

Forschungsberichte aus dem  
**wbk** Institut für Produktionstechnik  
Karlsruher Institut für Technologie (KIT)

Florian Stamer

**Dynamische Lieferzeit-Preisgestaltung  
in variantenreicher Produktion**  
Ein adaptiver Ansatz mithilfe von  
Reinforcement Learning

Band 260

Forschungsberichte aus dem  
wbk Institut für Produktionstechnik  
Karlsruher Institut für Technologie (KIT)

Hrsg.: Prof. Dr.-Ing. Jürgen Fleischer  
Prof. Dr.-Ing. Gisela Lanza  
Prof. Dr.-Ing. habil. Volker Schulze

Florian Stamer

**Dynamische Lieferzeit-Preisgestaltung in  
variantenreicher Produktion**  
Ein adaptiver Ansatz mithilfe von Reinforcement  
Learning

Band 260



**Dynamische Lieferzeit-Preisgestaltung in  
variantenreicher Produktion  
Ein adaptiver Ansatz mithilfe von Reinforcement Learning**

Zur Erlangung des akademischen Grades eines  
**Doktors der Ingenieurwissenschaften (Dr.-Ing.)**  
von der KIT-Fakultät für Maschinenbau des  
Karlsruher Instituts für Technologie (KIT)

angenommene

**Dissertation**

von

M.Sc. Florian Stamer

Tag der mündlichen Prüfung: 11.07.2022

Hauptreferent: Prof. Dr.-Ing. Gisela Lanza  
Korreferent: Prof. Dr.-Ing. Kai Furmans  
Korreferent: Prof. Dr. Torbjörn Netland

**Bibliografische Information der Deutschen Nationalbibliothek**

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

Zugl.: Karlsruhe, Karlsruher Institut für Technologie, Diss., 2022

Copyright Shaker Verlag 2022

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany.

ISBN 978-3-8440-8803-8

ISSN 0724-4967

Shaker Verlag GmbH • Am Langen Graben 15a • 52353 Düren  
Telefon: 02421 / 99 0 11 - 0 • Telefax: 02421 / 99 0 11 - 9  
Internet: [www.shaker.de](http://www.shaker.de) • E-Mail: [info@shaker.de](mailto:info@shaker.de)

## **Vorwort des Herausgebers**

Die schnelle und effiziente Umsetzung innovativer Technologien wird vor dem Hintergrund der Globalisierung der Wirtschaft der entscheidende Wirtschaftsfaktor für produzierende Unternehmen. Universitäten können als "Wertschöpfungspartner" einen wesentlichen Beitrag zur Wettbewerbsfähigkeit der Industrie leisten, indem sie wissenschaftliche Grundlagen sowie neue Methoden und Technologien erarbeiten und aktiv den Umsetzungsprozess in die praktische Anwendung unterstützen.

Vor diesem Hintergrund soll im Rahmen dieser Schriftenreihe über aktuelle Forschungsergebnisse des Instituts für Produktionstechnik (wbk) am Karlsruher Institut für Technologie (KIT) berichtet werden. Unsere Forschungsarbeiten beschäftigen sich sowohl mit der Leistungssteigerung von Fertigungsverfahren und zugehörigen Werkzeugmaschinen- und Handhabungstechnologien als auch mit der ganzheitlichen Betrachtung und Optimierung des gesamten Produktionssystems. Hierbei werden jeweils technologische wie auch organisatorische Aspekte betrachtet.

Prof. Dr.-Ing. Jürgen Fleischer

Prof. Dr.-Ing. Gisela Lanza

Prof. Dr.-Ing. habil. Volker Schulze



## **Vorwort des Verfassers**

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am wbk Institut für Produktionstechnik des Karlsruher Instituts für Technologie (KIT).

Mein besonderer Dank gilt Frau Prof. Dr.-Ing. Gisela Lanza für die Betreuung meiner wissenschaftlichen Arbeit als Hauptreferentin und das mir entgegengebrachte Vertrauen in den vergangenen Jahren. Ihre immerwährende Förderung ermöglichte nicht nur diese Arbeit, sondern auch meine persönliche Weiterentwicklung. Weiter danke ich Herrn Prof. Dr. Torbjørn Netland und Herrn Prof. Dr.-Ing. Kai Furmans für ihr Interesse an meiner Arbeit und der Übernahme des Korreferats sowie Herrn Prof. Dr. rer. nat. Frank Gauterin für den Prüfungsvorsitz.

Allen Kolleginnen und Kollegen des wbk in Karlsruhe, insbesondere im Bereich PRO, danke ich für die fortwährende Unterstützung und den kollegialen Austausch in den vergangenen Jahren. Namentlich seien an dieser Stelle Rainer Silbernagel, Sina Peukert, Shun Yang, Martin Benfer, Leonard Overbeck und Bastian Verhaelen erwähnt. Die gemeinsame Zeit wird mir immer in guter Erinnerung bleiben. Für das Korrekturlesen dieser Arbeit danke ich Sina Peukert, Bastian Verhaelen, Marvin May und Rainer Silbernagel. Besonders danken möchte ich außerdem Matthias Henzi, Fabian Ochs, Alexander Werle, Peter Lorenz und Paul Bizenberger.

Meinen Eltern Peter und Inge Stamer sowie meinen Geschwistern Felix Stamer, Julia Stamer und Philine Stamer danke ich von Herzen für ihre immerwährende Unterstützung und Rücksichtnahme - ohne Euch stünde ich nicht da, wo ich heute stehe.

Karlsruhe, im Juli 2022

Florian Stamer



## **Zusammenfassung**

Das Wertschöpfungsnetzwerk ist heute zur dominierenden Wertschöpfungsstruktur im Bereich der Produktion geworden. Für produzierende Unternehmen in solchen Netzwerken ist das Auftragsmanagement eine zentrale Aufgabe, die sich in einem Umfeld wachsender Herausforderungen befindet. Zum einen setzen kürzere Produktlebenszyklen und zunehmende Individualisierung die Unternehmen unter Druck. In diesem Zusammenhang hat die variantenreiche Produktion an Bedeutung gewonnen. Andererseits steigen die Wahrscheinlichkeit und die Auswirkungen von Störungen, insbesondere in Netzwerken, was die Leistungsfähigkeit von Produktionssystemen herausfordert.

Ein vielversprechender Ansatz in diesem Zusammenhang ist die Verwendung einer dynamischen Lieferzeit-Preisgestaltung über eine kontinuierliche Preis-Lieferzeit-Funktion. Durch den Einsatz einer dynamischen Lieferzeit-Preisgestaltung könnte es möglich sein, die Nachfrage im Netz mit den Fähigkeiten des Produktionssystems in Einklang zu bringen. Die Idee ist, Nachfragespitzen zu verschieben und die Kunden entsprechend ihrer individuellen Präferenzen zu beliefern. Auf diese Weise kann der Gesamtgewinn gesteigert werden, obwohl die Kapazitäten konstant bleiben.

Um die aufgestellten Hypothesen zu untersuchen, wird in dieser Arbeit eine Methodik zur dynamischen Lieferzeit-Preisgestaltung für die variantenreiche Produktion entwickelt. In einem ersten Schritt werden Leistungskennzahlen und eine Zielfunktion festgelegt. In einem zweiten Schritt wird ein Actor Critic Reinforcement-Learning-Verfahren als Grundlage für ein Lösungsmodell motiviert. Für das Training und das spätere Testen des resultierenden Lernagenten wird ein Simulationsmodell als Umgebung entwickelt und ein Verfahren zur Durchführung und Auswertung der Experimente ausgearbeitet.

Der Ansatz wurde im Rahmen der BMBF-Forschungsprojekte ReKoNeT und BaSys4SupplyQ diskutiert und angewendet. Die Ergebnisse aus zwei Anwendungsfällen zeigen, dass der Ansatz in der Lage ist, bei gleicher Kapazität deutlich höhere Gewinne zu erzielen und Nachfragespitzen in gewissem Umfang zu glätten. Gleichzeitig zeigen die Ergebnisse, dass das Einsparen von Kapazitäten keine dominante Strategie zu sein scheint. Dies spiegelt die anwendungsspezifische Berücksichtigung verschiedener Ziele wie Service, Kapazitätskosten und Marge wider.



## Abstract

Today, the value network has become the dominant value creation structure in the area of production. For manufacturing companies in such networks the order management is a central task which faces an environment of growing challenges. On the one hand, shorter product life cycles and increasing individualisation are putting companies under pressure. In this context, multi-variant production has gained in importance. On the other hand, the probability and impact of disruptions are increasing, especially in networks, which challenges the capabilities of production systems.

A promising approach in this context is the use of dynamic pricing via a continuous price-lead-time function. By using a dynamic pricing it would be possible to bring the demands in the network in line with the capabilities of the production system. The idea is to shift demand peaks and to supply customers according to their individual preferences. This way, the overall profit can be increased although the capacities remain constant.

To achieve and examine the described effects, the aim of this thesis is to develop a methodology for dynamic pricing for the make to order and assemble to order production. In the first step, key performance indicators and an objective function are established. Second, an actor critic reinforcement learning method is motivated, since reinforcement learning has already proven its potential to handle complex optimisation tasks efficiently in comparable domains. For the training and the later testing of the resulting learning agent, a simulation model is developed as an environment and a procedure for the execution and evaluation of the experiments is worked out.

The approach has been discussed and applied within the BMBF research projects ReKoNeT and BaSys4SupplyQ. The results from two use cases show that the approach is able to achieve significantly higher profits with the same capacity and to smooth demand peaks to a certain extent. At the same time, the results indicate that saving on capacities does not seem to be a dominant strategy. Instead, it is reasonable to maximise service for selected customers in order to justify higher margins. This reflects the application-specific consideration of various targets, such as service, capacity costs and margin. Thus, the results also reveal the central importance of aligning the parameterisation of the method with the strategy and goals of the applying company.



# Inhaltsverzeichnis

<b>Inhaltsverzeichnis</b>	<b>I</b>
<b>Abkürzungsverzeichnis</b>	<b>VI</b>
<b>Formelverzeichnis</b>	<b>VII</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Ausgangslage	1
1.2 Forschungsfrage	3
1.3 Zielsetzung	3
1.4 Aufbau der Arbeit	5
<b>2 Grundlagen</b>	<b>6</b>
2.1 Variantenreiche Produktion	6
2.1.1 Allgemeine Grundlagen zur Produktion	6
2.1.2 Grundlagen der variantenreichen Produktion	8
2.2 Ablaufplanung und Auftragsmanagement als Teil der Produktionsplanung und -steuerung	9
2.2.1 Ablaufplanung als Teil der Produktionsplanung und -steuerung	10
2.2.2 Auftragsmanagement als Teil der Produktionsplanung und -steuerung	12
2.3 Preisgestaltung und Kundenverhalten	13
2.3.1 Grundlagen der dynamischen Lieferzeit-Preisgestaltung	14
2.3.2 Modellierung von Kundenverhalten mit Entscheidungstheorie	16
2.4 Simulation im Produktionsumfeld	18
2.4.1 Grundlagen der Simulation im Produktionsumfeld	18
2.4.2 Statistische Versuchsplanung und Tests in einer Simulationsstudie	19
2.5 Methoden für die dynamische Lieferzeit-Preisgestaltung	22

2.6 Reinforcement Learning als Methode zur dynamischen Lieferzeit-Preisgestaltung	24
2.6.1 Grundlagen des Reinforcement Learning	24
2.6.2 Neuronale Netze zur Funktionsapproximation	28
2.6.3 Klassifikation von Reinforcement Learning Verfahren	31
2.6.4 Actor Critic Verfahren	32
<b>3 Stand der Forschung</b>	<b>35</b>
3.1 Anforderungen an den Lösungsansatz	35
3.2 Vorstellung der Lösungsansätze	36
3.2.1 Ansätze mit Reinforcement Learning	36
3.2.2 Ansätze mit analytischen Lösungsverfahren	38
3.2.3 Ansätze mit Entscheidungsbaumverfahren	41
3.2.4 Ansätze mit Heuristiken	44
3.2.5 Sonstige Ansätze zur dynamischen Lieferzeit-Preisgestaltung	46
3.3 Ableitung des Forschungsdefizits	48
<b>4 Überblick des Lösungsansatzes</b>	<b>52</b>
<b>5 Dynamische Lieferzeit-Preisgestaltung für die variantenreiche Produktion</b>	<b>55</b>
5.1 Zielsystem und Reinforcement Learning Modell	55
5.1.1 Erläuterung des Zielsystems	55
5.1.2 Reinforcement Learning Lösungsmodell	60
5.2 Umweltmodul	62
5.2.1 Netzwerkebene	63
5.2.2 Standortebene	75
5.3 Agentenmodul	84
5.3.1 Ableitung der Belohnungsfunktion für den Belohnungswert <i>Rt</i>	85
5.3.2 Beobachtbarer Zustand <i>St</i> der Umwelt	88

5.3.3	Aktion <i>At</i> des Agenten	93
5.3.4	Konzeptionelle Architektur von Actor und Critic	98
5.3.5	Trainingsprozess von Actor und Critic	100
5.4	<b>Versuchsmodul</b>	101
5.4.1	Funktionsweise und Aufbau des Versuchsmoduls	102
5.4.2	Hyperparameteroptimierung	104
5.4.3	Hypothesen zum Testen des Ansatzes hinsichtlich des Potenzials	109
5.5	<b>Implementierung, Nutzung und Verifikation</b>	111
5.5.1	Implementierung	111
5.5.2	Nutzung	114
5.5.3	Verifikation	114
<b>6</b>	<b>Erprobung des Lösungsansatzes</b>	<b>116</b>
6.1	Anwendung im Bereich der Hochpräzisionsgewichte	116
6.1.1	Vorstellung des Anwendungsfalls	116
6.1.2	Exemplarische Anwendung und Ergebnisse	122
6.2	Anwendung im Bereich der Pneumatikzylinderproduktion	139
6.2.1	Vorstellung des Anwendungsfalls	140
6.2.2	Exemplarische Anwendung und Ergebnisse	144
<b>7</b>	<b>Diskussion, kritische Würdigung und Ausblick</b>	<b>153</b>
7.1	Diskussion der Ergebnisse	153
7.2	Kritische Würdigung	155
7.3	Ausblick	158
<b>8</b>	<b>Zusammenfassung</b>	<b>161</b>
<b>9</b>	<b>Liste der eigenen Veröffentlichungen</b>	<b>163</b>
<b>10</b>	<b>Literaturverzeichnis</b>	<b>165</b>
<b>Abbildungsverzeichnis</b>		<b>I</b>

