesentliches Potenzial in der gleichzeitigen Optimierung von Nachfrageprognosen, Sicherheitsbeständen und Bestellmengen, um eine durchgängige kosten- und servicegradorientierte Planung zu gewährleisten.

Ein weiterer Ansatzpunkt für zukünftige Forschung besteht in den synthetisch generierten Daten zur Schaffung einer allgemeingültigen, übertragbaren und realitätsnahen Datengrundlage, welche in Praxis und Wissenschaft für weitere Forschung zur Planung und Steuerung logistischer Systeme verwendet werden kann. Dabei sollte der methodische Ansatz über die Bestellmengenplanung hinaus auf weitere logistische Planungs- und Steuerungsprozesse übertragen und untersucht werden.

Zudem sollte die Übertragbarkeit der Ergebnisse anhand umfangreicher empirischer Untersuchungen weiter validiert werden. Die bisherigen Validierungen mit zwei realen Unternehmensdatensätzen bestätigen zwar die Praxistauglichkeit der entwickelten Ansätze, erlauben jedoch noch keine allgemeingültigen Aussagen in Bezug auf unterschiedliche Branchen, Marktvolatilitäten oder Datenqualitäten. Künftige Studien sollten daher auch Implementierungsbarrieren, Interaktionskosten zwischen Mensch und Technik sowie organisatorische Akzeptanzfaktoren datenbasierter Dispositionslogiken adressieren.

Während die vorliegende Arbeit auf synthetisch generierten Daten basiert, um eine breite und übertragbare Datengrundlage zu gewährleisten, eröffnet eine datengetriebene Heuristikwahl auf realen Unternehmensdaten zusätzliche Erkenntnispotenziale. Unternehmensspezifische Muster, wie beispielsweise Nachfragemuster oder typische Prognosefehler, können durch Machine Learning Modelle präziser erfasst werden. Bei ausreichender Datenverfügbarkeit lässt sich somit gegebenenfalls ein spezifisches Modell entwickeln, das die Gesamtkosten zusätzlich reduziert. Darüber hinaus könnte die Nutzung realer Daten sowie die Anbindung an Echtzeitsysteme eine kontinuierliche Modellaktualisierung ermöglichen, wodurch die Leistungsfähigkeit der dynamischen Bestellmengenplanung weiter gesteigert werden kann.

Insgesamt leisten die Ergebnisse dieser Arbeit einen Beitrag zum wissenschaftlichen Diskurs zur kostenorientierten dynamischen Bestellmengenplanung und eröffnen zugleich vielfältige Anknüpfungspunkte für zukünftige Forschung. Insbesondere die Weiterentwicklung hin zu selbstlernenden, systemisch gekoppelten und zugleich interpretierbaren Planungssystemen stellt eine zentrale Herausforderung und zugleich eine bedeutende Chance für Wissenschaft und Praxis dar.

8 Literaturverzeichnis

Adam, Dietrich (1998): Produktions-Management. Wiesbaden: Gabler Verlag.

Alard, Robert (2002): Internetbasiertes Beschaffungsmanagement direkter Güter - Konzept zur Gestaltung der Beschaffung durch Nutzung internetbasierter Technologien: Konzept zur Gestaltung der Beschaffung durch Nutzung internetbasierter Technologien.

Alasali, F.; Haben, S.; Becerra, V.; Holdebraum, W. (2018): Day-ahead industrial load forecasting for electric RTG cranes. In: *J. Mod. Power Syst. Clean Energy* 6 (2), S. 223–234. DOI: 10.1007/s40565-018-0394-4.

Alicke, Knut (2005): Planung und Betrieb von Logistiknetzwerken. Unternehmensübergreifendes Supply Chain Management. 2., neu bearbeitete und erweiterte Auflage. Berlin, Heidelberg: Springer Berlin Heidelberg (SpringerLink Bücher).

Andler, Kurt (1929): Rationalisierung der Fabrikation und optimale Losgröße. Berlin, Boston: Oldenbourg Wissenschaftsverlag.

Andriolo, Alessandro; Battini, Daria; Grubbström, Robert W.; Persona, Alessandro; Sgarbossa, Fabio (2014): A century of evolution from Harris′s basic lot size model: Survey and research agenda. In: *International Journal of Production Economics* 155 (6), S. 16–38. DOI: 10.1016/j.ijpe.2014.01.013.

Armadani, Elvi; Jeffry, Daniel; As'adi, Muhammad; Widiatama, Yulizar (2025): Improving Inventory Control in the Electrical Sector Using Forecasting Models: A Comparative Study of ARIMA, Exponential Smoothing, Croston, and SBA. In: *Jurnal Serambi Engineering* 2025 (Vol. 10 No. 2). Online verfügbar unter https://jse.serambimekkah.id/index.php/jse/article/view/877.

Arnold, Dieter; Isermann, Heinz; Kuhn, Axel; Tempelmeier, Horst; Furmans, Kai (Hg.) (2008): Handbuch Logistik. Unter Mitarbeit von Inderfurth und Jensen. 3. Aufl. 2008. Berlin, Heidelberg: Springer Berlin Heidelberg (VDI-Buch). Online verfügbar unter http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1531831.

Arnolds, Hans; Heege, Franz; Röh, Carsten; Tussing, Werner (2016): Materialwirtschaft und Einkauf. Wiesbaden: Springer Fachmedien Wiesbaden.

Arrenberg, Jutta (2020): Wirtschaftsstatistik für Bachelor. 4., überarbeitete Auflage. München, Tübingen: UVK Verlag; Narr Francke Attempto Verlag GmbH & Co KG (utb Betriebs- und Volkswirtschaftslehre, 3914).

Axsäter, Sven (2015): Inventory Control. Cham: Springer International Publishing (225).

Babai, M. Z.; Dallery, Y. (2009): Dynamic versus static control policies in single stage production-inventory systems. In: *International Journal of Production Research* 47 (2), S. 415–433. DOI: 10.1080/00207540802426383.

Baciarello, Luca; D'Avino, Marco; Onori, Riccardo; Schiraldi, Massimiliano M. (2013): Lot Sizing Heuristics Performance. In: *International Journal of Engineering Business Management* 5 (3), S. 5. DOI: 10.5772/56004.

Bäck, Sabine; Tiefenbrunner, Martin; Grössle, Gernot (2008): Optimierung von logistischen Prozessen durch die Kombination von Simulation und neuronalen Netzen. In: Management komplexer Materialflüsse mittels Simulation. State-of-the-Art und innovative Konzepte. Wiesbaden: Gabler (Springer eBook Collection Business and Economics), S. 163–179.

Baetschmann, Gregori; Winkelmann, Rainer (2017): A dynamic hurdle model for zero-inflated count data. In: *Communications in Statistics - Theory and Methods* 46 (14), S. 7174–7187. DOI: 10.1080/03610926.2016.1146766.

Balci, Osman (1998): Verification, Validation, and Testing. In: Jerry Banks (Hg.): Handbook of Simulation: Wiley, S. 335–393.

Baltes, Oliver; Lakomy, Heribert; Spieß, Petra; Wörmann-Wiese, Elke (2017): SAP-Materialwirtschaft. Das Praxishandbuch. 1st ed. Bonn: Rheinwerk Verlag (SAP PRESS). Online verfügbar unter https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=6406098.

Bantel, Oliver (2014): Integriertes Bestandsmanagement in mehrstufigen Supply Chains. Zugl.: Köln, Univ., Diss., 2014. Aachen: Shaker-Verl. (Berichte aus der Logistik).

Barros, Júlio; Cortez, Paulo; Carvalho, M. Sameiro (2021): A systematic literature review about dimensioning safety stock under uncertainties and risks in the procurement process. In: *Operations Research Perspectives* 8 (1), S. 100192. DOI: 10.1016/j.orp.2021.100192.

Beck, Fabian G.; Grosse, Eric H.; Teßmann, Ruben (2015): An extension for dynamic lot-sizing heuristics. In: *Production & Manufacturing Research* 3 (1), S. 20–35. DOI: 10.1080/21693277.2014.985390.

Becker, Julian (2016): Dynamisches kennliniengestütztes Bestandsmanagement. Dissertation. Gottfried Wilhelm Leibniz Universität Hannover; TEWISS - Technik und Wissen GmbH.

Bergstra, James; Bengio, Yoshua (2012): Random Search for Hyper-Parameter Optimization. In: *Journal of Machine Learning Research* (13), S. 281–305.

Bertalanffy; Ludwig von (1949): Zu einer allgemeinen Systemlehre. In: *Biologia Generalis* (19), S. 114–129.