<b>Tabellenverzeichnis</b>		<b>IV</b>
<b>Anhang</b>		<b>IX</b>
A1	Ausführungen zu den Grundlagen der Simulation im Produktionsumfeld	IX
A1.1	Weiterführende Grundlagen zu Simulation im Produktionsumfeld	IX
A1.2	Vorgehensmodell zur Durchführung einer Simulationsstudie	XI
A2	Beispiel einer Exponentialverteilung als Bestellhäufigkeit	XIII
A3	Beispiel einer Poissonverteilung als Bestellmenge	XIII
A4	Beispiel einer Dreiecksverteilung als Fertigungszeit	XIV
A5	Algorithmus zur Ablaufplanung	XIV
A6	Ergänzung zur Bestimmung des Nutzenmaximums	XVI
A7	Erläuterungen zur Wahl der zufallsbasierten Suche	XVI
A8	Überlegungen zur Softwarearchitektur	XVII
A9	Erläuterungen zur Konfigurationsdatei	XXI
A10	Beispiel für ein Hochpräzisionsgewicht nach Norm	XXV
A11	Normtabelle der OIML	XXVI
A12	Prozessdiagramm für E1-Gewichte	XXVII
A13	Prozessdiagramm für E2-Gewichte	XXVIII
A14	Prozessdiagramm für F1-Gewichte	XXIX
A15	Prozessdiagramm für F2-Gewichte	XXX
A16	Prozessdiagramm für M-Gewichte	XXXI
A17	Beispiel für einen Fertigungsplan für Anwendungsfall 1	XXXII
A18	Fertigungszeiten am Beispiel E2 für Anwendungsfall 1	XXXIII
A19	Statistische Details – statische Lieferzeit – Szenario 1 – Anwendungsfall 1	XXXV
A20	Statistische Details – flexible Lieferzeit – Szenario 1 – Anwendungsfall 1	XXXVI

---

A21	Statistische Anwendungsfall 1	Details	–	statische	Lieferzeit	–	Szenario	2	–	XXXVII
A22	Statistische Anwendungsfall 1	Details	–	flexible	Lieferzeit	–	Szenario	2	–	XXXVIII
A23	Statistische Anwendungsfall 1	Details	–	statische	Lieferzeit	–	Szenario	3	–	XXXIX
A24	Statistische Anwendungsfall 1	Details	–	flexible	Lieferzeit	–	Szenario	3	–	XL
A25	Beispiel für einen Pneumatikzylinder								XLI	
A26	Herleitung der verkaufsrelevanten Daten im zweiten Anwendungsfall								XLII	
A27	Bestimmung der Bestellmenge einer Bestellung in Anwendungsfall 2								XLIII	
A28	Fertigungszeiten für den Pneumatikzylinder								XLIV	
A29	Herleitung der Maschinenanzahl								XLV	
A30	Maschineninvestitionskosten								XLVI	
A31	Ausführungen zur Potenzialanalyse in Anwendungsfall 2								XLVII	
A32	Statistische Anwendungsfall 2	Details	–	statische	Lieferzeit	–	Szenario	1	–	LII
A33	Statistische Anwendungsfall 2	Details	–	flexible	Lieferzeit	–	Szenario	1	–	LIII
A34	Statistische Anwendungsfall 2	Details	–	statische	Lieferzeit	–	Szenario	2	–	LIV
A35	Statistische Anwendungsfall 2	Details	–	flexible	Lieferzeit	–	Szenario	2	–	LV
A36	Statistische Anwendungsfall 2	Details	–	statische	Lieferzeit	–	Szenario	3	–	LVI
A37	Statistische Anwendungsfall 2	Details	–	flexible	Lieferzeit	–	Szenario	3	–	LVII

## Abkürzungsverzeichnis

Abkürzung	Beschreibung
ATO	Assemble-To-Order
ATP	Available-To-Promise
BMBF	Bundesministerium für Bildung und Forschung
BTO	Build-To-Order
CTO	Configure-To-Order
EDD	Earliest Due Date
ETO	Engineer-To-Order
FCFS	First Come, First Serve
MTO	Make-To-Order
MTS	Make-To-Stock
OIML	Organisation Internationale de Métrologie Légale
OPP	Order Penetration Point
PPS	Produktionsplanung und -steuerung
ReLU	Rectified Linear Unit
SPT	Shortest Processing Time

## Formelverzeichnis

Formelzeichen	Bedeutung	Einheit
$a$	Eine spezifische Aktion.	-
$A$	Menge der Aktionen.	-
$a_{i,v}$	Erster, freier Parameter der Nutzenfunktion. Beeinflusst die rechtsseitige Kurvensteilheit.	-
$a_{skaliert}$	Skalierte Aktion des Agenten.	-
$A_t$	Eine spezifische Aktion zum Zeitpunkt t.	-
$a_\Delta$	Untere Intervallgrenze einer Dreiecksverteilung für die Fertigungszeit einer Station für ein Produkt.	Minute
$b_f$	Bearbeitungszeitfaktor abhängig vom Automatisierungsgrad.	-
$\bar{b}_f$	Mittlerer Bearbeitungszeitfaktor.	-
$BH_{i,j}$	Produktvariantenspezifische Bestellhäufigkeit pro Kunde.	-
$b_{i,v}$	Zweiter, freier Parameter der Nutzenfunktion. Stellt den Hochpunkt der Funktion bzw. die präferierte Lieferzeit in Tagen ein.	-
$BM_{i,j}$	Produktvariantenspezifisch Bestellmenge pro Kunde.	-
$b_\Delta$	Obere Intervallgrenze einer Dreiecksverteilung für die Fertigungszeit einer Station für ein Produkt.	Minute
$c_{i,v}$	Dritter, freier Parameter der Nutzenfunktion. Beeinflusst die linksseitige Kurvensteilheit.	-
$c_\Delta$	Häufigster Wert einer Dreiecksverteilung für die Fertigungszeit einer Station für ein Produkt.	Minute
$\bar{D}$	Differenz der Gewinnmittelwerte, erhoben aus einer Stichprobe.	€
$d_{aufmaß}$	Aufmaß auf den Zerspandurchmesser.	mm
$d_{i,v}$	Vierter, freier Parameter der Nutzenfunktion. Stellt die Preisbereitschaft des Kunden ein.	€
$d_{i,v}^u$	Begrenzter Preisbereitschaftsfaktor als Substitut für $d_{i,v}$ .	€
$d_j$	Mittlerer Durchmesser eines spezifischen Produkts.	mm

---

$D_t$	Differenz zweier Gewinnwerte, jeweils erhoben aus einer Stichprobe bei gleichen Bedingungen.	€
$E[\cdot]$	Erwartungswert einer Zufallsvariable.	-
$E_{\pi}[\cdot]$	Erwartungswert einer Zufallsvariable bei gegebener Politik $\pi$ .	-
$e_{\text{datum}}$	Zeitstempel, welcher das Datum der Bestellung wiedergibt.	-
$F_{\text{exp}}(x_{\text{exp}})$	Exponentialverteilungsfunktion zur Modellierung der Bestellhäufigkeit.	-
$F_{\text{plan},j}$	Fertigungsplan in Form eines azyklischen Graphen für ein Produkt $j$ .	-
$F_{\text{pois}}(n_{\text{pois}})$	Poissonverteilungsfunktion zur Modellierung der Bestellmenge pro Bestellung.	-
$f_{\text{rechenzeit}}$	Übersetzungsfaktor von Simulationszeit zu realer Rechenzeit.	-
$fz_{j,q,S_{\text{typ}}}$	Fertigungszeit eines spezifischen Produkts $j$ in Zustand $q$ auf Stationstyp $S_{\text{typ}}$ .	Minute
$F_{\Delta}(x_{\Delta})$	Dreiecksverteilungsfunktion zur Modellierung von Fertigungszeiten an Stationen.	Minute
$g$	Spezifischer Gewinn.	€
$\bar{G}$	Mittelwert des Gewinns.	€
$G_t$	Ertrag zum Zeitpunkt $t$ .	€
$H_0$	Nullhypothese.	-
$H_1$	Alternativhypothese.	-
$h_{i,v}$	Binärvariable, ob ein Auftrag eines Kunden mit bestimmter Lieferzeit angenommen wurde.	-
$i$	ID eines Kunden.	-
$I$	Grundgesamtheit aller Kunden.	-
$i^*$	ID eines Kunden, der eine spezifische Variante $v$ nachfragt.	-
$i_{\text{one-hot}}$	ID eines Kunden One-Hot-kodiert.	-
$j$	ID eines einzelnen Produkts.	-
$k$	Eine Zahl aus der Menge der natürlichen Zahlen.	-
$k_0$	Konstante für den zu testenden Erwartungswert des Hypothesenpaares.	-

$k_{abbau}$	Abbaurate in einer exponentiellen Lernratenabbaufunktion.	-
$k_{fix}$	Summe aller Fixkosten im Rahmen einer Unternehmensgewinnrechnung (vgl. Tabelle 5-1).	€
$k_{Instand}$	Instandhaltungskosten.	€
$K_{i,v}$	Erster, freier Parameter der Preisfunktion. Beeinflusst die linksseitige Kurvensteilheit der Lieferzeit-Preis-Funktion.	-
$k_{i,v}$	Kunden- und auftragsspezifische Kosten der Auftragsbearbeitung – zum Beispiel Material- und Maschinenkosten.	€
$k_{Kapital}$	Kapitalkosten.	€
$k_{pois}$	Iteratorvariable der Poissonverteilung.	-
$k_{var}$	Variable Kosten im Rahmen einer Unternehmensgewinnrechnung.	€
$k_{Verfrühung,i,v}$	Lagerkosten bei zu früher Produktion.	€
$k_{Verspätung,i,v}$	Kosten aus Vertragsstrafen bzw. Verspätungskosten.	€
$L_{i,v}$	Zweiter, freier Parameter der Preisfunktion. Stellt ein, zu welcher Lieferzeit der niedrigste Preis angeboten wird.	-
$l_j$	Länge eines spezifischen Produkts.	mm
$Loss_{Critic}$	Verlustfunktion des Critic.	-
$M_{i,v}$	Dritter, freier Parameter der Preisfunktion. Beeinflusst die rechtsseitige Kurvensteilheit der Lieferzeit-Preis-Funktion.	-
$max$	Der Maximalwert der Stichprobe bzw. der betrachteten Zahlenmenge.	-
$m_{i,v}$	Kunden- und auftragsspezifische Bestellmenge.	-
$min$	Der Minimalwert der Stichprobe bzw. der betrachteten Zahlenmenge.	-
$MS_{typ}$	Die Menge aller Stationstypen.	-
$N(\mu; \sigma)$	Normalverteilung, welche mit einem Erwartungswert und einer Standardabweichung definiert wird.	-
$n_{aa}$	Anzahl entgangener bzw. abgelehnter Aufträge.	-
$n_{bf}$	Anzahl der Bearbeitungszeitfaktoren.	-

$NF_{i,j}$	Nachfrage eines Kunden für ein spezifisches Produkt.	-
$n_{freigrad}$	Freiheitsgrade der Stichprobe. Im Kontext dieser Arbeit ergibt sich dieser zu $n_{stichprobe} - 1$ .	-
$N_{i,v}$	Vierter, freier Parameter der Preisfunktion. Beeinflusst den permanenten, Preisaufschlag zum statischen Basispreis.	€
$N_{i,v}^l$	Begrenzter Preisaufschlagsfaktor als Substitut für $N$ .	€
$n_{konf}$	Anzahl (zu prüfender) Konfigurationen.	-
$n_{pois}$	Variable in der Poisson-Verteilungsfunktion, welche angibt bis zu welcher Höhe an Bestellungen die Wahrscheinlichkeit berechnet werden soll.	-
$n_{schonstation}$	Eingesparte Stationen	-
$n_{stichprobe}$	Anzahl an Stichproben (Stichprobenumfang).	-
$n_{verspätet}$	Anzahl verspäteter Aufträge.	-
$n_{wdh}$	Anzahl Wiederholungen mit neuen Zufallszahlen	-
$O$	Menge aller Stationen.	-
$o$	Eine Station aus der Menge aller Stationen $O$ .	-
$OG$	Obere Grenze zur Skalierung der Aktion.	-
$O_{Styp}$	Menge aller Stationen mit einem Stationstyp $S_{typ}$ .	-
$p^*$	Statischer Basispreis eines Produkts. Entspricht typischerweise einem Listenpreis.	€
$p(s'   s, a)$	Übergangswahrscheinlichkeitsfunktion, welche die Wahrscheinlichkeit angibt, in den Zustand $s'$ zu wechseln, sofern $s$ und $a$ gegeben sind.	-
$pa$	Ein einzelner Produktionsauftrag.	-
$pa_{aktuell,o}$	Der Auftrag, der aktuell auf einer betrachteten Station $o$ in Bearbeitung ist.	-
$PA_v$	Menge aller Produktionsaufträge, die zu einem (Kunden-)Auftrag $v$ gehören.	-
$p_{i,v}$	Kunden- und auftragsspezifischer Preis mit auftragsspezifischer Lieferzeit für eine Einheit.	€
$p_{i,v}(m_{i,v}, x_{i,v})$	Eine kunden- und auftragsspezifische Preisfunktion, abhängig von Menge $m_{i,v}$ sowie kundenauftrags- und variantenspezifischer Lieferzeit $x_{i,v}$ .	€

$P_{q_0, q_1}$	Übergangswahrscheinlichkeit für ein Produktionsauftrag im Zustand $q_0$ in den Zustand $q_1$ .	-
$p_{rel, i, v}$	Relativer, kunden- und auftragsspezifischer Preis- aufschlag für spezifische Lieferzeit.	-
$q$	Produktionsauftragszustand.	-
$q_{\pi}(s, a)$	Zustand-Aktion-Wertfunktion. Gibt den erwarteten Ertrag bei gegebener Politik $\pi$ , Zustand $s$ und Aktion $a$ wieder.	-
$R$	Menge aller Belohnungswerte.	-
$r$	Ein spezifischer Belohnungswert. Typischerweise definiert durch $R_t$ .	-
$r(s, a)$	Belohnungsfunktion in Abhängigkeit eines Zustandes $s$ und einer Aktion $a$ .	-
$R_t$	Belohnungswert zum Zeitpunkt $t$ .	-
$S$	Menge der Umweltzustände.	-
$s$	Spezifischer Umweltzustand, welcher $S_t$ entspricht	-
$s'$	Spezifischer Umweltzustand, welcher $S_{t+1}$ entspricht.	-
$sg$	Servicegrad.	-
$sig$	Die Sigmoid-Funktion (vgl. Kapitel 2.6.2).	-
$S_{standort}$	Referenz zum Produktionsstandort.	-
$S_t$	Umweltzustand zum Zeitpunkt $t$ .	-
$S_{typ}$	Der Stationstyp einer Station.	-
$t$	Freie Zeitvariable. Definiert einen spezifischen Zeitpunkt.	Minute
$tanh$	Tangens-Hyperbolicus-Funktion.	-
$T^*$	Teststatistik des t-Tests.	-
$t_{i,pa}$	Verbleibende Durchlaufzeit des Produktionsauftrags.	Minute
$T_{Jahr}$	Gibt die Menge aller Zeitpunkte eines vergangenen Jahres ab Betrachtungszeitpunkt an.	
$t_{la}$	Summe der Lieferterminabweichungen.	Minute
$t_{Liefer, plan, i, v}$	Gibt die geplanten Lieferzeit eines Auftrags wieder.	Minute
$t_{Liefer, real, i, v}$	Gibt die reale Lieferzeit eines Auftrags wieder.	Minute