Besenfelder, Christoph (2018): Quantifizierte Bewertung des Fertigungsstrukturwandels von Produktionssystemen. Vorgehen zur Modellierung und Bewertung multipler Ausprägungen des produktionslogistischen Gestaltungsraums in bestehenden Produktionssystemen. Dissertation. Technische Universität Dortmund, Dortmund. Fakultät Maschinenbau.

Bichler, Klaus; Krohn, Ralf; Riedel, Guido; Schöppach, Frank (2010): Beschaffungs- und Lagerwirtschaft. Praxisorientierte Darstellung der Grundlagen, Technologien und Verfahren. 9., aktualisierte und überarb. Aufl. Wiesbaden: Gabler.

Biedermann, Hubert (2008): Ersatzteilmanagement. Berlin, Heidelberg: Springer Berlin Heidelberg.

Bishop, Christopher M. (2006): Pattern recognition and machine learning. New York, NY: Springer (Computer science). Online verfügbar unter http://www.loc.gov/catdir/enhancements/fy0818/2006922522-d.html.

Bishop, Christopher M. (2009): Pattern recognition and machine learning. Corrected at 8th printing 2009. New York, NY: Springer (Information science and statistics). Online verfügbar unter https://github.com/ctgk/PRML.

Black, Jason E.; Kueper, Jacqueline K.; Williamson, Tyler S. (2023): An introduction to machine learning for classification and prediction. In: *Family practice* 40 (1), S. 200–204. DOI: 10.1093/fampra/cmac104.

Blackburn, Joseph D.; Millen, Robert A. (1980): Heuristic Lot-Sizing Performance in a Rolling-Schedule Environment. In: *Decision Sciences* 11 (4), S. 691–701. DOI: 10.1111/j.1540-5915.1980.tb01170.x.

Blackburn, Joseph D.; Millen, Robert A. (1985): A methodology for predicting single-stage lot-sizing performance: Analysis and experiments. In: *Journal of Operations Management* 5 (4), S. 433–448. DOI: 10.1016/0272-6963(85)90024-5.

Bodt, Marc A. de; Gelders, Ludo F.; van Wassenhove, Luk N. (1984): Lot sizing under dynamic demand conditions: A review. In: *Engineering Costs and Production Economics* 8 (3), S. 165–187. DOI: 10.1016/0167-188x(84)90035-1.

Bonamente, Massimiliano (2017): Statistics and Analysis of Scientific Data. New York, NY: Springer New York.

Box, George E. P.; Jenkins, Gwilym M. (1970): Time series analysis. Forecasting and control. London: Holden-Day (Holden-Day series in time series analysis).

Boylan, J. E.; Syntetos, A. A.; Karakostas, G. C. (2008): Classification for forecasting and stock control: a case study. In: *Journal of the Operational Research Society* 59 (4), S. 473–481. DOI: 10.1057/palgrave.jors.2602312.

Brahimi, Nadjib; Absi, Nabil; Dauzère-Pérès, Stéphane; Nordli, Atle (2017): Single-item dynamic lot-sizing problems: An updated survey. In: *European Journal of Operational Research* 263 (3), S. 838–863. DOI: 10.1016/j.ejor.2017.05.008.

Brahm, Amira; Hadj-Alouane, Atidel B.; Sboui, Sami (2020): Dynamic and reactive optimization of physical and financial flows in the supply chain. In: *10.5267/j.ijiec*, S. 83–106. DOI: 10.5267/j.ijiec.2019.6.003.

Breiman, Leo (2001): Random Forests. In: *Machine Learning* 45 (1), S. 5–32. DOI: 10.1023/A:1010933404324.

Breiman, Leo; Friedman, Jerome H.; Olshen, Richard A.; Stone, Charles J. (2017): Classification And Regression Trees: Routledge.

Bretzke, Wolf-Rüdiger (2015): Logistische Netzwerke. Berlin, Heidelberg: Springer Berlin Heidelberg.

Büchel, Jan; Engels, Barbara (2022): Datenbewirtschaftung von Unternehmen in Deutschland. IW-Trends 1/2022. Hg. v. Institut der deutschen Wirtschaft Köln Medien GmbH. Köln. Online verfügbar unter https://ieds-projekt.de/wp-content/uploads/2022/05/IW-Trends_2022-Buechel-Engels.pdf.

Burggräf, Peter (2021): Fabrikplanung. Handbuch Produktion und Management 4. Unter Mitarbeit von Günther Schuh. 2nd ed. Berlin, Heidelberg: Springer Berlin / Heidelberg. Online verfügbar unter https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=6523258.

Burgin, T. A. (1975): The Gamma Distribution and Inventory Control. In: *Operational Research Quarterly (1970-1977)* 26 (3), S. 507. DOI: 10.2307/3008211.

Bushuev, Maxim A.; Guiffrida, Alfred; Jaber, M. Y.; Khan, Mehmood (2015): A review of inventory lot sizing review papers. In: *Management Research Review* 38 (3), S. 283–298. DOI: 10.1108/MRR-09-2013-0204.

Çalışkan, Cenk (2021): A simple derivation of the optimal solution for the EOQ model for deteriorating items with planned backorders. In: *Applied Mathematical Modelling* 89 (1), S. 1373–1381. DOI: 10.1016/j.apm.2020.08.037.

Callarman, Thomas E.; Hamrin, Robert S. (1984): A Comparison of Dynamic Lot Sizing Rules for Use in a Single Stage MRP System with Demand Uncertainty. In: *International Journal of Operations & Production Management* 4 (2), S. 39–48. DOI: 10.1108/eb054713.

CAPS RESEARCH (2021): Metrics of Supply Management Cross-Industry Report. Hg. v. Metrics of Supply Management Cross-Industry Report.

Carletta, Jean; Hirschberg, Julia (1996): Assessing Agreement on Classification Tasks: The Kappa Statistic. In: *Computational Linguistics* 22 (2), S. 249–254. Online verfügbar unter https://aclanthology.org/J96-2004/.

Carson, J. S. (2002): Model verification and validation. In: Proceedings of the Winter Simulation Conference. 2002 Winter Simulation Conference. San Diego, CA, USA, 8-11 Dec. 2002: IEEE, S. 52–58.

Casella, George; Berger, Roger L. (2002): Statistical inference. Second edition. Andover, Melbourne, Mexico City, Stamford, CT, Toronto, Hong Kong, New Delhi, Seoul, Singapore, Tokyo: CENGAGE Learning (Duxbury advanced series).

Cattani, Kyle D.; Jacobs, F. Robert; Schoenfelder, Jan (2011): Common inventory modeling assumptions that fall short: Arborescent networks, Poisson demand, and single-echelon approximations. In: *Journal of Operations Management* 29 (5), S. 488–499. DOI: 10.1016/j.jom.2010.11.008.

Chang, Xin; Kwok, Wing Chun; Wong, George (2024): Demand uncertainty, inventory, and cost structure. In: *Contemporary Accting Res* 41 (1), S. 226–254. DOI: 10.1111/1911-3846.12918.

Chen, Rung-Ching; Dewi, Christine; Huang, Su-Wen; Caraka, Rezzy Eko (2020): Selecting critical features for data classification based on machine learning methods. In: *J Big Data* 7 (1). DOI: 10.1186/s40537-020-00327-4.

Chowdhury, Abdur Razzak; Paul, Rajesh; Rozony, Farhana Zaman (2025): A Systematic Review of Demand Forecasting Models for Retail E-Commerce Enhancing Accuracy in Inventory and Delivery Planning. In: *IJSIR* 06 (01), S. 1–27. DOI: 10.63125/mbbfw637.

Churchman, Charles West; Ackoff, Russel Lincoln; Arnoff, Leonard (1957): Introduction to Operations Research. New York: John Wiley & Sons.

Cohen, Jacob (1960): A Coefficient of Agreement for Nominal Scales. In: *Educational and Psychological Measurement* 20 (1), S. 37–46. DOI: 10.1177/001316446002000104.

Cohen, Jacob (2013): Statistical Power Analysis for the Behavioral Sciences: Routledge.

Cohen, Jacob; Cohen, Patricia; West, Stephen G.; Aiken, Leona S. (2003): Applied multiple regression/correlation analysis for the behavioral sciences. Third edition. New York, London: Routledge Taylor & Francis Group. Online verfügbar unter http://www.loc.gov/catdir/enhancements/fy0634/2002072068-d.html.

Crone, Sven F. (2010): Neuronale Netze zur Prognose und Disposition im Handel. Wiesbaden: Gabler (Gabler Research : Betriebliche Forschung zur Unternehmensführung).

Dangelmaier, Wilhelm (2017): Produktionstheorie 3. Berlin, Heidelberg: Springer Berlin Heidelberg.

Dantas, Tiago Mendes; Cyrino Oliveira, Fernando Luiz; Varela Repolho, Hugo Miguel (2017): Air transportation demand forecast through Bagging Holt Winters methods. In: *Journal of Air Transport Management* 59 (4), S. 116–123. DOI: 10.1016/j.jairtraman.2016.12.006.

Das, Panchanan (2019): Econometrics in Theory and Practice. Analysis of Cross Section, Time Series and Panel Data with Stata 15. 1. Singapore: Springer Singapore Pte. Limited. Online verfügbar unter https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=5892703.

Deng, Kan; Moore, A. W.; Nechyba, M. C. (1997): Learning to recognize time series: combining ARMA models with memory-based learning. In: Proceedings 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation CIRA'97. 'Towards New Computational Principles for Robotics and Automation'. 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation CIRA'97. 'Towards New Computational Principles for Robotics and Automation'. Monterey, CA, USA, 10-11 July 1997: IEEE Comput. Soc. Press, S. 246–251.

Dietterich, Thomas G. (2000): Ensemble Methods in Machine Learning 1857, S. 1–15. DOI: 10.1007/3-540-45014-9_1.

Djunaidi, M.; Devy, B. A. R.; Setiawan, E.; Fitriadi, R. (2019): Determination of lot size orders of furniture raw materials using dynamic lot sizing method. In: *IOP Conf. Ser.: Mater. Sci. Eng.* 674 (1), S. 12050. DOI: 10.1088/1757-899X/674/1/012050.

Dunke, Fabian; Nickel, Stefan (2021): Online optimization with gradual look-ahead. In: *Oper Res Int J* 38 (6), S. 522. DOI: 10.1007/s12351-019-00506-z.

Dyckhoff, H. (2006): Produktionstheorie. Grundzuge industrieller Produktionswirtschaft. 5., uberarbeitete Aufl. Berlin, New York: Springer.

Dzakiyullah, Nur Rachman; Hussin, Burairah; Saleh, Chairul; Handani, Aditian Maytri (2014): Comparison Neural Network and Support Vector Machine for Production Quantity Prediction. In: *adv sci lett* 20 (10), S. 2129–2133. DOI: 10.1166/asl.2014.5708.

Eckstein, Peter P. (2016): Angewandte Statistik mit SPSS. Wiesbaden: Springer Fachmedien Wiesbaden.

Engelmeyer, Torben (2016): Managing Intermittent Demand. Wiesbaden: Springer Fachmedien Wiesbaden.

Favre, Liliana (2003): UML and the unified process. Hershey, Pa.: IRM.

Fildes, R.; Kingsman, B. (2011): Incorporating demand uncertainty and forecast error in supply chain planning models. In: *Journal of the Operational Research Society* 62 (3), S. 483–500. DOI: 10.1057/jors.2010.40.

Forbes, Catherine; Evans, Merran; Hastings, Ncholas; Peacock, Brian (2010): Statistical distributions. 4th ed. Oxford: Wiley-Blackwell. Online verfügbar unter http://onlinelibrary.wiley.com/book/10.1002/9780470627242.

Frees, Edward W. (2012): Regression Modeling with Actuarial and Financial Applications: Cambridge University Press.

Freitas Costa, Eduardo de; Schneider, Silvana; Carlotto, Giulia Bagatini; Cabalheiro, Tainá; Oliveira Júnior, Mauro Ribeiro de (2021): Zero-inflated-censored Weibull and gamma regression models to estimate wild boar population dispersal distance. In: *Jpn J Stat Data Sci* 4 (2), S. 1133–1155. DOI: 10.1007/s42081-021-00124-0.

Friese, Fabian (2022): Multi-kriterielle stochastische kapazitätsbeschränkte Losgrößenplanung. Wiesbaden: Springer Fachmedien Wiesbaden.

Ganas, Ioannis S.; Papachristos, Sotirios (1997): Analytical evaluation of heuristics performance for the single-level lot-sizing problem for products with constant demand. In: *International Journal of Production Economics* 48 (2), S. 129–139. DOI: 10.1016/S0925-5273(96)00074-6.

Gartner (2023): Gartner Identifies Top Trends Shaping the Future of Data Science and Machine Learning. Analysts Explore Industry Trends at Gartner Data & Analytics Summit in Sydney. Hg. v. Gartner. Sydney, Australia. Online verfügbar unter https://www.gartner.com/en/newsroom/press-releases/2023-08-01-gartner-identifies-top-trends-shaping-future-of-data-science-and-machine-learning, zuletzt geprüft am 25.10.2025.

Georgii, Hans-Otto (2009): Stochastik. Einführung in die Wahrscheinlichkeitstheorie und Statistik. 4., überarb. und erw. Aufl. Berlin: De Gruyter (De-Gruyter-Lehrbuch). Online verfügbar unter http://www.reference-global.com/isbn/978-3-11-021527-4.

Gleißner, Harald; Femerling, J. Christian (2008): Logistik. Grundlagen, Übungen, Fallbeispiele. 2., akt. u. erw. Aufl. Wiesbaden: Springer Gabler.

Glock, Christoph H.; Grosse, Eric H.; Ries, Jörg M. (2014): The lot sizing problem: A tertiary study. In: *International Journal of Production Economics* 155 (4), S. 39–51. DOI: 10.1016/j.ijpe.2013.12.009.

Gonçalves, João N.C.; Sameiro Carvalho, M.; Cortez, Paulo (2020): Operations research models and methods for safety stock determination: A review. In: *Operations Research Perspectives* 7 (8–9), S. 100164. DOI: 10.1016/j.orp.2020.100164.

Goyal, Mandeep; Mahmoud, Qusay H. (2024): A Systematic Review of Synthetic Data Generation Techniques Using Generative AI. In: *Electronics* 13 (17), S. 3509. DOI: 10.3390/electronics13173509.

Grandini, Margherita; Bagli, Enrico; Visani, Giorgio (2020): Metrics for Multi-Class Classification: an Overview.

Grinsztajn, Léo; Oyallon, Edouard; Varoquaux, Gaël (2022): Why do tree-based models still outperform deep learning on tabular data? Online verfügbar unter http://arxiv.org/pdf/2207.08815.

Grundstein, Sebastian (2017): Fertigungsregelung flexibler Fließfertigungen und Werkstattfertigungen zur Einhaltung von Lieferterminen. Dissertation. Universität Bremen.

Gu, Qiong; Zhu, Li; Cai, Zhihua (2009): Evaluation Measures of the Classification Performance of Imbalanced Data Sets 51, S. 461–471. DOI: 10.1007/978-3-642-04962-0_53.

Gutenschwager, Kai; Rabe, Markus; Spieckermann, Sven; Wenzel, Sigrid (2017): Simulation in Produktion und Logistik. Berlin, Heidelberg: Springer Berlin Heidelberg.

Hao, Jiangang; Ho, Tin Kam (2019): Machine Learning Made Easy: A Review of Scikit-learn Package in Python Programming Language. In: *Journal of Educational and Behavioral Statistics* 44 (3), S. 348–361. DOI: 10.3102/1076998619832248.

Härdler, Jürgen (Hg.) (2012): Betriebswirtschaftslehre für Ingenieure. Lehr- und Praxisbuch für Ingenieure und Wirtschaftsingenieure. 5., aktualisierte Aufl. München: Fachbuchverl. Leipzig im Hanser-Verl. (Hanser eLibrary). Online verfügbar unter http://www.hanser-elibrary.com/doi/book/10.3139/9783446442429.

Harris, F. W. (1913): How many parts to make at once. In: *The Magazine of Management* (10), S. 135.

Hartmann, Horst (2005): Materialwirtschaft. Organisation - Planung - Durchführung - Kontrolle. 9th ed. Berlin: Duncker & Humblot. Online verfügbar unter https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=5898478.

Hartmann, Horst (2017): Bestandsmanagement und -controlling. Optimierungsstrategien mit Beispielen aus der Praxis. 3. erweiterte Auflage. Gernsbach: Deutscher Betriebswirte-Verlag GmbH (Praxisreihe Einkauf Materialwirtschaft, Band 8).

Hartung, Joachim (2012): Statistik. Lehr- und Handbuch der angewandten Statistik. Munchen: Oldenbourg Wissenschaftsverlag. Online verfügbar unter https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=3046024.