---

$t_{n-1,1-\alpha}$	t-Wert bei $n$ Datenpunkten bzw. $n - 1$ Freiheitsgraden und $\alpha$ Signifikanzniveau.	-
$t_{ref,haupt}$	Zerspanhauptzeit für das Referenzbauteil.	Minute
$t_{ref,neben}$	Zerspannebenzeit für das Referenzbauteil.	Minute
$t_{schlupf,pa,o}$	Schlupfzeit (des einzuplanenden Auftrags).	Minute
$t_{schlupf,pa,o}^-$	Schlupfzeit des vorherigen Auftrags.	Minute
$t_{schlupf,pa,o}^+$	Schlupfzeit des nachfolgenden Auftrags.	Minute
$T_{simmin}$	Zu simulierende Zeit.	Minute
$t_{fz,min}$	Minimale Fertigungszeit im Produktionssystem für einen neuen Auftrag auf der Skala von null bis eins skaliert. Bei stochastischer Planung ist dieser Wert unsicherheitsbehaftet.	Minute
$T_{Tag}$	Gibt die Menge aller Zeitpunkte eines vergangenen Tages ab Betrachtungszeitpunkt an.	-
$t_{verfrüht}$	Summe der Zeit aller verfrühten Aufträge.	Minute
$t_{verpätet}$	Summe der Zeit aller verspäteten Aufträge.	Minute
$T_{versuch}$	Dauer der Versuchsdurchführung (Rechenzeit).	Minute
$t_{verteil}$	Verteilzeit.	Minute
$u(x)$	Nutzenfunktion in Abhängigkeit der Alternative $x$ .	-
$U(UG; OG)$	Gleichverteilung mit einer unteren und oberen Grenze	
$UG$	Untere Grenze.	-
$u_i$	Nutzenfunktion eines spezifischen Kunden.	-
$v$	Ein einzelner Auftrag.	-
$V$	Menge aller Aufträge.	-
$V_{44mm}$	Referenzvolumen des Beispielwerkstücks mit einem Durchmesser von 44mm.	$mm^3$
$V_j$	Zerspanvolumen eines spezifischen Produkts.	$mm^3$
$V_{ref}$	Referenzzerspanvolumen.	$mm^3$
$\vec{w}$	Gewichtungsvektor. Im Kontext eines neuronalen Netzwerks handelt es sich hierbei um Gewichtungsparameter der Neuronen.	-
$w_t$	Gewichtungsparameter zum Zeitpunkt $t$ . Im Kontext eines neuronalen Netzwerks handelt es sich hierbei um Gewichtungsparameter der Neuronen.	-

---

$W_1$	Gewichtungsfaktor des Bestrafungsterms für entgangene Aufträge.	-
$W_2$	Gewichtungsfaktor des Bestrafungsterms für Verfrühungen.	-
$W_3$	Gewichtungsfaktor für Verspätungen, die sich bereits bei der Einplanung ergeben.	-
$W_4$	Gewichtungsfaktor des Belohnungsterms für verhandelte Preisaufschläge.	-
$W_5$	Bestrafungsfaktor für eingetretene Verspätungen in der Belohnungsfunktion des Agenten.	-
$W_6$	Gewichtungsfaktor für mögliche Kapazitätseinsparungen durch nicht Nutzung für einen Tag.	-
$W_7$	Gewichtungsfaktor für realisierte Kapazitätseinsparungen durch nicht Nutzung für ein Jahr.	-
$W_{schlange,o}$	Warteschlange an einer Station $o$ .	-
$x^i$	Alternative $i$ im Kontext einer Nutzenfunktion.	-
$x^j$	Alternative $j$ im Kontext einer Nutzenfunktion.	-
$x_{exp}$	Parameter für die Exponentialverteilung.	Minute
$x_{i,v}$	Der kundenspezifische und auftragsspezifische Lieferzeitpunkt.	Tag
$x_{i,v}^{opt}$	Der optimale, kundenspezifische Lieferzeitpunkt gegeben eine Lieferzeit-Nutzen-Funktion.	Tag
$x_{rnd}$	Zufallszahl aus der Gleichverteilung $[0; 1]$ .	-
$x_\Delta$	Parameter der Dreiecksverteilung zur Bestimmung einer Fertigungszeit.	Minute
$y_{o,t}$	Variable, die 1 annimmt, wenn Station $o$ zum Zeitpunkt $t$ verwendet, sonst 0.	-
$z$	Zielfunktionswert.	
$\alpha^w$	Lernrate des Critic.	-
$\alpha^\theta$	Lernrate des Actors.	-
$\alpha_t$	Lernrate zum Zeitpunkt $t$ .	-
$\gamma$	Diskontierungsfaktor.	-
$\overline{\Delta G}$	Mittelwert der Gewinndifferenz von Agent und Vergleichssystem.	€
$\delta_t$	Der „Vorteil“ (engl.: Advantage) bzw. die Differenz, die sich aus approximiertem und realisiertem Ertrag ergibt.	-

$\Delta x_{i,v}$	Abweichung von der vereinbarten, kundenspezifischen Lieferzeit.	Tag
$\varepsilon$	Wert, welcher sehr viel kleiner als 1 und größer als 0 ist.	-
$\vec{\theta}$	Politik-Parameter. Im Kontext eines neuronalen Netzwerks handelt es sich hierbei um Gewichtungsparameter der Neuronen.	-
$\theta_t$	Politik-Parameter zum Zeitpunkt $t$ des Actor. Im Kontext eines neuronalen Netzwerks handelt es sich hierbei um Gewichtungsparameter der Neuronen.	-
$\lambda_{exp}$	Erwartungswert der Bestellhäufigkeit als Teil einer Exponentialverteilung.	1/Tag
$\lambda_{pois}$	Erwartungswert und Varianz der Bestellmenge als Teil einer Poissonverteilung.	-
$\lambda_{pois,i}$	Erwartungswert und Varianz der Bestellmenge als Teil einer Poissonverteilung für einen spezifischen Kunden $i$ .	-
$\lambda_{pois,v}$	Erwartungswert und Varianz der Bestellmenge als Teil einer Poissonverteilung für ein spezifisches Produkt bzw. eine spezifische Produktvariante $v$ pro Monat.	-
$\mu$	Allgemeines Formelzeichen für den Erwartungswert einer Grundgesamtheit.	-
$\mu_D$	Erwartungswert einer Differenz in der Grundgesamtheit.	-
$\mu_{\Delta G}$	Erwartungswert der spezifischen Gewinndifferenz von Agent und Vergleichssystem.	€
$\nu$	Substitution für $e^{x_{i,v}}$ .	-
$v_\pi(s)$	Zustandswertfunktion, welche den erwarteten Ertrag angibt, der unter der gegebenen Politik $\pi$ und dem Zustand $s$ noch erreicht werden kann.	-
$\hat{v}_\pi(s)$	Approximation der Zustandswertfunktion (Schätzer) in Abhängigkeit des Umweltzustandes zum Zeitpunkt $t$ unter der gegebenen Politik $\pi$ .	-
$\hat{v}_{\pi,L}(s)$	Approximation der Zustandswertfunktion für die Politik $\pi$ zur Bestimmung des Parameters $L_{i,v}$ .	-
$\hat{v}_{\pi,N}(s)$	Approximation der Zustandswertfunktion für die Politik $\pi$ zur Bestimmung des Parameters $N_{i,v}$ .	-

$\hat{v}_\pi(S_t, \vec{w})$	Approximation der Zustandswertfunktion (Schätzer) in Abhängigkeit des Umweltzustandes zum Zeitpunkt $t$ und den Gewichtungsparametern eines neuronalen Netzes unter der gegebenen Politik $\pi$ .	-
$\pi(a   s)$	Die Politik des Agenten. Sie gibt die Wahrscheinlichkeit für eine Aktion $a$ gegeben ein Zustand $s$ wieder.	-
$\sigma$	Standardabweichung.	-
$\sigma_g$	Empirische Standardabweichung des Gewinns.	$\epsilon$
$\hat{\sigma}_D$	Empirischer Standardfehler einer Differenz.	$\epsilon$
$\hat{\sigma}_{\Delta G}$	Empirischer Standardfehler der Gewinndifferenz.	$\epsilon$



# 1 Einleitung

Wirtschaftlicher Erfolg wird wesentlich durch die Effizienz der zugrundeliegenden Wertschöpfungsstrukturen bestimmt (acatech 2017, S. 15; Wu, Chuang & Hsu 2014, S. 122). Im Bereich der Produktion ist das Wertschöpfungsnetzwerk zur dominierenden Wertschöpfungsstruktur geworden (Lanza & Ferdows et al. 2019, S. 823). Die Effizienz eines Wertschöpfungsnetzwerks wird wesentlich beeinflusst durch das Auftragsmanagement jedes beteiligten Unternehmens. Das Auftragsmanagement entscheidet über die Annahme von Aufträgen und deren Konditionen, also insbesondere auch die Lieferzeit und den Preis (Schuh, Brandenburg & Cuber 2012, S. 61f). Hierbei muss das Auftragsmanagement in externer Sicht Änderungen der Auftragslage im Netzwerk berücksichtigen. Gleichzeitig muss es in interner Sicht die aktuellen Fähigkeiten und Möglichkeiten des Produktionssystems zur Fertigung von Produkten kennen, um die Aufträge mit der Produktion in Einklang zu bringen. Dem Auftragsmanagement wird somit bei der Lieferzeit-Preisgestaltung eine stark integrative Aufgabe zuteil (Schuh, Brandenburg & Cuber 2012, S. 60–65; Ivanov, Tsipoulanidis & Schönberger 2019, S. 527f).

## 1.1 Ausgangslage

Das Auftragsmanagement agiert im Umfeld wachsender Herausforderungen. Die Herausforderungen können hierbei von zwei Seiten betrachtet werden.

Auf der einen Seite setzen kürzer werdende Produktlebenszyklen und weiter steigende Individualisierung der Produkte die Unternehmen unter Druck (Lanza & Nyhuis et al. 2018, S. 6; Slamanig 2011, S. 124–126). In der Folge kam es bereits zu einer stärkeren Kundenorientierung und damit zu einer Erhöhung der Variantenzahl bei gleichzeitig kürzer werdenden Lieferzeiten (Piller 2008, S. 127f). Aus Auftrags- und Lagerfertigern wurden somit zunehmend Variantenfertiger mit ausgewähltem Produktpool (Helmig, Schmidt & Kompa 2012, S. 205f; Piller 2008, S. 127f). Die Individualisierung nimmt noch weiter zu, sodass sich die Variantenfertigung weiterentwickeln muss. Fraglich ist daher, wie die kundenindividuellen, heterogenen Anforderungen noch besser bedient werden können. Eine reine Erhöhung der Variantenzahl kann keine Lösung sein, weil jede Variantenverdopplung bisher zu einem Kostenanstieg von 20-35 % führte (Bertagnoli 2018, S. 11). Dies widerspricht der wirtschaftlichen Notwendigkeit, Komplexitäts- und Flexibilitätskosten unter Kontrolle zu halten.

Auf der anderen Seite steigen auch die Wahrscheinlichkeit und der Einfluss von Störungen der Produktion, auf welche Unternehmen reagieren müssen (Ivanov & Pavlov et al. 2016, S. 7f; Schuh, Schenk & Servos 2015, S. 426; Queiroz & Ivanov et al. 2020, S. 2f). Die Störereignisse haben beispielsweise spontane Kapazitätsengpässe, Qualitätsprobleme oder Auftragsänderungen zur Folge, die wiederum einen erheblichen Einfluss auf den Zustand des einzelnen Produktionssystems haben (Lanza & Treber 2019, S. 439; Ouazene & Yalaoui et al. 2017, S. 1). Die Fähigkeit eines Produktionsstandorts, beauftragte Produkte zu fertigen, wird somit zunehmend volatil. Eine Erhöhung der Kapazitäten und Lagerbestände als Lösung zum Ausgleich der Störungen widerspricht abermals der Wirtschaftlichkeit, weil hierdurch die Kapitalkosten steigen. In der Folge sind Unternehmen gezwungen, Entscheidungen im Auftragsmanagement adaptiver und dynamischer zu treffen (Lanza & Nyhuis et al. 2018, S. 6f; Lanza & Ferdows et al. 2019, S. 837; Huber 2014, S. 127).

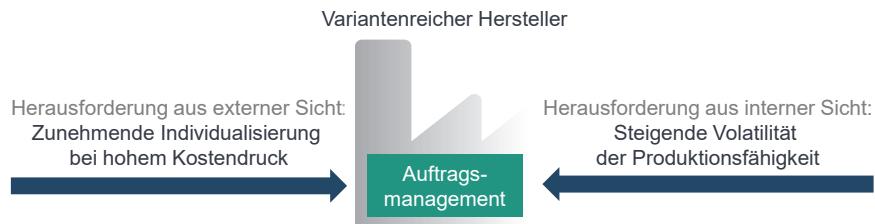


Abbildung 1-1: Überblick über die Herausforderungen im Auftragsmanagement

Im Kontext dieser Herausforderungen obliegt dem Auftragsmanagement insbesondere die Aufgabe, dynamisch zu entscheiden, zu welcher Lieferzeit und welchem Preis Aufträge angenommen werden (Schuh, Brandenburg & Cuber 2012, S. 61f; Kunath & Winkler 2019, S. 276). Unter dem Begriff „dynamisch“ wird hierbei die Berücksichtigung von Informationen über den Produktionsstandort und das Netzwerk im Zeitpunkt der Anfrage verstanden. Die vom Auftragsmanagement durchgeführte Lieferzeit-Preisgestaltung muss folglich schnell im Verhältnis zu einer angemessenen Angebotsbearbeitungszeit und adaptiv auf Basis der Informationsgrundlage sein. Um einem solchen Anspruch an die Dynamik gerecht zu werden, sollte die Lieferzeit-Preisgestaltung IT-basiert erfolgen (Ivanov, Tsipoulanidis & Schönberger 2019, S. 565; Kunath & Winkler 2019, S. 269f).

## 1.2 Forschungsfrage

Ausgehend von der Ausgangslage ergibt sich die leitende Forschungsfrage:

*Wie kann eine dynamische Lieferzeit-Preisgestaltung im Bereich der variantenreichen Produktion unter Berücksichtigung von Produktionsstandort und Netzwerk vorteilhaft gestaltet werden?*

## 1.3 Zielsetzung

Ausgehend von dieser Forschungsfrage ist das zentrale Ziel der Arbeit, ein ausführbares Lösungsmodell zur dynamischen Lieferzeit-Preisgestaltung im Bereich der variantenreichen Produktion unter Berücksichtigung von Produktionsstandort und Netzwerk zu entwickeln.

Der Zielsetzung der Arbeit liegt die Hypothese zugrunde, dass es ein Lösungsmodell zur dynamischen Lieferzeit-Preisgestaltung geben kann, welches sowohl der Herausforderung steigender Individualisierung als auch der Herausforderung steigender Volatilität im Produktionssystem begegnet. Hinter dieser Hypothese steht die Überlegung, dass eine dynamische Lieferzeit-Preisgestaltung auf zweierlei Art und Weise wirkt. Zum einen sollte es mit einer geeigneten Lieferzeit-Preisgestaltung möglich sein, kundenindividuelle Präferenzen zu antizipieren, einen geeigneten Kompromiss zwischen diesen Präferenzen und den eigenen Fähigkeiten zur Produktion zu finden und somit sowohl für den Hersteller als auch für den Kunden einen höheren Gewinn bzw. Nutzen zu erzielen (Ivanov, Tsipoulanidis & Schönberger 2019, S. 565). Zum anderen sollte es durch eine dynamische Lieferzeit-Preisgestaltung möglich sein, die Nachfrage zu beeinflussen bzw. in begrenztem Rahmen zeitlich zu verschieben, sodass es zu einer Kapazitätsnivellierung bei gleichzeitig optimalem Preiskompromiss kommt (Ouazene & Yalaoui et al. 2017, S. 1; Leimeister 2020, S. 319). Eine erfolgreiche Kapazitätsnivellierung erlaubt sodann die Reduktion teurer Flexibilität, zum Beispiel von Maschinen, oder ermöglicht es, eine größere Anzahl an Aufträgen zu erfüllen ohne in anderen, für das Auftragsmanagement relevanten Zielgrößen Einbußen hinnehmen zu müssen. In der Folge soll durch die dynamische Lieferzeit-Preisgestaltung die Wirtschaftlichkeit und Reaktionsgeschwindigkeit des Produktionsstandorts steigen (Ouazene & Yalaoui et al. 2017, S. 1).