Hastie, Trevor; Tibshirani, Robert; Friedman, Jerome H. (2017): The elements of statistical learning. Data mining, inference, and prediction. Second edition. New York, NY: Springer (Springer Series in Statistics). Online verfügbar unter https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=6314834.

Hecker, Dirk; Voss, Angi; Paaß, Gerhard; Wirtz, Tim (2023): Big Data 2.0 – mit synthetischen Daten KI-Systeme stärken. In: *Wirtsch Inform Manag* 15 (2), S. 161–167. DOI: 10.1365/s35764-022-00437-z.

Hernawaty; Yossy Fadly; Livia Purnama (2024): Analysis of Raw Material Inventory Control Using the EOQ Method at PT Wijaya Karya Beton Tbk 2024 (Vol. 1 (2024): 1st ICANEAT), S. 253–258.

Herrmann, Frank (2011): Operative Planung in IT-Systemen für die Produktionsplanung und -steuerung. Wirkung, Auswahl und Einstellhinweise von Verfahren und Parametern. Wiesbaden: Vieweg+Teubner Verlag / Springer Fachmedien Wiesbaden GmbH Wiesbaden.

Herrmann, Frank (2018): Übungsbuch Losbildung und Fertigungssteuerung. Aufgaben zur operativen Produktionsplanung und -steuerung. Wiesbaden: Springer Fachmedien Wiesbaden; Imprint Springer Gabler.

Hilbe, Joseph M. (2014): Modeling Count Data: Cambridge University Press.

Ho, Y. C.; Pepyne, D. L. (2002): Simple Explanation of the No-Free-Lunch Theorem and Its Implications. In: *Journal of Optimization Theory and Applications* 115 (3), S. 549–570. DOI: 10.1023/A:1021251113462.

Hofstra, Nienke; Spiliotopoulou, Eirini; Leeuw, Sander de (2024): Ordering decisions under supply uncertainty and inventory record inaccuracy: An experimental investigation. In: *Decision Sciences* 55 (3), S. 303–318. DOI: 10.1111/deci.12564.

Holderied, Cornelius (2005): Güterverkehr, Spedition und Logistik. Managementkonzepte für Güterverkehrsbetriebe, Speditionsunternehmen und logistische Dienstleister. München: Oldenbourg. Online verfügbar unter http://dx.doi.org/10.1524/9783486700299.

Hollmann, Noah; Müller, Samuel; Purucker, Lennart; Krishnakumar, Arjun; Körfer, Max; Hoo, Shi Bin et al. (2025): Accurate predictions on small data with a tabular foundation model. In: *Nature* 637 (8045), S. 319–326. DOI: 10.1038/s41586-024-08328-6.

Hoppe, Marc (2012): Bestandsoptimierung mit SAP. Mit SAP ERP und SAP SCM Bestände optimieren und Bestandskosten senken ; Prognosegenauigkeit und Planung

verbessern ; mit zahlreichen Praxisbeispielen. 3., aktualisierte und erw. Aufl. Bonn: Galileo Press (SAP betriebswirtschaftlich).

Hosen, Shahadat Mohammed; Islam, Raisul; Naeem, Zain; Folorunso, Esther; Son Chu, Thai; Al Mamun et al. (2024): Data-Driven Decision Making: Advanced Database Systems for Business Intelligence. In: *nano-ntp* 20 (S3). DOI: 10.62441/nano-ntp.v20is3.51.

Huber, Andreas; Laverentz, Klaus (2019): Logistik. 2., überarbeitete und korrigierte Auflage. München: Verlag Franz Vahlen (Vahlens Kurzlehrbücher). Online verfügbar unter https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1933263.

Hutzschenreuter, Thomas (2015): Allgemeine Betriebswirtschaftslehre. Wiesbaden: Springer Fachmedien Wiesbaden.

Hyndman, R. J.; Athanasopoulos, G. (2018): Forecasting: principles and practice. 2nd. Melbourne, Australia: OTexts. Online verfügbar unter https://otexts.com/fpp2/.

Hyndman, Rob J.; Koehler, Anne B. (2006): Another look at measures of forecast accuracy. In: *International Journal of Forecasting* 22 (4), S. 679–688. DOI: 10.1016/j.ijforecast.2006.03.001.

Ihme, Joachim (2006): Logistik im Automobilbau. Logistikkomponenten und Logistiksysteme im Fahrzeugbau. München, Wien: Hanser. Online verfügbar unter http://www.hanser-elibrary.com/doi/book/10.3139/9783446408623.

Indah, Yunna Mentari; Aristawidya, Rafika; Fitrianto, Anwar; Erfiani, Erfiani; Jumansyah, L. Risman DwiM. (2025): Comparison of Random Forest, XGBoost, and LightGBM Methods for the Human Development Index Classification. In: *Jambura J. Math* 7 (1), S. 14–18. DOI: 10.37905/jjom.v7i1.28290.

Inman, R. Anthony; Green, Kenneth W. (2022): Environmental uncertainty and supply chain performance: the effect of agility. In: *JMTM* 33 (2), S. 239–258. DOI: 10.1108/JMTM-03-2021-0097.

DIN IEC 60050-351:2013, 2014: Internationales Elektrotechnisches Wörterbuch.

Islam, Sadeka; Keung, Jacky; Lee, Kevin; Liu, Anna (2012): Empirical prediction models for adaptive resource provisioning in the cloud. In: *Future Generation Computer Systems* 28 (1), S. 155–162. DOI: 10.1016/j.future.2011.05.027.

Japkowicz, Nathalie (2011): Evaluating Learning Algorithms. A classification perspective. Cambridge, New York: Cambridge University Press. Online verfügbar unter https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=366108.

Jeunet, Jully (2006): Demand forecast accuracy and performance of inventory policies under multi-level rolling schedule environments. In: *International Journal of Production Economics* 103 (1), S. 401–419. DOI: 10.1016/j.ijpe.2005.10.003.

Jeunet, Jully; Jonard, Nicolas (2000): Measuring the performance of lot-sizing techniques in uncertain environments. In: *International Journal of Production Economics* 64 (1-3), S. 197–208. DOI: 10.1016/S0925-5273(99)00058-4.

Johnson, Steven (2004): Emergence. The connected lives of ants, brains, cities and software. 1. Scribner trade paperback ed. New York, NY: Scribner.

Jose, Victor Richmond R. (2017): Percentage and Relative Error Measures in Forecast Evaluation. In: *Operations Research* 65 (1), S. 200–211. DOI: 10.1287/opre.2016.1550.

Kalaya, P.; Termsuksawad, P.; Wasusri, T. (2019): Forecasting Lumpy Demand for Planning Inventory: The Case of Community Hospitals in Thailand, S. 686–690. DOI: 10.1109/IEEM44572.2019.8978541.

Kämmerer, Christian (2017): Fehlmengenkosten in der Distributionslogistik: Analyse und Modellierung aus Sicht der Investitionsgüterindustrie. Dissertation. Würzburg.

Kaul, Deepak; Khurana, Rahul (2022): AI-Driven Optimization Models for E-commerce Supply Chain Operations: Demand Prediction, Inventory Management, and Delivery Time Reduction with Cost Efficiency Considerations 7, S. 59–77.

Kazan, Osman; Nagi, Rakesh; Rump, Christopher M. (2000): New lot-sizing formulations for less nervous production schedules. In: *Computers & Operations Research* 27 (13), S. 1325–1345. DOI: 10.1016/S0305-0548(99)00076-3.

Kelleher, John D.; MacNamee, Brian; D'Arcy, Aoife (2015): Fundamentals of machine learning for predictive data analytics. Algorithms, worked examples, and case studies. Cambridge, Massachusetts, London, England: The MIT Press.

Kern, Werner (1996): Handwörterbuch der Produktionswirtschaft. 2., völlig neu gestaltete Aufl. Stuttgart: Schäffer-Poeschel (Enzyklopädie der Betriebswirtschaftslehre, 7).

Kersten, Wolfgang; Seiter, Mischa; See, Birgit von; Hackius, Niels; Maurer, Timo (2017): Chancen der digitalen Transformation. Trends und Strategien in Logistik und Supply Chain Management. Hamburg: DVV Media Group GmbH.

Kian, Ramez; Berk, Emre; Gürler, Ülkü; Rezazadeh, Hassan; Yazdani, Baback (2021): The effect of economies-of-scale on the performance of lot-sizing heuristics in rolling horizon basis. In: *International Journal of Production Research* 59 (8), S. 2294–2308. DOI: 10.1080/00207543.2020.1730464.

Kilger, Christoph; Wagner, Michael (2008): Demand Planning. In: Hartmut Stadtler und Christoph Kilger (Hg.): Supply Chain Management and Advanced Planning, Bd. 4. Berlin, Heidelberg: Springer Berlin Heidelberg, S. 133–160.

Kistner, Klaus-Peter; Steven, Marion (2001): Produktionsplanung. Mit 33 Tabellen. 3., vollst. überarb. Aufl. Heidelberg: Physica-Verl. (Physica-Lehrbuch).

Klaus, Peter; Krieger, Winfried; Krupp, Michael (2012): Gabler Lexikon Logistik. Management logistischer Netzwerke und Flüsse. 5. Aufl. 2012. Wiesbaden: Gabler Verlag.

Kluck, Dieter (2008): Materialwirtschaft und Logistik. Lehrbuch mit Beispielen und Kontrollfragen. 3., überarb. Aufl. Stuttgart: Schäffer-Poeschel (Praxisnahes Wirtschaftsstudium).

Knuth, Tobias (2021): Lernende Entscheidungsbäume. In: *Informatik Spektrum* 44 (5), S. 364–369. DOI: 10.1007/s00287-021-01398-0.

Korponai, János; Tóth, Ágota Bányainé; Illés, Béla (2017): The Effect of the Safety Stock on the Occurrence Probability of the Stock Shortage. In: *Management and Production Engineering Review* 8 (1), S. 69–77. DOI: 10.1515/mper-2017-0008.

Korstanje, Joos (2021): Advanced Forecasting with Python. Berkeley, CA: Apress.

Krobath, Peter Michael (2010): Die Kostenoptimale Betsellpolitik für die Beschaffung von Legiertem Halbzeug aus Stahl. Ein praktikables Verfahren für mehrfache, diskrete Preisänderungen und restriktiver Bestellemengenauswahl unter Einbezug von Servicegraderwartungen. Dissertation. Leoben.

Lan, Ting; Hu, Hui; Jiang, Chunhua; Yang, Guobin; Zhao, Zhengyu (2020): A comparative study of decision tree, random forest, and convolutional neural network for spread-F identification. In: *Advances in Space Research* 65 (8), S. 2052–2061. DOI: 10.1016/j.asr.2020.01.036.

Lasch, Rainer (2019): Strategisches und operatives Logistikmanagement: Beschaffung. Wiesbaden: Springer Fachmedien Wiesbaden.

Lau, Hon-Shiang; Lau, Amy Hing-Ling (2003): Nonrobustness of the normal approximation of lead-time demand in a (Q, R) system. In: *Naval Research Logistics* 50 (2), S. 149–166. DOI: 10.1002/nav.10053.

Law, Averill M. (2015): Simulation Modeling and Analysis. Fifth Edition. New York: McGraw-Hill US Higher Ed USE Legacy. Online verfügbar unter https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=6327699.

Lennox, B.; Rutherford, P.; Montague, G. A.; Haughin, C. (1995): A novel fault prediction technique using model degradation analysis. In: *Proceedings of 1995 American Control Conference - ACC'95*, S. 3274–3278. DOI: 10.1109/ACC.1995.532208.

Li, Song; Goel, Lalit; Wang, Peng (2016): An ensemble approach for short-term load forecasting by extreme learning machine. In: *Applied Energy* 170, S. 22–29. DOI: 10.1016/j.apenergy.2016.02.114.

Liao, Xiaoqun; Cao, Nanlan; Li, Ma; Kang, Xiaofan (2019): Research on Short-Term Load Forecasting Using XGBoost Based on Similar Days. In: *2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, S. 675–678. DOI: 10.1109/ICITBS.2019.00167.

Lödding, Hermann (2016): Verfahren der Fertigungssteuerung. Grundlagen, Beschreibung, Konfiguration. 3. Auflage. Berlin: Springer Vieweg (VDI-Buch).

Luhmann, Niklas (2018): Soziale Systeme. Grundriß einer allgemeinen Theorie. 17. Auflage. Frankfurt am Main: Suhrkamp (Suhrkamp-Taschenbuch Wissenschaft, 666).

Lukesch, Maximilian; Kellner, Florian (2019): Übungsbuch Produktionswirtschaft. Berlin, Heidelberg: Springer Berlin Heidelberg; Imprint Springer Gabler.

Ma, Xiyuan; Rossi, Roberto; Archibald, Thomas (2019): Stochastic Inventory Control: A Literature Review. In: *IFAC-PapersOnLine* 52 (13), S. 1490–1495. DOI: 10.1016/j.ifacol.2019.11.410.

Maimon, Oded; Rokach, Lior (2005): Data Mining and Knowledge Discovery Handbook. New York: Springer-Verlag.

Melzer-Ridinger, Ruth (2009): Materialwirtschaft und Einkauf. Beschaffungsmanagement. 5., unverändte Aufl. Berlin, Boston: Oldenbourg Wissenschaftsverlag.

Mertens, Peter; Rässler, Susanne (2012): Prognoserechnung. Siebte, wesentlich überarbeitete und erweiterte Auflage. Berlin, Heidelberg: Physica-Verlag.

Meyer, Dennis (2024): Entwicklung eines Vorgehensmodells zur Einführung von KI im Einkauf. Dissertation.

Meyer, Jan Christoph; Sander, Ulrich (Hg.) (2009): Bestände senken, Lieferservice steigern - Ansatzpunkt Bestandsmanagement. 2., durchges. Aufl. Aachen: FIR (Praxis-Edition, Bd. 12).

Mitchell, Tom M. (1997): Machine learning. International ed., [Reprint.]. New York, NY: McGraw-Hill (McGraw-Hill series in computer science).

Mittag, Hans-Joachim (2017): Statistik. Eine Einführung mit interaktiven Elementen. 5., wesentlich überarbeitete Auflage. Berlin: Springer Spektrum (Lehrbuch).

VDI 4465, 2021: Modellierung und Simulation Modellbildungsprozess.

Mohd-Lair, Noor Ajian; Pang, Chuan Kian; Liew, Willey Y.H.; Semui, Hardy; Yew, Loh Zhia (2013): An EOQ Based Multi-Storage Location of Spare Part Inventories: A Case Study. In: *AMM* 315, S. 733–738. DOI: 10.4028/www.scientific.net/AMM.315.733.

Monteiro, Paulo; Lino, José; Araújo, Rui Esteves; Costa, Louelson (2024): Comparison between LightGBM and other ML algorithms in PV fault classification. In: *EAI Endorsed Trans Energy Web* 11. DOI: 10.4108/ew.4865.

Mumm, Mirja (2015): Kosten- und Leistungsrechnung. Berlin, Heidelberg: Springer Berlin Heidelberg.

Olaniyi, Oluwadamilare Abiodu; Pugal, Paul Sundar (2024): Optimising Inventory Management Strategies for Cost Reduction in Supply Chains: A Systematic Review. In: *JAB* 10 (1), S. 48–55. DOI: 10.31289/jab.v10i1.11678.

Ortmann, Christian; Siebeking, Ingo (2000): Heuristiken zur Losgrößenplanung in PPS-Systemen. Prämierte Diplomarbeit. Universität Osnabrück, Osnabrück. Fachbereich Wirtschaftswissenschaften.

Parmezan, Antonio Rafael Sabino; Souza, Vinicius M.A.; Batista, Gustavo E.A.P.A. (2019): Evaluation of statistical and machine learning models for time series prediction: Identifying the state-of-the-art and the best conditions for the use of each model. In: *Information Sciences* 484 (5–6), S. 302–337. DOI: 10.1016/j.ins.2019.01.076.

Patzak, Gerold (1982): Systemtechnik - Planung komplexer innovativer Systeme. Grundlagen, Methoden, Techniken. Berlin, Heidelberg: Springer Berlin Heidelberg. Online verfügbar unter http://dx.doi.org/10.1007/978-3-642-81893-6.

Pawellek, Günther (2013): Integrierte Instandhaltung und Ersatzteillogistik. Vorgehensweisen, Methoden, Tools. Berlin, Heidelberg: Springer (Springer eBook Collection).

Pedregosa, Fabian; Varoquaux, Gaël; Gramfort, Alexandre; Michel, Vincent; Thirion, Bertrand; Grisel, Olivier et al. (2011): Scikit-learn: Machine Learning in Python. Online verfügbar unter http://arxiv.org/pdf/1201.0490v4.

Peffers, Ken; Tuunanen, Tuure; Rothenberger, Marcus A.; Chatterjee, Samir (2007): A Design Science Research Methodology for Information Systems Research. In: *Journal of Management Information Systems* 24 (3), S. 45–77. DOI: 10.2753/MIS0742-1222240302.