Um die Korrektheit der Überlegung nachweisen zu können, ist das Teilziel 1 der Arbeit ein Zielsystem, welches geeignete Zielgrößen, wie zum Beispiel den Gewinn oder den

Servicegrad, umfasst. Anhand der Zielgrößen des Zielsystems wird die Leistung der dynamischen Lieferzeit-Preisgestaltung gemessen. Der dynamische Lieferzeit-Preisgestaltungsansatz selbst wird als ausführbares Lösungsmodell konzipiert, um dem Anspruch der Dynamik gerecht zu werden. Aufgrund der hohen Komplexität und engen Zeitrestriktion der umzusetzenden Lieferzeit-Preisgestaltung ist es notwendig, auf adaptive, heuristische Verfahren zurückzugreifen (Juan & Keenan et al. 2021, S. 2). Die Auswahl eines entsprechenden Verfahrens und die Umsetzung zu einem Lösungsmodell sind Teilziel 2 und 3 dieser Arbeit. Mithilfe des Lösungsmodells ist dann die Durchführung von Simulationsexperimenten möglich. Hier besteht Teilziel 4 darin, geeignete Testhypotesen, ein Vergleichssystem und Testszenarien aufzustellen, um mithilfe des Zielsystems die Ausgangshypothese dieser Arbeit zu validieren.

Abbildung 1-2 fasst die Zielsetzung nochmals zusammen, indem sie zuerst linksseitig die Ausgangslage aufgreift, in der Mitte die Hypothese der Arbeit nochmals herausstellt und auf der rechten Seite die zentralen (Teil-)Ziele aufzeigt.

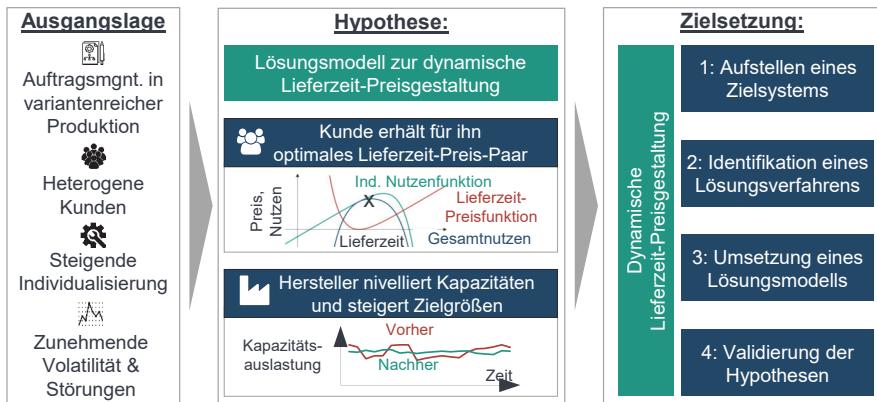


Abbildung 1-2: Überblick über Ausgangslage, Hypothesen und Zielsetzung der Arbeit

## 1.4 Aufbau der Arbeit

Ausgehend von der Zielsetzung werden in Kapitel 2 zunächst die relevanten Grundlagen dargelegt. Hierzu gehören die Themenfelder variantenreiche Produktion, Ablaufplanung, Auftragsmanagement und Preisgestaltung. Methodisch wird in die Bereiche der Simulation und des Reinforcement Learning eingeführt. Im Zuge der Darlegung der Grundlagen werden kontextspezifisch Anforderungen an den Lösungsansatz aufgestellt.

Die abgeleiteten Anforderungen werden in Kapitel 3 aufgegriffen, in welchem der Stand der Forschung dargelegt und sodann das Forschungsdefizit abgeleitet wird. Hierfür werden bestehende Lösungsansätze aus dem Stand der Forschung nach ihrer zugrundeliegenden Methodik strukturiert vorgestellt und hinsichtlich der Erfüllung der Anforderungen untersucht.

Das identifizierte Forschungsdefizit wird in den nachfolgenden Kapiteln mit dem eigenen Lösungsansatz adressiert. Kapitel 4 stellt den Lösungsansatz allgemein vor. Der Lösungsansatz wird dann im Laufe des Kapitels 5 zu einem Lösungsmodell ausgearbeitet und erläutert. Die Erprobung des Lösungsmodells erfolgt in Kapitel 6. Hier wird die Tauglichkeit in zwei Anwendungsfällen aus der industriellen Praxis anhand von Hypothesen geprüft.

Kapitel 7 bietet eine Diskussion der Ergebnisse, die kritische Würdigung der Arbeit sowie einen Ausblick auf mögliche zukünftige Schritte basierend auf den erlangten Erkenntnissen. Im abschließenden Kapitel 8 wird die gesamte Arbeit zusammengefasst.

## 2 Grundlagen

In diesem Kapitel werden die Grundlagen aufgeführt, auf denen der Lösungsansatz aufbaut. Zuerst werden die Grundlagen der variantenreichen Produktion vorgestellt (Kapitel 2.1). Als Zweites wird die Produktionsplanung und -steuerung mit Fokus auf dem Auftragsmanagement und der Ablaufplanung behandelt (Kapitel 2.2). Eine Aufgabe des Auftragsmanagements ist die Preisgestaltung, welche zusammen mit den Grundlagen zur Entscheidungstheorie in Kapitel 2.3 erläutert wird. Wichtige methodische Grundlagen zur dynamischen Lieferzeit-Preisgestaltung werden in den Kapiteln 2.4 – 2.6 vermittelt. Sofern sich anhand der Ausführungen in einem Kapitel Anforderungen an den Lösungsansatz ableiten lassen, werden diese am Ende des jeweiligen Kapitels erläutert und hervorgehoben.

### 2.1 Variantenreiche Produktion

Dieser Abschnitt schafft ein Verständnis zur Produktion allgemein (Kapitel 2.1.1) und widmet sich dann der variantenreichen Produktion im Speziellen (Kapitel 2.1.2).

#### 2.1.1 Allgemeine Grundlagen zur Produktion

Die Produktion steht im Mittelpunkt eines Industrieunternehmens (Bloech & Bogaschewsky et al. 2014, S. 3) und ist damit ein strategischer Erfolgsfaktor für ein solches Unternehmen (Gottmann 2016, S. 18). Die verfolgten Ziele im Rahmen der Produktion sind dabei vielzählig: Ein Denken in den drei Dimensionen Zeit, Kosten und Qualität ist verbreitet (Bloech & Bogaschewsky et al. 2014, S. 7f). Es kann auch noch die vierte Dimension Variabilität hinzugefügt werden, die sowohl Flexibilität als auch Wandlungsfähigkeit eines Unternehmens umfasst (Erlach 2010, S. 13f, 23f).

Aus Sicht eines betriebswirtschaftlichen Wertschöpfungsprozesses ist Produktion der Transformationsprozess von Materialien, Diensten, Rechten und Informationen (Dyckhoff 2006, S. 1–9). Bei diesem Transformationsprozess werden Produktionsfaktoren als Input genutzt, um ein Produkt als Output zu erzeugen (Bloech & Bogaschewsky et al. 2014, S. 3f). Zu den Produktionsfaktoren zählen sowohl direkte Faktoren wie Produktionsanlagen, Material, Mitarbeiter, als auch indirekte Faktoren wie Organisation, Logistik, Materialwirtschaft, Planung und Steuerung (Gottmann 2016, S. 18; Eversheim 1996, S. 3f). Die zur Produktion gezählten Vorgehensweisen haben darüber hinaus Einfluss auf administrative Prozesse, Kosten und Flexibilität eines Unternehmens (Gottmann 2016, S. 1).

Aus Ressourcensicht lassen sich in Ebenen gegliederte Struktureinheiten der Produktion aufstellen (Wiendahl & ElMaraghy et al. 2007, S. 785; Schenk, Wirth & Müller 2014, S. 165). Beginnend mit der Station auf der untersten Ebene, bauen die Ebenen über die Zelle, Linie, Segment, Standort bis zum Netzwerk aufeinander auf (Wiendahl & ElMaraghy et al. 2007, S. 785). Abbildung 2-1 stellt die Ressourcensicht grafisch dar.

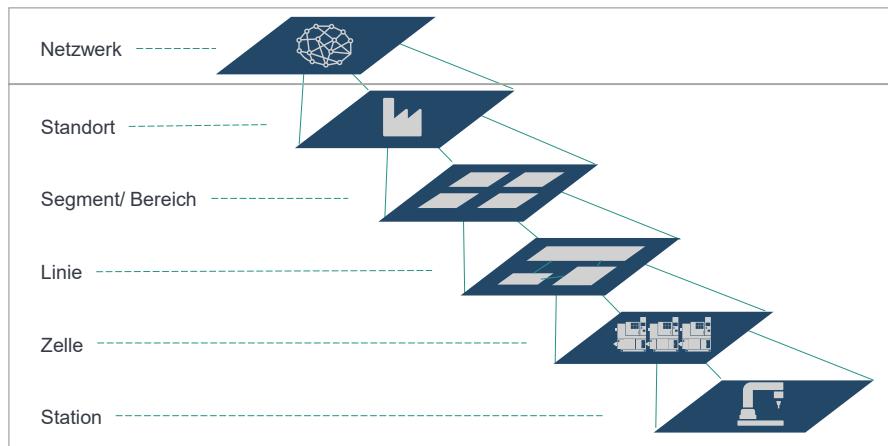


Abbildung 2-1: Ebenen der Produktion in Anlehnung an Wiendahl & ElMaraghy et al. (2007, S. 785f)

Auf der Netzwerkebene werden anders als auf den anderen Ebenen auch firmenübergreifende Standorte betrachtet, welche über Material-, Finanz- und Informationsströme miteinander verbunden sind (Lanza & Ferdows et al. 2019, S. 823; Váncza & Monostori et al. 2011, S. 799). Aus Sicht eines Standorts können alle anderen Teilnehmer des Netzwerks in Konkurrenten, Zulieferer und Kunden unterschieden werden. Die Aufgaben aus der Sicht eines einzelnen Standorts können gemäß dem SCOR-Modell in fünf Aufgaben unterteilt werden: Planen, Beschaffen, Herstellen, Liefern und Zurücknehmen (Werner 2020, S. 71f).

Als Fazit lässt sich zusammenfassen, dass an einem Produktionsstandort Stationen Transformationsprozesse durchführen, um Produkte herzustellen. Die Produktionsstandorte sind hierbei mit anderen Produktionsstandorten in einem Netzwerk verbunden. Diese anderen Standorte können als Konkurrenten, Zulieferer oder Kunden auftreten. Dieser Strukturlogik kann auch bei der Modellierung im Rahmen einer dynamischen Lieferzeit-Preisgestaltung gefolgt werden.

### **2.1.2 Grundlagen der variantenreichen Produktion**

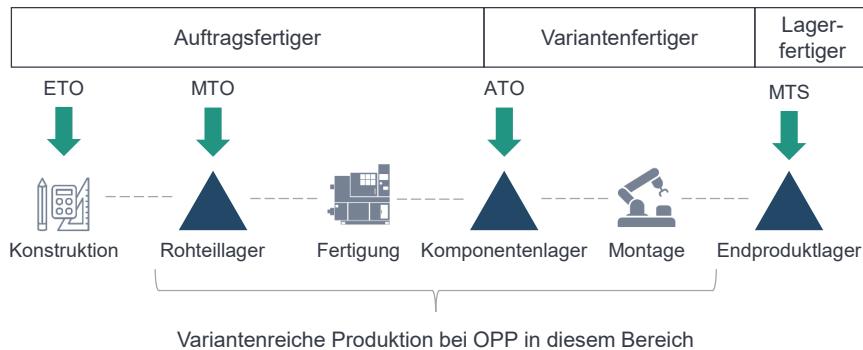
In den letzten Jahren gewann die variantenreiche Produktion zunehmend an Relevanz (Piller 2008, S. 127f; Bertagnolli 2018, S. 11). Das ist eine Folge des zunehmenden Marktdrucks und des Kundenwunsches nach mehr Individualisierung (Piller 2008, S. 127f; Bertagnolli 2018, S. 11). Als Folge der Individualisierung wird in der Literatur der Trend weg von der freien Auftrags- und starren Lagerfertigung hin zur Variantenfertigung beobachtet (Piller 2008, S. 127; Helmig, Schmidt & Kompa 2012, S. 205f). Sie geht einher mit einer Vielzahl an Produkten und einer entsprechenden Anzahl an Produktionsressourcen, um diese zu fertigen (Piller 2008, S. 130f; Volling 2009, S. 11; Helmig, Schmidt & Kompa 2012, S. 169; Schuh, Schmidt & Helmig 2012, S. 169).

Zur besseren Abgrenzung der Variantenfertigung von der Lager- und Auftragsfertigung, kann sie als ein Fertigungstyp betrachtet werden. Unter einem Fertigungstyp werden Unternehmen, deren Produktionssysteme markante Ähnlichkeiten aufweisen, zusammengefasst (Schuh & Schmidt 2014, S. 33). Hierbei konnte beobachtet werden, dass ein wesentliches Unterscheidungsmerkmal eines Fertigungstyps der Kundenentkoppelungspunkt (engl. Order Penetration Point, kurz OPP) ist (Schuh & Schmidt 2014, S. 33f; Kellner, Lienland & Lukesch 2020, S. 9f). Folglich hängt der OPP mit dem Fertigungstyp und dem Produktionssystem zusammen. Ein Auftragsfertiger ist charakterisiert durch einen OPP im Rohteillager (Schuh & Schmidt 2014, S. 35–37). Er produziert keine Produkte vor, sondern produziert nur basierend auf Kundenaufträgen. Der Variantenfertiger hat eine kundenanonyme Vorproduktion und montiert dann vorbestimmte Varianten auftragsbezogen (Schuh & Schmidt 2014, S. 40–42). Er hat seinen OPP typischerweise unmittelbar vor der Endmontage. Damit kann der Variantenfertiger als Hybrid zwischen Auftragsfertiger und Lagerfertiger bezeichnet werden. Der Lagerfertiger hat entsprechend seinen OPP im Endproduktlager (Schuh & Schmidt 2014, S. 42). Er produziert folglich komplett kundenanonym.

Im englischsprachigen Raum wird eine abweichende Nomenklatur verwendet, deren Verständnis für die korrekte Einordnung der Literatur relevant ist. Hier werden Engineer-To-Order (ETO), Make-to-Order (MTO), Assemble-To-Order (ATO) bzw. Build-To-Order (BTO) oder auch Configure-To-Order (CTO) und Make-To-Stock (MTS) unterschieden, um eine Produktion zu charakterisieren (Kellner, Lienland & Lukesch 2020, S. 9f; Geier 2014, S. 34; Jammerlegg, Reiner & Wakolbinger 2018, S. 71). Sowohl

ETO als auch MTO kann dem Auftragsfertiger zugeordnet werden, da der Auftragsfertiger sowohl ein festes Erzeugnisspektrum vorsehen kann, als auch komplett neue Produkte fertigt (Schuh, Schmidt & Helming 2012, S. 137). ATO, BTO und CTO können dem Variantenfertiger zugeordnet werden, wohingegen die MTS-Fertigung dem Lagerfertiger entspricht.