Petropoulos, Fotios; Apiletti, Daniele; Assimakopoulos, Vassilios; Babai, Mohamed Zied; Barrow, Devon K.; Ben Taieb, Souhaib et al. (2022): Forecasting: theory and practice. In: *International Journal of Forecasting* 38 (3), S. 705–871. DOI: 10.1016/j.ijforecast.2021.11.001.

Qin, Tao (2020): Dual Learning. Singapore: Springer Singapore Pte. Limited. Online verfügbar unter https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=6396086.

Quinlan, J. R. (1986): Induction of decision trees. In: *Mach Learn* 1 (1), S. 81–106. DOI: 10.1007/BF00116251.

Rabe, Markus; Spieckermann, Sven; Wenzel, Sigrid (2008): Verifikation und Validierung für die Simulation in Produktion und Logistik. Vorgehensmodelle und Techniken. 1. Aufl. Berlin, Heidelberg: Springer. Online verfügbar unter https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=364359.

Ramaekers, Katrien; Janssens, Gerrit K. (2008): On the choice of a demand distribution for inventory management models. In: *EJIE* 2 (4), Artikel 18441, S. 479. DOI: 10.1504/EJIE.2008.018441.

Ramaekers, Katrien; Janssens, Gerrit K. (2014): Optimal policies for demand Optimal policies for demand forecasting and inventory management of goods with intermittent demand. In: *Journal of Applied Operational Research* 2014, 2014 (6 (2)), S. 111–123.

Rashedi, Jonas (2022): Das datengetriebene Unternehmen. Erfolgreiche Implementierung einer data-driven Organization. Wiesbaden, Heidelberg: Springer Gabler.

Rathipriya, R.; Abdul Rahman, Abdul Aziz; Dhamodharavadhani, S.; Meero, Abdelrhman; Yoganandan, G. (2023): Demand forecasting model for time-series pharmaceutical data using shallow and deep neural network model. In: *Neural computing & applications* 35 (2), S. 1945–1957. DOI: 10.1007/s00521-022-07889-9.

REFA (1985): Methodenlehre der Planung und Steuerung. 4. Aufl. München: Hanser.

Richards, Gwynne; Grinsted, Susan (2016): The logistics and supply chain toolkit. Second edition. London, Philadelphia, New Delhi: Kogan Page.

Robinson, D. (2017): The incredible growth of Python. https://stackoverflow.blog/2017/09/06/incredible-growth-python/, zuletzt geprüft am 27.08.2025.

Rokach, Lior (2010): Ensemble-based classifiers. In: *Artif Intell Rev* 33 (1-2), S. 1–39. DOI: 10.1007/s10462-009-9124-7.

Rolf, Benjamin; Jackson, Ilya; Müller, Marcel; Lang, Sebastian; Reggelin, Tobias; Ivanov, Dmitry (2023): A review on reinforcement learning algorithms and applications in supply chain management. In: *International Journal of Production Research* 61 (20), S. 7151–7179. DOI: 10.1080/00207543.2022.2140221.

Ross, David Frederick (2018): Distribution planning and control. Managing in the era of supply chain management. Third edition, corrected publication. New York, Heidelberg,

Dordrecht, London: Springer. Online verfügbar unter http://gbv.eblib.com/patron/FullRecord.aspx?p=2094422.

Rueda, Fernando Moya; Fink, Gernot A. (2020): From Human Pose to On-Body Devices for Human-Activity Recognition. In: 2020 25th International Conference on Pattern Recognition (ICPR). 2020 25th International Conference on Pattern Recognition (ICPR). Milan, Italy, 10.01.2021 - 15.01.2021: IEEE, S. 10066–10073.

Rufo, Derara Duba; Debelee, Taye Girma; Ibenthal, Achim; Negera, Worku Gachena (2021): Diagnosis of Diabetes Mellitus Using Gradient Boosting Machine (LightGBM). In: *Diagnostics (Basel, Switzerland)* 11 (9). DOI: 10.3390/diagnostics11091714.

Sahin, Funda; Narayanan, Arunachalam; Robinson, E. Powell (2013): Rolling horizon planning in supply chains: review, implications and directions for future research. In: *International Journal of Production Research* 51 (18), S. 5413–5436. DOI: 10.1080/00207543.2013.775523.

SAP (2022): Verbrauchsgesteuerte Disposition (MM-CBP). Optimierende Losgrößenverfahren. Hg. v. SAP. online. Online verfügbar unter https://help.sap.com/saphelp_me60/helpdata/de/da/97b6535fe6b74ce10000000a174cb4/frameset.htm, zuletzt geprüft am 26.01.2022.

Sargent, Robert (2010): Proceedings of the 2010 Winter Simulation Conference (WSC). 5 - 8 Dec. 2010, Baltimore, MD, USA ; [incorporating the] MASM (Modeling and Analysis for Semiconductor Manufacturing) Conference. Online verfügbar unter http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=5672636.

Schenk, Michael; Müller, Egon; Wirth, Siegfried (2014): Fabrikplanung und Fabrikbetrieb. Methoden für die wandlungsfähige, vernetzte und ressourceneffiziente Fabrik. 2., vollst. überarb. u. erw. Aufl. 2014. Berlin, Heidelberg: Springer Vieweg (SpringerLink Bücher). Online verfügbar unter http://dx.doi.org/10.1007/978-3-642-05459-4.

Schlittgen, Rainer; Streitberg, Bernd H. J. (2001): Zeitreihenanalyse. 9., unwesentlich veränd. Aufl. München, Oldenbourg: Oldenbourg (Lehr- und Handbücher der Statistik). Online verfügbar unter http://dx.doi.org/10.1524/9783486710960.

Schmidt, Klaus-Jürgen; Schützdeller, Klaus; Gröner, Lothar; Venitz, Markus; Zeilinger, Peter; Skrowronek, Gerhard et al. (1993): Logistik. Grundlagen, Konzepte, Realisierung. Hg. v. Klaus-Jürgen Schmidt. Wiesbaden: Vieweg+Teubner Verlag (Springer eBook Collection Computer Science and Engineering).

Schmidt, Matthias; Nyhuis, Peter (2021): Produktionsplanung und -steuerung im Hannoveraner Lieferkettenmodell. Berlin, Heidelberg: Springer Berlin Heidelberg.

Schmidt, Michael (2018): Distribution Center Design Process. Dissertation, Dortmund.

Schneeweiß, Christoph (1981): Modellierung industrieller Lagerhaltungssysteme. Einführung und Fallstudien. Berlin, Heidelberg: Springer Berlin Heidelberg. Online verfügbar unter http://dx.doi.org/10.1007/978-3-642-67901-8.

Schönsleben, Paul (2016): Integral Logistics Management. Operations and Supply Chain Management Within and Across Companies, Fifth Edition. 5th ed. Boca Raton: CRC Press. Online verfügbar unter http://gbv.eblib.com/patron/FullRecord.aspx?p=4711495.

Schönsleben, Paul (2024): Bestandsmanagement und stochastisches Materialmanagement. In: Paul Schönsleben (Hg.): Handbuch Integrales Logistikmanagement. Berlin, Heidelberg: Springer Berlin Heidelberg, S. 467–512.

Schuh, Günther (2006): Produktionsplanung und -steuerung. Grundlagen, Gestaltung und Konzepte. 3., völlig neu bearb. Aufl. Berlin: Springer (VDI-Buch).

Schuh, Günther; Schmidt, Carsten (2014): Produktionsmanagement. Berlin, Heidelberg: Springer Berlin Heidelberg.

Schuh, Günther; Stich, Volker (2013): Logistikmanagement. Berlin, Heidelberg: Springer Berlin Heidelberg.

Schulte, Gerd (2001): Material- und Logistikmanagement. 2., wesentlich erw. und verb. Aufl. München, Wien: Oldenbourg.

Schulz, Thomas; Ayaz, BorisJohannes Kröckel; Huber, Marco; Kröckel, Johannes; Ingold, Remo; Oppermann, Henrik et al. (Hg.) (2022): Analytics in der Industrie. Schlüsseltechnologie für die digitale Transformation. Vogel Business Media. 1. Auflage. Würzburg: Vogel Communications Group. Online verfügbar unter https://www.content-select.com/index.php?id=bib_view&ean=9783834362933.

scikit-learn (2025a): Scikit documentation - 1. Supervised learning. 1.10. Decision Trees. Online verfügbar unter https://scikit-learn.org/stable/modules/tree.html#tree-algorithms.

scikit-learn (2025b): Scikit documentation - 1. Supervised learning. 1.11. Ensembles: Gradient boosting, random forests, bagging, voting, stacking. Online verfügbar unter https://scikit-learn.org/stable/modules/ensemble.html.

Seeck, Stephan (2010): Erfolgsfaktor Logistik. Klassische Fehler erkennen und vermeiden. 1. Aufl. Wiesbaden: Gabler.

Sharma, Himani; Kumar, Sunil (2016): A Survey on Decision Tree Algorithms of Classification in Data Mining. In: *IJSR* 5.

Sheskin, David J. (2020): Handbook of Parametric and Nonparametric Statistical Procedures: Chapman and Hall/CRC.

Silver, Edward A.; Pyke, David F.; Thomas, Douglas J. (2016): Inventory and Production Management in Supply Chains: CRC Press.

Simpson, N. C. (2001): Questioning the relative virtues of dynamic lot sizing rules. In: *Computers & Operations Research* 28 (9), S. 899–914. DOI: 10.1016/S0305-0548(00)00015-0.

Snyder, R. D. (1984): Inventory control with the gamma probability distribution. In: *European Journal of Operational Research* 17 (3), S. 373–381. DOI: 10.1016/0377-2217(84)90133-4.

Sokolova, Marina; Lapalme, Guy (2009): A systematic analysis of performance measures for classification tasks. In: *Information Processing & Management* 45 (4), S. 427–437. DOI: 10.1016/j.ipm.2009.03.002.

Stadtler, Hartmut (Hg.) (2010): Supply Chain Management und Advanced Planning. Konzepte, Modelle und Software. Berlin, Heidelberg: Springer.

Stadtler, Hartmut (2023): An improved Groff criterion for myopic lot sizing in both regular and sporadic demand. 5. Aufl. Hg. v. Wolfgang Brüggemann, Malte Fliedner, Knut Haase und Guido Voigt. Universität Hamburg. Hamburg (Operations & Supply Chain Management).

Stadtler, Hartmut; Kilger, Christoph; Meyr, Herbert (2015): Supply Chain Management and Advanced Planning. Berlin, Heidelberg: Springer Berlin Heidelberg.

Steurer, Elmar (1996): Prognose von 15 Zeitreihen der DGOR mit Neuronalen Netzen. In: *OR Spektrum* 18 (2), S. 117–125. DOI: 10.1007/BF01539737.

Stölzle, Wolfgang; Heusler, Klaus Felix; Karrer, Michael (2004): Erfolgsfaktor Bestandsmanagement. Konzept, Anwendung, Perspektiven. 1. Aufl. Zürich: Versus Verlag.

Sun, Yige; Li, Jing; Xu, Yifan; Zhang, Tingting; Wang, Xiaofeng (2023): Deep learning versus conventional methods for missing data imputation: A review and comparative study. In: *Expert Systems with Applications* 227, S. 120201. DOI: 10.1016/j.eswa.2023.120201.

Sutton, Richard S.; Barto, Andrew (2020): Reinforcement learning. An introduction. Second edition. Cambridge, Massachusetts, London, England: The MIT Press (Adaptive computation and machine learning).

Syntetos, A. A.; Boylan, J. E.; Croston, J. D. (2005): On the categorization of demand patterns. In: *Journal of the Operational Research Society* 56 (5), S. 495–503. DOI: 10.1057/palgrave.jors.2601841.

Tadikamalla, Pandu R. (1984): A comparison of several approximations to the lead time demand distribution. In: *Omega* 12 (6), S. 575–581. DOI: 10.1016/0305-0483(84)90060-4.

Tallón-Ballesteros, Antonio J.; Riquelme, José C. (2014): Data Mining Methods Applied to a Digital Forensics Task for Supervised Machine Learning 555, S. 413–428. DOI: 10.1007/978-3-319-05885-6_17.

Tempelmeier, Horst (2006): Material-Logistik. Modelle und Algorithmen für die Produktionsplanung und -steuerung in Advanced Planning-Systemen ; mit 127 Tabellen. 6., neubearb. Aufl. Berlin, Heidelberg: Springer. Online verfügbar unter http://deposit.dnb.de/cgi-bin/dokserv?id=2681056&prov=M&dok_var=1&dok_ext=htm.

Tempelmeier, Horst (2016): Supply Chain Management und Produktion. Übungen und Mini-Fallstudien. Norderstedt: Books on Demand.

Tempelmeier, Horst (2018): Bestandsmanagement in Supply Chains. 6., wesentlich erweiterte und verbesserte Auflage. Norderstedt: BoD - Books on Demand.

Tempelmeier, Horst; Bantel, Oliver (2015): Integrated optimization of safety stock and transportation capacity. In: *European Journal of Operational Research* 247 (1), S. 101–112. DOI: 10.1016/j.ejor.2015.05.069.

The pandas development team (2025): pandas-dev/pandas: Pandas: Zenodo.

Thomas A Caswell; Antony Lee; Elliott Sales de Andrade; Michael Droettboom; Tim Hoffmann; Jody Klymak et al. (2023): matplotlib/matplotlib: REL: v3.7.1: Zenodo.

Thommen, Jean-Paul; Achleitner, Ann-Kristin; Gilbert, Dirk Ulrich; Hachmeister, Dirk; Jarchow, Svenja; Kaiser, Gernot (2023): Allgemeine Betriebswirtschaftslehre. Umfassende Einführung aus managementorientierter Sicht. 10., überarbeitete und aktualisierte Auflage. Wiesbaden, Heidelberg: Springer Gabler (Lehrbuch). Online verfügbar unter http://www.springer.com/.

Thommen, Jean-Paul; Achleitner, Ann-Kristin; Gilbert, Dirk Ulrich; Hachmeister, Dirk; Kaiser, Gernot (2016): Allgemeine Betriebswirtschaftslehre. Umfassende Einführung aus managementorientierter Sicht / Jean-Paul Thommen, Ann-Kristin Achleitner, Dirk Ulrich Gilbert, Dirk Hachmeister, Gernot Kaiser. Eighth edition. Wiesbaden: Springer Gabler.

Thomopoulos, Nick T. (2018): Probability distributions. With truncated, log and bivariate extensions. Cham: Springer.

Thonemann, Ulrich (2015): Operations management. Konzepte, Methoden und Anwendungen. 3., aktualisierte Auflage. Hallbergmoos: Pearson (Always learning). Online verfügbar unter https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=5583783.

Thorstenson, A. (2018): As Simple as Possible but no Simpler-An Inquiry into Approximations for a Re-order Point Inventory Control Model with Gamma-distributed Demand. In: 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). Bangkok, 16.12.2018 - 19.12.2018: IEEE, S. 153–157.

Tiedtke, Jürgen R. (2007): Allgemeine BWL. Betriebswirtschaftliches Wissen für kaufmännische Berufe \u2014 Schritt für Schritt. 2., überarbeitet Auflage. Wiesbaden: Betriebswirtschaftlicher Verlag Dr. Th. Gabler | GWV Fachverlage GmbH Wiesbaden.

Tirkolaee, Erfan Babaee; Sadeghi, Saeid; Mooseloo, Farzaneh Mansoori; Vandchali, Hadi Rezaei; Aeini, Samira (2021): Application of Machine Learning in Supply Chain Management: A Comprehensive Overview of the Main Areas. In: *Mathematical Problems in Engineering* 2021, S. 1–14. DOI: 10.1155/2021/1476043.

Trigg, D. W. (1964): Monitoring a Forecasting System. In: *OR* 15 (3), S. 271. DOI: 10.2307/3007215.

Tsoumakas, Grigorios; Katakis, Ioannis; Vlahavas, Ioannis (2010): Mining Multi-label Data. In: *in Data mining and knowledge discovery handbook*, S. 667–685. DOI: 10.1007/978-0-387-09823-4_34.

Tu, Jun; Zhou, Guofu (2004): Data-generating process uncertainty: What difference does it make in portfolio decisions? In: *Journal of Financial Economics* 72 (2), S. 385–421. DOI: 10.1016/j.jfineco.2003.05.003.

Ulrich, Hans (1982): Anwendungsorientierte Wissenschaft. Vol. 36, No. 1. Nomos Verlagsgesellschaft mbH (Die Unternehmung).

United States Census Bureau (2025): Manufacturing and Trade Inventories and Sales: July 2025. Hg. v. United States Census Bureau. United States Census Bureau. Online verfügbar unter https://www.census.gov/mtis/current/index.html.