Es lässt sich aus dieser Darstellung ableiten, dass die eingangs beschriebenen Charakteristika einer variantenreichen Produktion auf Unternehmen vom Auftragsfertiger mit festem Produktspektrum bis zum typischen Variantenfertiger mit OPP vor der Endmontage zutreffen können. Dies sei definiert als der Bereich variantenreicher Produktion (vgl. Abbildung 2-2).



*Abbildung 2-2: Einordnung der variantenreichen Produktion über den Kundenentkopplungspunkt und die zugehörigen Fertigungstypen*

Die auf diese Weise definierte variantenreiche Produktion steckt das in dieser Arbeit fokussierte Anwendungsgebiet der dynamischen Lieferzeit-Preisgestaltung ab. Ein Lösungsansatz muss entsprechend eine **dynamische Lieferzeit-Preisgestaltung in der variantenreichen Produktion** behandeln. Zudem muss ein solcher Lösungsansatz mit einer **Vielzahl an Produkten und Produktionsressourcen** umgehen können.

## 2.2 Ablaufplanung und Auftragsmanagement als Teil der Produktionsplanung und -steuerung

Die Produktionsplanung und -steuerung (PPS) hat allgemein das Ziel, das Produktionsystem optimal zu betreiben (Schuh, Brosze & Brandenburg 2012, S. 11). Sie soll hierzu die Vorgaben aus dem strategischen Produktionsmanagement umsetzen und

sich stets am Kundennutzen ausrichten (Schuh, Brosze & Brandenburg 2012, S. 11; Kellner, Lienland & Lukesch 2020, S. 166). Die spezifischen Ziele der PPS lassen sich als eine hohe und gleichmäßige Ressourcenauslastung, geringe Kapitalbindung, niedrige Beschaffungskosten, hohe Flexibilität, sowie eine hohe Termintreue und Lieferbereitschaft festhalten (Kellner, Lienland & Lukesch 2020, S. 162; Schuh, Brandenburg & Cuber 2012, S. 29). Die PPS behandelt hierzu alle Aufgaben im Kontext des Kapazitäts-, mengen- und terminbezogenen Produktionsablaufs (Schuh, Brandenburg & Cuber 2012, S. 29). Diese Aufgaben unterteilen Schuh, Brosze & Brandenburg (2012, S. 21) in die drei Bereiche Netzwerkaufgaben, Kernaufgaben und Querschnittsaufgaben (vgl. Abbildung 2-3).

Im Folgenden werden die Ablaufplanung und das Auftragsmanagement der Produktionsplanung und -steuerung vertieft, die für die Arbeit von hoher Relevanz sind. Die Ablaufplanung (Kapitel 2.2.1) kann der Eigenfertigungsplanung und -steuerung zugeordnet werden und hat einen wesentlichen Einfluss auf die Einhaltung von Lieferzeiten. Das Auftragsmanagement (Kapitel 2.2.2), dessen Bedeutung bereits in der Einleitung motiviert wurde, muss mit der Ablaufplanung harmonieren, damit nur Aufträge angenommen werden, die noch rechtzeitig bearbeitet werden können.



Abbildung 2-3: Übersicht über die Aufgaben der PPS in Anlehnung an Schuh, Brosze & Brandenburg (2012, S. 21) und markierte Schwerpunkte von Kapitel 2.2

## 2.2.1 Ablaufplanung als Teil der Produktionsplanung und -steuerung

Die Ablaufplanung in der Produktionsdomäne kann als Teil der PPS angesehen werden (Bloech & Bogaschewsky et al. 2014, S. 2). Hinter der Ablaufplanung verbirgt sich ein Optimierungsproblem, welches die zeitliche, operative Zuordnung von Aufträgen und Ressourcen behandelt (Jaehn & Pesch 2019, S. 17; Pinedo 2016, S. 1). Das Ziel der Ablaufplanung ist es, die anfallenden Aufträge im Produktionssystem derart zu allokalieren, dass die Abarbeitung effizient durchgeführt und gleichzeitig Endtermine eingehalten werden können (Schuh, Brandenburg & Cuber 2012, S. 56; Kummer 2013, S. 286;

Bloech & Bogaschewsky et al. 2014, S. 236f). Die Effizienz misst sich hierbei an verschiedenen Zielgrößen, wie zum Beispiel Rüstkosten oder Liefertermintreue (Bloech & Bogaschewsky et al. 2014, S. 236f). Aufgrund von verschiedenen, zufallsbasierten Einflüssen, wie zum Beispiel Eilaufträgen oder Maschinenausfällen, muss die Ablaufplanung laufend wiederholt werden, um sich den Gegebenheiten anzupassen (Pinedo 2016, S. 4f).

Die Optimierung der Ablaufplanung ist in vielen Fällen NP-schwer und kann nicht mehr in polynomieller Zeit exakt durchgeführt werden, wodurch exakte Verfahren schnell an ihre praktischen Grenzen stoßen (Blazewicz & Ecker et al. 2019, S. 2, 37). Eine Alternative ist die Anwendung von Prioritätsregeln. Sie weisen Aufträgen auf Basis ihrer Eigenschaften eine Priorität zu, welche die Reihenfolge der Bearbeitung bestimmt (Bloech & Bogaschewsky et al. 2014, S. 257). Nachfolgend wird eine Auswahl gängiger Prioritätsregeln vorgestellt. Diese und weitere Prioritätsregeln finden sich in (Blazewicz & Ecker et al. 2019, S. 367; Kellner, Lienland & Lukesch 2020, S. 305; Kummer 2013, S. 285).

- Die „First Come, First Serve“-Regel (FCFS) priorisiert die Aufträge auf Basis ihres Eintreffens.
- Die „Earliest Due Date“- Regel (EDD) priorisiert Aufträge nach ihrem Fälligkeitsdatum.
- Die „Shortest Processing Time“ - Regel (SPT) priorisiert Aufträge basierend auf ihrer Bearbeitungszeit.
- Die „Minimum Slack“-Regel priorisiert Aufträge basierend auf ihrem Schlupf. Der Schlupf bezeichnet hier die Differenz zwischen Fälligkeitsdatum und Durchlaufzeit.

Im Folgenden wird zwischen fristbasierten und nicht-fristbasierten Methoden der Ablaufplanung unterschieden. Fristbasierte Methoden, wie zum Beispiel die EDD-Regel und die „Minimum Slack“-Regel, zeichnen sich dadurch aus, dass sie das Fälligkeitsdatum bei der Einplanung berücksichtigen und dadurch Verspätungen minimieren können. Sie sind daher besonders bei Aufträgen mit verschiedenen Lieferzeiten sinnvoll, wie sie bei kundenindividuellen Angeboten auftreten.

Um kurzfristige Aufträge mit individuellen Lieferzeiten entsprechend einplanen zu können, ergibt sich die Anforderung nach einer **fristbasierten Ablaufsteuerung**, mit der eine dynamischen Lieferzeit-Preisgestaltung harmonieren muss.

## 2.2.2 Auftragsmanagement als Teil der Produktionsplanung und -steuerung

Unter dem Auftragsmanagement werden die Aufgaben zur Auftragsplanung, -steuerung und -überwachung zusammengefasst (Schuh, Brandenburg & Cuber 2012, S. 60). Für den Varianten- und Auftragsfertiger mit festem Variantenspektrum, welcher in Kapitel 2.1.2 der variantenreichen Produktion zugeordnet wurde, findet sich in der Literatur ein Sollprozess für das Auftragsmanagement (Schuh, Schmidt & Helwig 2012, S. 143, 172). Dieser ist in Abbildung 2-4 dargestellt.

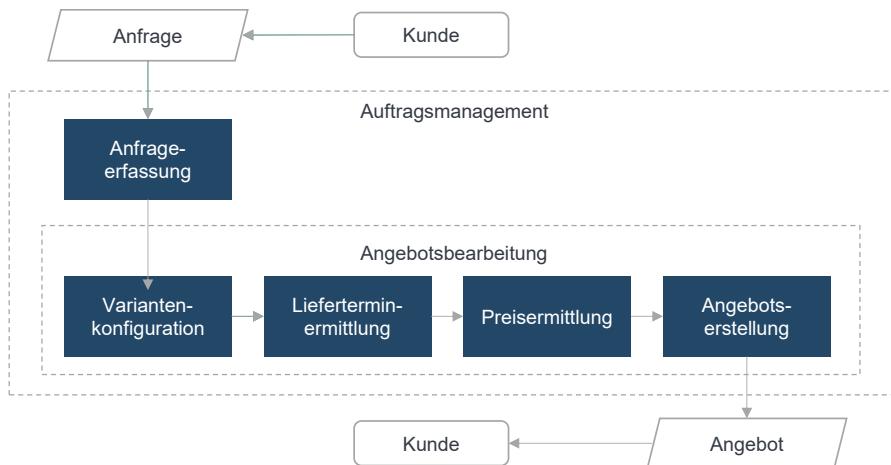


Abbildung 2-4: Soll-Prozess des Auftragsmanagements in der variantenreichen Produktion in Anlehnung an Schuh, Schmidt & Helwig (2012, S. 143, 172)

Gemäß dem Soll-Prozess stellt der Kunde eine Anfrage für ein bestimmtes Produkt in bestimmter Menge. Im Rahmen des Auftragsmanagements wird dann ein möglicher Liefertermin und Verkaufspreis ermittelt (Schuh, Schmidt & Helwig 2012, S. 143, 172). Das Ergebnis der Ermittlung wird dem Kunden in Form eines Angebots übergeben. In der variantenreichen Produktion läuft das Auftragsmanagement häufig (teil-)automatisiert ab (Schuh, Schmidt & Helwig 2012, S. 171; Salewski 2018, S. 235). Hierbei ist die Preiskalkulation basierend auf Standard- oder Listenpreisen üblich (Schuh, Schmidt &

Adema 2014, S. 122). Für kurze Lieferzeiten bieten einige Unternehmen spezielle Optionen mit Preisaufschlägen an (Wiendahl 2011, S. 123). Auch kann ein Unternehmen ein permanentes Lieferversprechen verfolgen, über welches eine Lieferdauer für die Standardbestellung fest garantiert wird. Hierdurch wird versucht, den Kunden eine Planungssicherheit zu geben und sich damit gegen die unstete Konkurrenz durchzusetzen (Werner 2015, S. 22; Weichselbaum & Rothböck 2007, S. 167, 198). Neben der Preis- und Lieferterminermittlung kann auch die Selektion von Aufträgen eine große Rolle spielen. Je nach Markt- und Auftragslage kann es Sinn ergeben, die Aufträge nach Profitabilität zu filtern, da die Kapazitäten begrenzt sind (Hintsches 2012, S. 15, 159; Volling 2009, S. 65).

In der Selektion der Aufträge und im gesamten Zusammenspiel zwischen Vertrieb und Produktion gibt es noch offenes Potenzial, um die Profitabilität der variantenreichen Produktion zu steigern (Kunath & Winkler 2019, S. 287f). Das offene Potenzial besteht in der Praxis häufig bei der Integration der verschiedenen Funktionen (Schuh, Brandenburg & Cuber 2012, S. 60). Durch eine dynamische Lieferzeit-Preisgestaltung soll diese Integration gelingen. Zentral ist hierbei die zunehmende Individualisierung, die von den heterogenen Kunden gefordert wird (Kunath & Winkler 2019, S. 269). Gleichzeitig besteht der Wunsch nach Just-in-Time Lieferungen, sodass nicht allein Verspätungen, sondern auch verfrühte Lieferungen schadhaft sind (Piller 2008, S. 7, 43).

Als Fazit lässt sich aus diesem Kapitel festhalten, dass eine dynamische Lieferzeit-Preisgestaltung, welche eine Verbesserung im hier erläuterten Soll-Prozesses erzielen soll, **individuelle und heterogene Kunden** berücksichtigen muss. Diese haben Präferenzen hinsichtlich des Preises und der Lieferzeit. Gleichzeitig muss dem Wunsch nach einer Just-In-Time-Produktion Rechnung getragen werden, welche eine **beidseitige Liefertermintreue** bedingt.

## 2.3 Preisgestaltung und Kundenverhalten

Die Preisgestaltung als wichtiger Teil der Angebotsbearbeitung und Schwerpunktthema der Arbeit wird in diesem Abschnitt behandelt. Als erstes werden Grundlagen der dynamischen Preisgestaltung, welche dem Revenue Management entstammt, dargelegt (Kapitel 2.3.1). Zum Verständnis, wann ein Kunde einen dynamischen Preis akzeptiert oder ablehnt, wird anschließend noch das Thema der Entscheidungstheorie zur Modellierung von Kundenverhalten behandelt (Kapitel 2.3.2).

### 2.3.1 Grundlagen der dynamischen Lieferzeit-Preisgestaltung

Unternehmen mit eingeschränkten Kapazitäten stehen vor der Herausforderung, diese optimal zu nutzen (Lohnert & Fischer 2019, S. 235). Diese Herausforderung ist besonders dann von hoher Bedeutung, wenn die Kapazitäten nicht kurzfristig angepasst werden können, zum Beispiel aufgrund von langen Beschaffungszeiten für Maschinen und Anlagen. Eine solche Situation herrscht in der variantenreichen Produktion vor (Barut & Sridharan 2005, S. 288). In diesen Situationen bietet sich der Einsatz von Methoden des Revenue Management, speziell der dynamischen Lieferzeit-Preisgestaltung an, um die vorhandenen, fixen Kapazitäten optimal auszunutzen (Lohnert & Fischer 2019, S. 238; Guadix & Cortés et al. 2010, S. 519). Zu beachten ist, dass sich in der Literatur häufig nur der Begriff der dynamischen Preisgestaltung findet. In dieser Arbeit liegt der Fokus auf der Gestaltung des Preises in Abhängigkeit von der Lieferzeit. Daher wird stets von dynamischer Lieferzeit-Preisgestaltung gesprochen. In jedem Fall handelt es sich um eine Entscheidungsunterstützung mit dem Ziel der Umsatzmaximierung (Ivanov, Tsipoulanidis & Schönberger 2019, S. 528). Häufig kann durch eine gute dynamische Lieferzeit-Preisgestaltung sowohl der Kundennutzen als auch der Unternehmensgewinn gesteigert werden (Helmold 2020, S. 178-180).

Für die Anwendung einer dynamischen Lieferzeit-Preisgestaltung gibt es neben quasi-fixen Kapazitäten weitere begünstigende Merkmale. So sollten Kapazitäten vorbuchbar, Kunden oder Märkte segmentierbar und die Nachfrage großen Schwankungen unterlegen sein (Klein & Steinhardt 2008, S. 180; Tomczak & Heidig 2014, S. 4f). Außerdem sind ein Verlust ungenutzter Kapazitäten und ein hoher Quotient aus fixen zu variablen Kosten zu nennen (Tomczak & Heidig 2014, S. 4f). Da diese Merkmale typischerweise auf die variantenreiche Produktion zutreffen, kann sie als sinnvolles Anwendungsgebiet einer dynamischen Lieferzeit-Preisgestaltung angesehen werden (Zatta 2016, S. 19f; Salewski 2018, S. 243).

Im Rahmen der dynamischen Lieferzeit-Preisgestaltung soll die Frage beantwortet werden, welche Aufträge zu welchem Preis angenommen werden sollen (Capocchi 2019, S. 104). Eine gute dynamische Lieferzeit-Preisgestaltung schafft es, die allgemeine Marktlage, also Nachfrage und Angebot, und die aktuelle Kapazitätsauslastung in die Lieferzeit-Preisbildung zu integrieren (Tomczak & Heidig 2014, S. 85; Klein & Steinhardt 2008, S. 174f). Durch dieses Vorgehen soll es dem Anwender der dynamischen Liefer-

zeit-Preisgestaltung möglich werden, die sich verändernden Marktbedingungen effizient zu adaptieren (Helmold 2020, S. 46). Es kommt aus makroökonomischer Sicht zu einem besseren Ausgleich von Angebot und Nachfrage (Klein & Steinhardt 2008, S. 180).

So viel Potenzial zur Steigerung des Gewinns eine dynamische Lieferzeit-Preisgestaltung auch bietet, ist es wichtig, die Risiken zu berücksichtigen: Zentral ist hierbei der Kunde. Zum einen kann ein Gefühl der Ungleichbehandlung durch eine dynamische Lieferzeit-Preisgestaltung beim Kunden auftreten, zum anderen können die resultierenden Preisschwankungen auch rational als Planungsrisiko wahrgenommen werden (Tomczak & Heidig 2014, S. 93; Barut & Sridharan 2005, S. 288). Dies gilt es bei der Auslegung der dynamischen Lieferzeit-Preisgestaltung zu berücksichtigen.

Neben dem Preis für ein Produkt spielt auch die Zeit eine wesentliche Rolle: Zu welchem Zeitpunkt eine Leistung bezogen wird, kann somit ebenfalls in die Preisbildung integriert werden (Kimes & Wirtz 2015, S. 7f). Dies liegt an der individuellen Preisbereitschaft und Lieferzeitpräferenz der Kunden (Kimes & Wirtz 2015, S. 7f). Aufgrund der Informationsasymmetrie sind die Präferenzen der Kunden jedoch a priori unbekannt (Váncza & Monostori et al. 2011, S. 801; Vosooghidizaji, Taghipour & Canel-Depitre 2020, S. 1805). Die Präferenzen können jedoch in Analogie eines Signalspiels über mehrere Interaktionen hinweg offenbart werden (Bartholomae & Wiens 2016, S. 172). Hierbei sollte die Anzahl der angebotenen Alternativen mindestens der Anzahl der Präferenzgruppen entsprechen, um eine Unterscheidung aller Präferenzgruppen anhand der Wahl der Alternative zu erreichen. Hieraus lässt sich ableiten, dass der theoretische Idealfall eine kontinuierliche (Lieferzeit-Preis-)Funktion ist, welche durch ihre Kontinuität prinzipiell unendlich viele, beliebige Präferenzgruppen segregieren kann (Winter 2019, S. 181).

Als Fazit dieses Kapitels ergibt sich die Anforderung an den Lösungsansatz, dass dieser eine **Informationsasymmetrie** berücksichtigen muss. Dies umfasst nicht nur die reine Modellierung, sondern auch entsprechende Maßnahmen zum Umgang mit Informationsasymmetrie, wie zum Beispiel die Nutzung einer kontinuierlichen Lieferzeit-Preisfunktion im Sinne eines Signalspiels. Weiterhin wäre es von Vorteil, wenn der Lösungsansatz eine Möglichkeit zur manuellen Begrenzung von Lieferzeiten und Preisen bietet, um dem Anwender eine Möglichkeit zu geben, gezielt die Unsicherheit für den Kunden einzustellen.

### 2.3.2 Modellierung von Kundenverhalten mit Entscheidungstheorie

Ob eine dynamische Lieferzeit-Preisgestaltung den gewünschten Effekt erzielen kann, hängt wesentlich mit den Kundenpräferenzen zusammen. Nur wenn unter den verschiedenen Preisoptionen diejenige ist, welche die Kundenpräferenz erfüllt, wird sich ein Kunde für diese Option entscheiden. Mit dieser Thematik beschäftigt sich die Entscheidungstheorie. (Kolmar 2017, S. 165f)

Ein zentrales Werkzeug der Entscheidungstheorie ist die Nutzenfunktion. Eine Nutzenfunktion dient der Abbildung von Präferenzordnungen (Kolmar 2017, S. 174). Eine Präferenz ist die Bevorzugung einer Alternative, beispielsweise einer Merkmalsausprägung, gegenüber einer anderen durch einen individuellen Entscheider (Kolmar 2017, S. 166). Präferenzen können zur Erklärung der Kaufbereitschaft bzw. -entscheidung eines individuellen Entscheiders herangezogen werden (Giesecking 2009, S. 60; Rao 2014, S. 1). Produkte können in diesem Kontext als Bündel von Merkmalsausprägungen verstanden werden (Backhaus & Erichson et al. 2018, S. 500). Präferenzen für Merkmalsausprägungen sind hierbei nicht statisch, weisen jedoch üblicherweise eine gewisse zeitliche Stabilität auf (Hillig 2006, S. 14).

Eine Präferenzordnung gibt eine ordinale Sortierung hinsichtlich der Bevorzugung verschiedener Alternativen an (Kolmar 2017, S. 174). Eine zugehörige Nutzenfunktion muss nun in der Lage sein, diese ordinale Sortierung wiederzugeben. Es kann daher eine Vielzahl an Nutzenfunktionen zur Wiedergabe einer Präferenzordnung geben, so lange sie folgende zwei Bedingungen erfüllt (Kolmar 2017, S. 174):

$$x^i > x^j \Leftrightarrow u(x^i) > u(x^j) \quad \text{Formel 2-1}$$

$$x^i \sim x^j \Leftrightarrow u(x^i) = u(x^j) \quad \text{Formel 2-2}$$

Hierbei gibt  $x$  die Alternative an und  $u(x)$  repräsentiert die entsprechende Nutzenfunktion in Abhängigkeit der Alternative. Eine Alternative kann wiederum auch ein Tupel aus Merkmalen sein.

Die absoluten Werte haben folglich keine eigene Bedeutung über die Reihenfolge hinaus. Insbesondere führt die direkte Verrechnung von Nutzenwerten, zum Beispiel mittels Addition oder Subtraktion, im Allgemeinen zu keiner ökonomischen Aussage (Kolmar 2017, S.174f). Nutzenfunktionen lassen sich daher auch als Präferenzmodelle bezeichnen, da sie die Präferenzordnung wiedergeben (Bichler & Trommsdorff 2021,

S.81f). Um einzelne Nutzenfunktionen sinnvoll zu einer Gesamtfunktion aggregieren zu können, werden Verknüpfungsfunktionen verwendet (Bichler & Trommsdorff 2021, S.81f). Dies kann beispielsweise hilfreich sein, wenn die Kaufbereitschaft für ein Produkt bewertet werden soll, jedoch nur die Nutzenfunktionen zu einzelnen Produktmerkmalen bekannt sind. Verknüpfungsfunktionen werden auch als Präferenzintegrationsmodelle bezeichnet (Bichler & Trommsdorff 2021, S. 81).

Präferenzintegrationsmodelle lassen sich in kompensatorisch und nicht-kompensatorisch einteilen (Bichler & Trommsdorff 2021, S. 88). Diese Unterscheidung bezieht sich auf die Substituierbarkeit der Nutzenbeiträge von Merkmalen (Hillig 2006, S. 30). Die kompensatorischen Modelle können wiederum in additive, nicht-additive und multiplikative Modelle unterschieden werden (Hillig 2006, S. 30). Die additiven Modelle sind die am weitest verbreitete Art (Bichler & Trommsdorff 2021, S. 89). Sie nehmen eine Austauschbarkeit bzw. Kompensation der verschiedenen Nutzenbeiträge an (Hillig 2006, S. 31). Für die Bestimmung der finalen Kaufentscheidung auf Basis des aggregierten Wertes gibt es verschiedene Ansätze (Hillig 2006, S. 31). Ein gängiger Ansatz ist die First-Choice-Regel (Hillig 2006, S. 31; Woratschek & Roth 2017, S. 338). Diese Regel besagt, dass ein Entscheider bzw. Kunde sich immer für die Alternative mit dem höchsten, aggregierten Gesamtwert entscheidet (Woratschek & Roth 2017, S. 338).

Im Rahmen einer dynamischen Lieferzeit-Preisgestaltung kann eine Verknüpfungsfunktion additiv aus drei einzelnen Nutzenfunktionen gebildet werden. Diese drei Nutzenfunktionen sind abhängig vom Preis, dem Nutzenbeitrag des Produkts und der Lieferzeit (Afèche & Pavlin 2016, S. 2416; Akan, Ata & Olsen 2012, S. 1507f; Yu, Zhao & Sun 2013, S. 125f). Die Klassifikation der Präferenzintegrationsmodelle ist in Abbildung 2-5 dargestellt, wobei die für eine dynamische Lieferzeit-Preisgestaltung relevanten Elemente hervorgehoben sind.

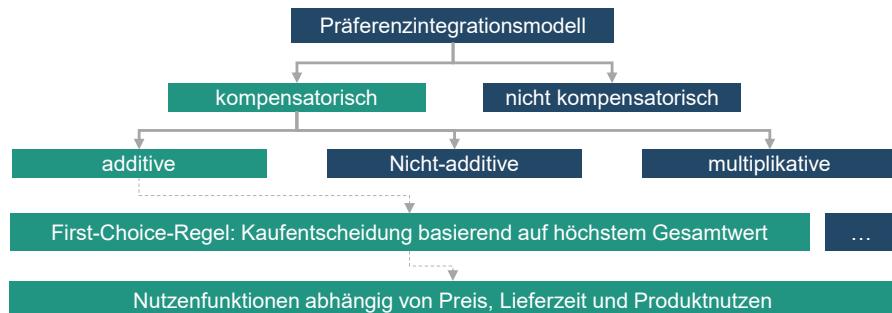


Abbildung 2-5: Überblick über die Einteilung von Präferenzintegrationsmodellen und Hervorhebung der im Kontext einer Lieferzeit-Preisgestaltung relevanten Elemente

Für eine Nutzenfunktion, welche abhängig von der Lieferzeit ist, kann im Kontext industrieller Produktion ein fallender Verlauf unterstellt werden, was mathematisch mit einer strikt negativen, zweiten Ableitung der Nutzenfunktion gleichzusetzen ist (Afèche & Pavlin 2016, S. 2416; Akan, Ata & Olsen 2012, S. 1507f; Yu, Zhao & Sun 2013, S. 125f). In diesem Fall wird auch von Verspätungskosten gesprochen (Yu, Zhao & Sun 2013, S. 126). Es wird zudem angenommen, dass ein Kunde nur dann ein Angebot realisiert, wenn der Gesamtnutzen positiv ist (vgl. Afèche & Pavlin 2016, S. 2416). Dies ist äquivalent mit einer Nicht-Kauf-Alternative mit einem Gesamtnutzen von null. Weitere gleichartige Arbeiten im Bereich der dynamischen Lieferzeit-Preisgestaltung, die diese Darstellung untermauern, finden sich im Stand der Forschung (vgl. Kapitel 3) und werden im Forschungsdefizit diskutiert.

## 2.4 Simulation im Produktionsumfeld

Ein wichtiges Werkzeug im Umfeld der Produktion ist die Simulation. Auch für diese Arbeit ist die Simulation von großer Bedeutung. Daher werden in Kapitel 2.4.1 zunächst die Grundlagen der Simulation behandelt. Sodann werden die statistische Versuchsplanung und statistische Tests fokussiert, die einen wichtigen Teil von Simulationsstudien darstellen (Kapitel 2.4.2).

### 2.4.1 Grundlagen der Simulation im Produktionsumfeld

Die Simulation ist eine Problemlösungsmethode (Gutenschwager & Rabe et al. 2017, S. 22). Mit ihr können Erkenntnisse über das Verhalten eines abgebildeten Systems

gewonnen werden, wobei kein Eingriff in das Realsystem notwendig ist (Gutenschwager & Rabe et al. 2017, S. 22). Aufgrund dieser vorteilhaften Eigenschaft können Hypothesen geprüft und Szenarien durchgespielt werden, ohne das betrachtete System zu beeinflussen (Banks 2010, S. 24). Die Simulation wird daher bei der Analyse, Kontrolle, Entwicklung und Planung von Systemen sowie zur Entscheidungsunterstützung eingesetzt (Rabe, Spiekermann & Wenzel 2008, S. 1; Banks 2010, S. 21; VDI 3633 Blatt 1, S. 8f). Das unterscheidet die Simulation von Experimenten im realen System, welche den regulären Ablauf gefährden können, Systemressourcen blockieren und mit einem höheren Kosten- und Zeitaufwand verbunden sind (VDI 3633 Blatt 1, S. 9). Der Einsatz einer Simulation bietet sich zudem an, wenn das zu analysierende System so komplex ist, dass andere Analyseverfahren nicht mehr anwendbar sind (Steven 2007, S. 483; VDI 3633 Blatt 1, S. 5). Bei manchen Systemen mit stochastischer Natur und zeitlich variablem Verhalten sind Simulationen die einzige Möglichkeit zur Analyse (Gutenschwager & Rabe et al. 2017, S. 38). Ein Produktionssystem ist ein solches System mit stochastischer Natur und komplexer Struktur (VDI 3633 Blatt 1, S. 5).

Im Kontext der dynamischen Lieferzeit-Preisgestaltung in der variantenreichen Produktion bietet sich der Einsatz einer Simulation aufgrund der genannten Eigenschaften an. Mit der Simulation wird es möglich, quantitativ **bewertbare Experimente** durchzuführen und die Allgemeingültigkeit eines Ansatzes nachzuweisen, indem die **Stabilität der Leistung in verschiedenen Szenarien** untersucht wird. Dies sind wichtige Anforderungen an den Lösungsansatz.

Zum weiteren Verständnis der Simulation an sich und zu der Frage, wie eine Simulationsstudie durchgeführt werden sollte, finden sich ergänzende Ausführungen in Anhang A1.

#### **2.4.2 Statistische Versuchsplanung und Tests in einer Simulationsstudie**

Ein wichtiger Bestandteil der Durchführung einer Simulationsstudie ist der Schritt „Experimente und Analyse“ (vgl. Anhang A1). Dahinter verbirgt sich die Bildung von Experimentplänen, das Aufstellen von Hypothesen und die Analyse der Ergebnisse, wodurch letztendlich Schlussfolgerungen gezogen werden können.

Das Aufstellen von Experimentplänen fällt in das Gebiet der statistischen Versuchsplanung. Das Ziel der Versuche ist es, den Zusammenhang zwischen Zielgrößen und für den Versuch ausgewählten Faktoren innerhalb eines Systems mit möglichst wenigen Versuchen herauszufinden (Siebertz, van Bebber & Hochkirchen 2017, S. 1f, 5–7). Die

Faktoren lassen sich unterscheiden in Steuerfaktoren, welche sich im Systembetrieb einstellen lassen, und Rauschfaktoren, welche bestenfalls im Experiment, aber nicht im Systembetrieb einstellbar sind (Kleppmann 2020, S. 13f). Das Wissen über den Zusammenhang zwischen Steuerfaktoren und Zielgrößen kann beispielsweise dazu eingesetzt werden, eine optimale Konfiguration des Systems zu finden oder das Antwortverhalten des Systems in bestimmten Szenarien vorherzusagen (Dean, Voss & Draguljić 2017, S. 1).