Vahrenkamp, Richard (2008): Produktionsmanagement. Unter Mitarbeit von Christoph Siepermann. 6., überarbeitete Auflage. München: Oldenbourg Verlag. Online verfügbar unter http://dx.doi.org/10.1524/9783486848243.

van Randen, Hendrik Jan; Bercker, Christian; Fieml, Julian (2016): Einführung in UML. Wiesbaden: Springer Fachmedien Wiesbaden.

Vania, Abigail; Yolina, Hanni (2021): Analysis Inventory Cost Jona Shop with EOQ Model. In: *EMACS Journal* 3 (1), S. 21–25. DOI: 10.21512/emacsjournal.v3i1.6847.

VDI 3633 (2014): Simulation von Logistik-, Materialfluss und Produktionssystemen. Blatt 1: Grundlagen ICS 03.100.10.

Veinott, Arthur F. (1966): The Status of Mathematical Inventory Theory. In: *Management Science* 12 (11), S. 745–777. DOI: 10.1287/mnsc.12.11.745.

Vidal, Germán Herrera (2023): Deterministic and Stochastic İnventory Models in Production Systems: a Review of the Literature. In: *Process Integr Optim Sustain* 7 (1-2), S. 29–50. DOI: 10.1007/s41660-022-00299-3.

Vogel, Jürgen (2015): Prognose von Zeitreihen. Wiesbaden: Springer Fachmedien Wiesbaden.

Vojdani, Nina; Knop, Mathias (2014): Adaptive Materialbereitstellung in flexiblen Produktionssystemen auf Grundlage einer agentenbasierten Transportsteuerung. In: *Logistics Journal : Proceedings* (Vol. 2014). DOI: 10.2195/lj_Proc_vojdani_de_201411_01.

Wagner, Harvey M.; Whitin, Thomson M. (1958): Dynamic Version of the Economic Lot Size Model. In: *Management Science* 5 (1), S. 89–96. DOI: 10.1287/mnsc.5.1.89.

Wagner, Michael (2003): Bestandsmanagement in Produktions- und Distributionssystemen. Zugl.: Augsburg, Univ., Diss., 2003. Aachen: Shaker (Berichte aus der Betriebswirtschaft).

Wang, Dehua; Zhang, Yang; Zhao, Yi (2017): LightGBM: An Effective miRNA Classification Method in Breast Cancer Patients. In: Proceedings of the 2017 International Conference on Computational Biology and Bioinformatics. ICCBB 2017: 2017 International Conference on Computational Biology and Bioinformatics. Newark NJ USA, 18 10 2017 20 10 2017. New York, NY, USA: ACM, S. 7–11.

Wannenwetsch, Helmut (2010): Integrierte Materialwirtschaft und Logistik. Beschaffung, Logistik, Materialwirtschaft und Produktion. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg (Springer-Lehrbuch).

Warwas, Dennis (2015): Dimensionierung logistischer Reserven. Siegen, Universität Siegen, Diss., 2014. Universitätsbibliothek der Universität Siegen, Siegen. Online verfügbar unter http://dokumentix.ub.uni-siegen.de/opus/volltexte/2015/868/.

Waskom, Michael (2021): seaborn: statistical data visualization. In: *JOSS* 6 (60), S. 3021. DOI: 10.21105/joss.03021.

Wassermann, Otto; Schwarzer, Michael (2012): Das intelligente Unternehmen. Schlummernde Potenziale realisieren. 6. Aufl. 2012. Berlin, Heidelberg: Springer Berlin Heidelberg (VDI-Buch). Online verfügbar unter http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1564093.

Weber, Jürgen (2012): Logistikkostenrechnung. Berlin, Heidelberg: Springer Berlin Heidelberg.

Wemmerlöv, Urban (1989): The behavior of lot-sizing procedures in the presence of forecast errors. In: *Journal of Operations Management* 8 (1), S. 37–47. DOI: 10.1016/S0272-6963(89)80004-X.

Wemmerlöv, Urban; WHYBARK, D. CLAY (1984): Lot-sizing under uncertainty in a rolling schedule environment. In: *International Journal of Production Research* 22 (3), S. 467–484.

Wenzel, Sigrid; Weiß, Matthias; Collisi-Böhmer, Simone; Pitsch, Holger; Rose, Oliver (2008): Qualitätskriterien für die Simulation in Produktion und Logistik. Planung und Durchführung von Simulationsstudien. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg (SpringerLink Bücher). Online verfügbar unter http://dx.doi.org/10.1007/978-3-540-35276-1.

Westkämper, Engelbert (2006): Einführung in die Organisation der Produktion. Berlin, Heidelberg: Springer Berlin Heidelberg (SpringerLink Bücher).

Wiendahl, Hans-Peter (2014): Betriebsorganisation für Ingenieure. Mit 3 Tabellen. 8., überarb. Aufl. München: Hanser (Hanser eLibrary).

Wiener, Norbert; Trawny, Peter (Hg.) (2022): Mensch und Menschmaschine. Mit einem Vorwort von Peter Trawny. Nomos eLibrary (Online service). 1. Auflage 2022. Frankfurt am Main: Klostermann (Klostermann Rote Reihe, 147). Online verfügbar unter https://philportal.de/resolver/v1/ebook?doi=10.5771/9783465145998.

Winters, Peter R. (1960): Forecasting Sales by Exponentially Weighted Moving Averages. In: *Management Science* 6 (3), S. 324–342. DOI: 10.1287/mnsc.6.3.324.

Wirtz, Markus Antonius; Nachtigall, Christof (2012): Deskriptive Statistik. Statistische Methoden für Psychologen Teil 1. 6., überarbeitete Aufl. Weinheim: Beltz (Juventa Paperback). Online verfügbar unter http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1113834.

Witten, Ian H.; Frank, Eibe; Hall, Mark A.; Pal, Christopher J. (2017): Data mining. Practical machine learning tools and techniques. Fourth edition. Amsterdam, Boston, Heidelberg, London, New York, Oxford, Paris, San Diego, San Francisco, Singapore, Sydney, Tokyo: Elsevier Morgan Kaufmann (Morgan Kaufmann series in data management systems). Online verfügbar unter https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=4708912.

Wolf, Alexander (2021): Produktionsplanung und -steuerung mit SAP S/4HANA. Das Praxishandbuch. Unter Mitarbeit von Christoph Sting. 1st ed. Bonn: Rheinwerk Verlag. Online verfügbar unter https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=6529730.

Wolpert, D. H.; Macready, W. G. (1997): No free lunch theorems for optimization. In: *IEEE Trans. Evol. Computat.* 1 (1), S. 67–82. DOI: 10.1109/4235.585893.

Wuttke, Alexander; Hunker, Joachim; Scheidler, Anne Antonia; Rabe, Markus (2022): Synthetic Demand Generation with Seasonality for Data Mining on a Data-Farmed Data Basis of a Two-Echelon Supply Chain. In: *Procedia Computer Science* 204, S. 226–234. DOI: 10.1016/j.procs.2022.08.027.

Yang, Xin-She; He, Xing-Shi; Fan, Qin-Wei (2020): Mathematical framework for algorithm analysis. In: Nature-Inspired Computation and Swarm Intelligence: Elsevier, S. 89–108.

Yeh, Quey-Jen; Chang, T.-P.; Chang, H.-C. (1997): An inventory control model with gamma distribution. In: *Microelectronics Reliability* 37 (8), S. 1197–1201. DOI: 10.1016/S0026-2714(96)00295-8.

Yılmaz Yalçıner, Ayten (2021): Determination of the Cost-Effective Lot-Sizing Technique for Perishable Goods: A Case Study. In: *International Journal of Management and Administration* 5 (9), S. 33–46.

Yuniasih, Annisa Wirrdiana; A'yuni, Nur Rohmah Lufti (2024): Literature Review of Inventory with Probabilistic Economic Order Quantity (EOQ). In: *JTM* 22 (1), S. 83–92. DOI: 10.52330/jtm.v22i1.220.

Zelewski, Stephan; Hohmann, Susanne; Hügens, Torben; Peters, Malte L. (Hg.) (2008): Produktionsplanungs- und -steuerungssysteme: Oldenbourg Wissenschaftsverlag GmbH.

Zhao, Qiankun; Cai, Ximing; Li, Yu (2019): Determining Inflow Forecast Horizon for Reservoir Operation. In: *Water Resour. Res.* 55 (5), S. 4066–4081. DOI: 10.1029/2019WR025226.

Zhou, Zhi-Hua (2012): Ensemble Methods: Chapman and Hall/CRC.

Zoller, K.; Robrade, A. (1987): Dynamische Bestellmengen - und Losgrößenplanung. Verfahrensübersicht und Vergleich. In: *OR Spektrum* 1987, 1987 (9), S. 219–233.