Unter den Versuchsplänen werden zwei Arten unterschieden: Vollfaktorielle Versuchspläne und teilstatistische Versuchspläne. Vollfaktorielle Versuchspläne zeichnen sich dadurch aus, dass sie alle Faktorkombinationen untersuchen und damit alle Wechselwirkungen sicher aufdecken (Siebertz, van Bebber & Hochkirchen 2017, S. 6; Kleppmann 2020, S. 129). Vollfaktorielle Versuchspläne haben jedoch den Nachteil, dass sie zu einer nicht mehr handhabbaren Versuchsanzahl führen können (Siebertz, van Bebber & Hochkirchen 2017, S. 28; Kleppmann 2020, S. 129). In vielen Fällen ist eine vollfaktorielle Versuchspläne jedoch nicht notwendig. Die teilstatistischen Versuchspläne versprechen hier Abhilfe. Sie reduzieren die Versuchszahl, fokussieren die Identifikation der wichtigsten Faktoren und geben ein Indiz über die Einflussrichtung verschiedener Parameter (Siebertz, van Bebber & Hochkirchen 2017, S. 129).

Um schlussendlich eine Aussage über die Leistung eines justierten Systems statistisch gesichert treffen zu können, bieten sich das Aufstellen von Hypothesen und das Anwenden statistischer Tests an. Zur Beurteilung der Leistungsfähigkeit eines neuen Systems wird typischerweise das neue System mit dem System im Ausgangszustand verglichen (Banks 2010, S. 481).

Nach Eley (2012, S. 22) besteht das Vorgehen bei Hypothesentests aus fünf Schritten:

1. Aufstellen der Nullhypothese  $H_0$  und Festlegung eines Signifikanzniveaus  $\alpha$ .
2. Bestimmung der Testgröße  $T^*$  und der Verteilung der Testgröße.
3. Bestimmung des Ablehnungsbereichs für das bestimmte Signifikanzniveau.
4. Berechnung des Testgrößenwerts.
5. Entscheidung über die Hypothese und Interpretation des Ergebnisses.

Die Nullhypothese  $H_0$  ist so aufzustellen, dass sie falsifizierbar ist (Crawley 2015, S. 3).

Zu jeder Nullhypothese gehört zudem eine Gegenhypothese  $H_1$  (Sibbertsen & Lehne 2021, S. 340). Die Gegenhypothese formuliert üblicherweise das Gegenteil der Nullhypothese (Sibbertsen & Lehne 2021, S. 340). Das Signifikanzniveau  $\alpha$  gibt die maximale

Wahrscheinlichkeit an, einen Fehler erster Art zu begehen, also die Nullhypothese  $H_0$  abzulehnen, obwohl sie korrekt ist (Sibbertsen & Lehne 2021, S. 343). Ein Fehler zweiter Art liegt vor, wenn die Nullhypothese  $H_0$  angenommen wird, obwohl sie falsch ist (Sibbertsen & Lehne 2021, S. 343). Typische Werte für Signifikanzniveaus liegen bei 1 %, 5 % oder auch 10 % (Sibbertsen & Lehne 2021, S. 344).

Um die Leistungsfähigkeit eines Systems im Vergleich zu einem anderen zu testen, bietet es sich an, die Differenz der Leistungsmittelwerte  $\bar{D}$  beider Systeme auf einen signifikanten Unterschied von einer Konstanten  $k_0$  zu testen (Banks 2010, S. 482). Der Mittelwert einer Stichprobe ist hierbei ein Schätzer für den Erwartungswert  $\mu$  (Sibbertsen & Lehne 2021, S. 218). Für den Vergleich zweier Mittelwerte auf einen signifikanten Unterschied findet der t-Test Anwendung (Crawley 2015, S. 91). Eine wichtige Grundvoraussetzung zur Anwendung des t-Tests ist, dass die Verteilung der Mittelwerte einer Normalverteilung folgt. Für Stichproben einer gewissen Größe, in der Literatur angegeben sind 30 oder mehr Ziehungen, ist diese Grundvoraussetzung gemäß dem zentralen Grenzwertsatz erfüllt (Sibbertsen & Lehne 2021, S. 352).

Zur Formulierung der Nullhypothese für einen Lagevergleich der Differenz zweier unbekannter Erwartungswerte  $\mu_D$  kann ein rechtsseitiger Test verwendet werden (Sibbertsen & Lehne 2021, S. 347, 366):

$$\text{Rechtsseitiger Test: } H_0: \mu_D \leq k_0 \text{ und } H_1: \mu_D > k_0 \quad \text{Formel 2-3}$$

$k_0$  gibt hierbei eine beliebige zu testende Konstante an. Im Kontext eines Systemvergleichs bietet sich ein rechtsseitiger Test mit einer Konstanten  $k_0 = 0$  an, um zu zeigen, dass das neue System signifikant besser als das bisherige ist. Anders ausgedrückt formuliert die Nullhypothese in diesem Fall, dass die Differenz der Erwartungswerte signifikant kleiner als null ist. Das Ziel ist es, diese Hypothese verwerfen zu können. Zur Überprüfung der Hypothese werden der t-Test und die Stichprobenmittelwerte herangezogen. Die Nullhypothese kann verworfen werden, wenn folgendes gilt (Sibbertsen & Lehne 2021, S. 351):

$$\text{Rechtsseitiger Test: } T^* > t_{n_{freigrad}, 1-\alpha} \quad \text{Formel 2-4}$$

Hierbei entspricht  $t_{n_{freigrad}, 1-\alpha}$  einem Wert der t-Verteilung bei  $n_{freigrad}$  Freiheitsgraden und einem Signifikanzniveau von  $\alpha$ . Im Kontext des t-Tests ergibt sich der Freiheitsgrad gemäß Formel 2-5 (Sibbertsen & Lehne 2021, S. 351).

$$n_{freigrad} = n_{stichprobe} - 1 \quad \text{Formel 2-5}$$

Die Teststatistik  $T^*$  des t-Tests ist wie folgt definiert (Sibbertsen & Lehne 2021, S. 299, 350f):

$$T^* = \frac{\bar{D} - k_0}{\hat{\sigma}_D} \cdot \sqrt{n_{stichprobe}} \quad \text{Formel 2-6}$$

In der Formel 2-6 kommt die Standardabweichung der Stichprobenverteilung  $\hat{\sigma}_D$  neu hinzu. Diese berechnet sich gemäß Formel 2-7.

$$\hat{\sigma}_D^2 = \frac{\sum_{t=1}^{n_{stichprobe}} (D_t - \bar{D})^2}{n_{stichprobe} - 1} \quad \text{Formel 2-7}$$

$D_t$  ist hierbei ein einzelner Differenzwert aus der Stichprobe mit einem Umfang von  $n_f$  Ziehungen.

Es kann geschlussfolgert werden, dass der t-Test geeignet ist, um in einem Vergleich von einem Lösungsansatz mit dynamischer Lieferzeit-Preisgestaltung und einem Status Quo zu einer statistisch gesicherten Aussage zu kommen. Hierzu wird ein Stichprobenumfang von mindestens 30 Ziehungen benötigt.

## 2.5 Methoden für die dynamische Lieferzeit-Preisgestaltung

Wie in der Einleitung motiviert, muss die dynamische Lieferzeit-Preisgestaltung IT-basiert funktionieren. Entsprechend wird die Auswahl der Methodengruppen von vornherein auf solche eingegrenzt. Des Weiteren sind wichtige Merkmale der Methoden die Skalierbarkeit und Adaptivität (Helmold 2020, S. ix, 2). Diese spiegeln sich in der Rechenzeit und dem Aufwand, um das Modell aktuell zu halten, wider. Neben der Ergebnisgüte entscheiden diese beiden Merkmale über die spätere Anwendbarkeit eines auf die Methode aufbauenden Lösungsansatzes. Um eine geeignete Methode für eine dynamische Lieferzeit-Preisgestaltung auszuwählen, bietet es sich an, zunächst einen Überblick über mögliche Methoden zu schaffen. In diesem Kapitel werden dementsprechend kurz verschiedene, relevante Methodengruppen vorgestellt. Im Einzelnen werden Reinforcement Learning, analytische Verfahren, Entscheidungsbaumverfahren und Heuristiken als Methodengruppen unterschieden.

Die erste Gruppe von Methoden ist das Reinforcement Learning. Das Reinforcement Learning ist ein Teilbereich des maschinellen Lernens und speziell für den Einsatz bei Problemen mit verketteten Entscheidungen, wie der dynamischen Lieferzeit-Preisgestaltung, geeignet (Rebala, Ravi & Churiwala 2019, S. 1f; Stasinski 2020, S. 471). Es besitzt hierbei ein hohes Maß an Adaptivität (Sutton & Barto 2018, S. 1f). Des Weiteren

ist das Reinforcement Learning geprägt durch vergleichsweise schnelle Ausführungszeiten (Plaat 2020, S. 56, 59; Sutton & Barto 2018, S. 73; Zhang & Yu 2020, S. 128). Dementsprechend spielt das Reinforcement Learning als Methode zur dynamischen Lieferzeit-Preisgestaltung in der Forschung bereits eine herausragende Rolle (Stasinski 2020, S. 471).

Als zweite Gruppe lassen sich die nicht-iterativen, analytischen Methoden identifizieren. Sie zeichnen sich durch manuelle, ableitungsbasierte Umformung von Gleichungen aus, um eine Lösungsformel zu finden. Die Lösungsformel führt dann durch Einsetzen der Modellparameter zur reproduzierbaren, optimalen Lösung (Pieper 2017, S. 11f). Voraussetzung für Methoden dieser Gruppe ist ein kontinuierliches Problem, wofür gegebenenfalls vereinfachende Annahmen getroffen werden müssen (Stein 2021, S. 23f; Pieper 2017, S. 11). Typische Vertreter bei den zugrundeliegenden Modellen sind einfache Warteschlangenmodelle (Domschke & Drexl et al. 2015, S. 224–230). Ein Beispiel hierfür ist das M/M/1-Modell, welches eine einzelne Kapazität, eine exponentialverteilte Bearbeitungszeit und die Ankunftsrate als Poisson-Prozess abbildet (Domschke & Drexl et al. 2015, S. 231).

Die dritte Gruppe von Methoden sind die Entscheidungsbaumverfahren, welche ebenfalls in einem vorgegebenen Modell optimale Lösungen finden (Domschke & Drexl et al. 2015, S. 134). Ein prominenter Vertreter der Entscheidungsbaumverfahren ist das Branch-and-Bound, welches allgemein für linear gemischt-ganzzahlige Modelle zum Einsatz kommen kann (Domschke & Drexl et al. 2015, S. 134, 148; Zimmermann 2008, S. 233). Ein weiterer Vertreter dieser Gruppe ist die dynamische Programmierung, die vor allem auf diskrete, mehrstufige Modelle angewendet wird (Domschke & Drexl et al. 2015, S. 169). Ist das mehrstufige Modell stochastischer Natur, so wird von stochastischer dynamischer Programmierung gesprochen (Zimmermann 2008, S. 245). Bei der dynamischen Programmierung wird im Gegensatz zum Reinforcement Learning von einem perfekten, finiten Modell ausgegangen (Sutton & Barto 2018, S. 73). Die dynamische Programmierung ist folglich mit einer entsprechenden Modellierungspflege verbunden, um gute Ergebnisse zu erzielen. Allgemein gilt, dass die Rechenzeit exakter Entscheidungsbaumverfahren mit der Komplexität eines Modells exponentiell steigen kann, was ihre Einsatzmöglichkeiten einschränkt (Jin, Li & Sivakumar 2013, S. 353; Domschke & Drexl et al. 2015, S. 133).

Im Fall zu langer Rechenzeiten, wie sie bei komplexen Modellen auftreten können, kann auch auf die vierte Gruppe von Methoden, die Heuristiken, zurückgegriffen werden (Jin,

Li & Sivakumar 2013, S. 353). Sie garantieren jedoch keine optimale Lösung, was sie von den analytischen Verfahren und den Entscheidungsbaumverfahren abgrenzt (Domschke & Drexl et al. 2015, S. 127; Werners 2013, S. 9). In dieser Arbeit seien die Heuristiken vom Reinforcement Learning insofern zu unterscheiden, dass sie keinen Lernmechanismus haben, sodass ihre Adaptivität auf eine manuelle Anpassung beschränkt ist.

Über die vier Gruppen hinaus könnten noch weitere Gruppen eingeführt werden. Hierzu gehören zum Beispiel Methoden der Spieltheorie oder physikalische Ersatzmodelle. Es hat sich jedoch gezeigt, dass diese im Stand der Forschung eine untergeordnete Rolle spielen. Sie werden daher nicht im Detail aufgeführt.

Aus diesem Kapitel lässt sich abschließend festhalten, dass der Aufwand zur Modellpflege und die Rechenzeit relevante Kriterien für die Wahl einer geeigneten Methode zur dynamischen Lieferzeit-Preisgestaltung sind. Nur, wenn ein Lösungsansatz ausreichend **skalierbar und adaptiv** ist, bleiben Rechenzeit und Modellpflege in einem angemessenen Rahmen. Hier sticht das Reinforcement Learning hervor, welches nachfolgend detailliert betrachtet wird.

## **2.6 Reinforcement Learning als Methode zur dynamischen Lieferzeit-Preisgestaltung**

Das Reinforcement Learning ist aufgrund seiner Eigenschaften eine aussichtsreiche Methode für eine dynamische Lieferzeit-Preisgestaltung (vgl. Kapitel 2.5). Zum besseren Verständnis des Reinforcement Learning werden nachfolgend dessen Grundlagen vermittelt (Kapitel 2.6.1) und ein Exkurs zum neuronalen Netzwerk als wichtiges Werkzeug gegeben (Kapitel 2.6.2). Anschließend wird eine Klassifikation der verschiedenen Reinforcement Learning Methoden vorgestellt (Kapitel 2.6.3) und das Actor Critic Verfahren als hybrider Ansatz im Detail erläutert (Kapitel 2.6.4).

### **2.6.1 Grundlagen des Reinforcement Learning**

Das Reinforcement Learning gehört zur Kategorie des maschinellen Lernens. Es ist weder dem überwachten noch dem unüberwachten Lernen zuzuordnen, sondern bildet eine eigenständige Kategorie des maschinellen Lernens (Sutton & Barto 2018, S. 2). Im konzeptionellen Modell des Reinforcement Learning kann ein Agent für eine Interaktion, auch Aktion genannt, bestimmte Parameter der Umwelt  $S_t$  zu einem Zeitpunkt  $t$  beobachten. Er wählt dann eine Aktion  $A_t$  basierend auf der Beobachtung aus. Die

Umwelt wird durch die Aktion des Agenten beeinflusst, sodass sich der Umweltzustand und damit die beobachtbaren Parameter nach der Aktion verändern ( $S_{t+1}$ ). Basierend auf einer Belohnungsfunktion erhält der Agent zudem eine Belohnung  $R_{t+1}$  in Abhängigkeit seiner Aktion bzw. der daraus folgenden Veränderung der Umwelt. Danach startet der Prozess von vorn. Das Ziel besteht nun darin, die kumulative Belohnung über einen Zeithorizont zu maximieren, sodass der Agent die beste Verhaltensstrategie entwickelt. (Sutton & Barto 2018, S. 6f)

Das beschriebene, allgemeine Modell des Reinforcement Learning wird in Abbildung 2-6 dargestellt.

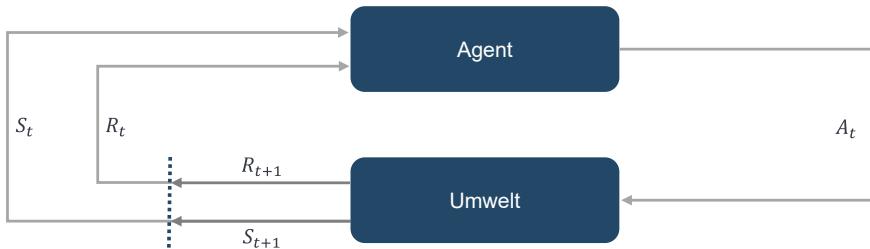


Abbildung 2-6: Konzeptionelles Modell des Reinforcement Learning nach (Sutton & Barto 2018, S. 48)

Alle Reinforcement Learning Methoden haben gemein, dass ihnen ein Markow-Entscheidungsproblem zugrunde liegt (Sutton & Barto 2018, S. 2). Es beschreibt mathematisch, wie ein lernender Agent mit einer Umwelt interagiert (Plaat 2020, S. 43f; Sutton & Barto 2018, S. 47-53). Eine zentrale Eigenschaft und gleichzeitig Voraussetzung für die Klassifikation als Markow-Entscheidungsproblem ist, dass eine Zustandsänderung einzig vom aktuellen Umweltzustand und der gewählten Aktion abhängt (Plaat 2020, S. 44). Das bedeutet, dass alle für die Zukunft relevanten Informationen vorheriger Umweltzustände im aktuellen Umweltzustand enthalten sein müssen. Der Umweltzustand erfüllt dann das Markow-Kriterium bzw. die Markow-Eigenschaft. (Sutton & Barto 2018, S. 49)

Mathematisch lässt sich das Markow-Entscheidungsproblem über folgendes Tupel darstellen (Plaat 2020, S. 44f):

$$< S, A, R, p, \gamma >$$

Formel 2-8

Mit  $S = \{s_1, s_2, \dots\}$  als der Menge möglicher Umweltzustände und  $S_t \in S$ .  $A = \{a_1, a_2, \dots\}$  als die Menge möglicher Aktionen. Hier gilt analog  $A_t \in A$ .  $R$  ist die Menge der möglichen Belohnungswerte. Auch hier gilt  $R_t \in R$ . Zur Ermittlung eines konkreten Belohnungswerts  $r = R_t$  kann eine Belohnungsfunktion  $r(s, a)$  aufgestellt werden. Sie ordnet jedem Zustand-Aktion-Paar einen erwarteten Belohnungswert aus der Menge aller möglichen Belohnungswerte  $R$  zu. Sie lässt sich wie folgt notieren (Sutton & Barto 2018, S. 49):

$$r(s, a) = E[R_t | s = S_{t-1}, a = A_{t-1}] = \sum_{r \in R} r \sum_{s' \in S} p(s', r | s, a) \quad \text{Formel 2-9}$$

$p$  gibt die Übergangswahrscheinlichkeit an. Diese Funktion beschreibt die Wahrscheinlichkeit, von einem bestimmten Zustand  $s = S_t$  durch eine bestimmte Aktion  $a = A_t$  in den Zustand  $s' = S_{t+1}$  überzugehen. Für die Übergangswahrscheinlichkeitsfunktion gilt nach (Sutton & Barto 2018, S. 48f):

$$p(s' | s, a) = \sum_{r \in R} p(s', r | s, a) \quad \text{Formel 2-10}$$

Aus dem beschriebenen Tupel des Markow-Entscheidungsproblems ist nun noch  $\gamma$  zu erläutern. Hierbei handelt es sich um den Diskontierungsfaktor. Er gibt an, wie viel Einfluss zukünftige Belohnungen auf das Verhalten des Agenten haben. Die Wirkung von  $\gamma$  wird bei Betrachtung der Ertragsfunktion  $G_t$  klar. Der zukunftsgerichtete Ertrag ergibt sich durch Aufsummierung der Belohnungswerte über die Zeitschritte. Positive Erträge bestärken das Handeln des Agenten und vice versa. Für Markow-Entscheidungsprobleme mit einer großen Anzahl oder sogar unendlich vielen Zeitschritten bzw. Zuständen kann der Ertrag sehr groß werden oder gegen unendlich gehen. Zudem kann der zukünftige Ertrag bei Unsicherheit unmöglich zu berechnen sein. Um dies zu vermeiden, wird mit einem Diskontierungsfaktor  $\gamma$  gearbeitet, wodurch zukünftige Erträge weniger Einfluss nehmen, je weiter sie in der Zukunft liegen. Es lässt sich nun der Ertrag  $G_t$  sowohl als unendliche Summe als auch als rekursive Summe definieren (vgl. Formel 2-11) (Sutton & Barto 2018, S. 54f). Der Ertrag und die hier aufgestellte Definition spielen später für den Lernprozess des Agenten eine essenzielle Rolle.

$$\begin{aligned} G_t &= R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \\ &= R_{t+1} + \gamma G_{t+1} \end{aligned} \quad \text{Formel 2-11}$$

Nachdem das Markow-Entscheidungsproblem vollständig beschrieben ist, soll als nächstes die Frage geklärt werden, wie der Agent in diesem Entscheidungsproblem seine Aktion wählt und lernt. Hierzu sei die Politikfunktion  $\pi(a|s)$  definiert (Sutton & Barto 2018, S. 58). Die Politik gibt die Wahrscheinlichkeit einer Aktion bei einem gegebenen bestimmten Umweltzustand an. Eine Politik sei besser als eine andere Politik, wenn sie zu einem höheren Ertrag führt (Sutton & Barto 2018, S. 62). Eine Politik wird als optimal bezeichnet, wenn sie besser oder gleich jeder anderen Politik ist (Sutton & Barto 2018, S. 62). Im Zuge des Lernprozesses wird die Politik des Agenten angepasst, sodass die gewählten Aktionen den Ertrag maximieren. Hier besteht stets ein Zielkonflikt zwischen Ausbeute und Exploration (Sutton & Barto 2018, S. 27): Um etwas über seine Umwelt zu lernen, muss der Agent Aktionen wählen, die er zuvor nicht gewählt hat (Exploration) auf die Gefahr hin, dass die Aktionen schlechter als die bisher bekannten sind. Gleichzeitig ist es das Ziel, den Ertrag zu maximieren (Ausbeute), was über die bekannten Aktionen üblicherweise besser gelingt als durch unbekannte Aktionen (Sutton & Barto 2018, S. 3). Die richtige Balance zwischen Ausbeute und Exploration ist eine der bestehenden Herausforderungen beim Reinforcement Learning, deren Güte sich an dem Erreichen der Zielgrößen ablesen lässt (Sutton & Barto 2018, S. 27; Plaat 2020, S. 52f).

Zusätzlich zur Politik können weitere Funktionen definiert werden, die den Ertrag in Abhängigkeit der Politik wiedergeben. So kann die Zustandswertfunktion  $v_\pi(s)$  gebildet werden (vgl. Formel 2-12). In dieser Formel ist zu beachten, dass  $E_\pi$  den Erwartungswert angibt, sofern eine spezifische Politik  $\pi$  verfolgt wird. Insgesamt drückt die Zustandswertfunktion  $v_\pi(s)$  also aus, welcher Ertrag bei gegebenem Zustand und gegebener Politik zu erwarten ist (Sutton & Barto 2018, S. 58).

$$v_\pi(s) = E_\pi[G_t | s = S_t] = E_\pi[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | s = S_t], \forall s \in S \quad \text{Formel 2-12}$$

Mithilfe der rekursiven Definition des Ertrags (vgl. Formel 2-11) lässt sich auch die Zustandswertfunktion rekursiv definieren, sodass sich die Formel 2-13 ergibt (Sutton & Barto 2018, S. 59). Diese Formel wird auch als Bellmansche Gleichung bezeichnet (Sutton & Barto 2018, S. 59).

$$\begin{aligned} v_\pi(s) &= E_\pi[G_t | s = S_t] \\ &= E_\pi[R_{t+1} + \gamma G_{t+1} | s = S_t], \forall s \in S \end{aligned} \quad \text{Formel 2-13}$$

In Ergänzung zur Zustandswertfunktion (Formel 2-13) gibt die Zustand-Aktion-Wertfunktion  $q_\pi$  (Formel 2-14) an, welcher Ertrag zu erwarten ist, wenn in einem Zustand irgendeine Aktion gewählt und ansonsten einer bestimmten Politik gefolgt wird (Sutton & Barto 2018, S. 58). Sie ist hilfreich, um in einem Zustand die verschiedenen Alternativen bei einer bestehenden Politik zu bewerten und damit zu bestimmen, ob sich die Politik zum Zeitpunkt  $t$  verbessern ließe.

$$q_\pi(s, a) = E_\pi[G_t \mid s = S_t, a = A_t] \quad \text{Formel 2-14}$$

## 2.6.2 Neuronale Netze zur Funktionsapproximation

Die Aufgabe des Reinforcement Learning ist es, die optimale Politik innerhalb eines Markow-Entscheidungsproblems zu finden. Hierzu sollte der Agent in der Lage sein, basierend auf den Belohnungen seine Politik anzupassen. Diese Anpassung einer allgemeinen Funktion, sodass für einen spezifischen Input ein vorgegebener Sollwert ausgetragen wird, ist formal betrachtet eine Funktionsapproximation. Für diese Funktionsapproximation gibt es verschiedene Ansätze. Für kleine Problemräume können beispielsweise Tabellen verwendet werden, um Zustände, Aktionen und resultierende Belohnungen zu speichern und nachzuschlagen (Sutton & Barto 2018, S. 45). Die Tabellen können in manchen Fällen zum Optimum führen (Sutton & Barto 2018, S. 195). Für größere Problemräume sind Lösungen, wie zum Beispiel Tabellen, nicht mehr möglich, insbesondere bei kontinuierlichen Werten, die theoretisch zu unendlich großen Tabellen führen (Sutton & Barto 2018, S. 195). Stattdessen wird eine Funktionsapproximation benötigt, die in der Lage ist, zu generalisieren (Sutton & Barto 2018, S. 195). Ein solcher Funktionsapproximator für das Reinforcement Learning ist das neuronale Netz (Calin 2020, S. viii, 201; Sutton & Barto 2018, S. 195, 223).

Ein neuronales Netz besteht aus Neuronen, die über Kanten miteinander verbunden sind. Ein Neuron hat gewichtete Eingänge und einen Verzerrungsfaktor (Weber 2020, S. 45). Die Summe aus gewichteten Eingangswerten und Verzerrungsfaktor wird mithilfe einer Aktivierungsfunktion in einen Ausgabewert übersetzt (Weber 2020, S. 45). Aktuell verwendete Aktivierungsfunktionen sind der Tangens Hyperbolicus ( $tanh$ ), die Sigmoid-Funktion ( $sig$ ) oder auch die Gleichrichter-Funktion (engl. „Rectified Linear Unit“ oder kurz ReLU mit der besonderen Form leakyReLU) (Aggarwal 2018, S. 11f, 133). Die leakyReLU verhindert im Vergleich zu ReLU ein vollständiges Ausschalten einzelner Neuronen im Trainingsprozess bei negativen Werten (Aggarwal 2018, S.

133). Ein exemplarischer Funktionsverlauf von tanh, Sigmoid, ReLU und leakyReLU Aktivierungsfunktion ist in Abbildung 2-7 zu sehen.

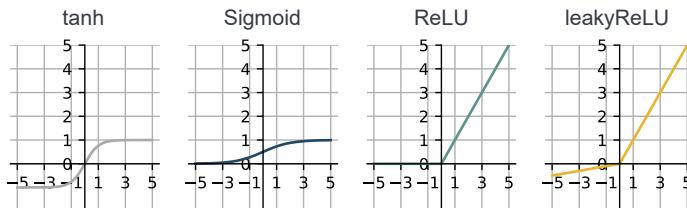


Abbildung 2-7: Darstellung von relevanten Aktivierungsfunktionen

Die Neuronen eines Netzwerks können in Schichten angeordnet und mehrere dieser Schichten hintereinander geschaltet werden (Aggarwal 2018, S. 17f). Es gibt hierbei mindestens eine Eingabe- und eine Ausgabeschicht. Dazwischen befinden sich beliebig viele Zwischenschichten. Typischerweise ist jedes Neuron einer Schicht mit allen Neuronen der nächsten Schicht verbunden (Aggarwal 2018, S. 17f). Es wird dann von einem vorwärts gerichteten neuronalen Netz gesprochen (Aggarwal 2015, S. 17). Abbildung 2-8 zeigt exemplarisch die Architektur eines vorwärtsgerichteten, neuronalen Netzes. Mithilfe einer Verlustfunktion wird im Rahmen des Lernprozesses die Diskrepanz der erzielten Ausgabe bei gegebenem Input und gewünschtem Output bestimmt (Aggarwal 2018, S. 7). Eine bekannte Verlustfunktion ist die aus der Regression bekannte Funktion der kleinsten Fehlerquadrate (Aggarwal 2018, S. 7).

Neben der Architektur des neuronalen Netzes ist die eigentliche Funktionsapproximation, umgangssprachlich Lernprozess genannt, ein kritischer Faktor für den erfolgreichen Einsatz eines neuronalen Netzes (Aggarwal 2018, S. 21–30). Bei vielschichtigen neuronalen Netzen ergeben sich zahlreiche mathematische Probleme im Lernprozess. So ist eine rechnerische Herausforderung im Lernprozess, den Gradienten der Verlustfunktion jedes Gewichtes zur Aktualisierung der Werte effizient zu finden. Eine partielle Ableitung und Lösung des resultierenden Gleichungssystems zur Findung der Nullstellen ist hierbei nicht effizient umsetzbar. Stattdessen wurde die Rückwärtspropagation (engl. Backpropagation) entwickelt, die Gradienten iterativ bestimmen kann (Aggarwal 2018, S. 21f). Diese Gradienten werden dann im Gradientenverfahren genutzt, um die Parameter von jedem Neuron in Richtung der optimalen Einstellung zu bewegen.

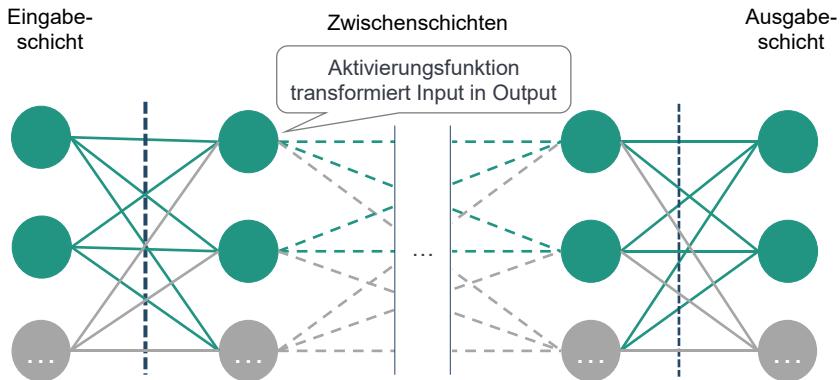


Abbildung 2-8: Grundlegende Architektur eines vorwärtsgerichteten neuronalen Netzes

Für die Aktualisierung der Neuronenparameter durch den Gradienten der Verlustfunktion wurden verschiedene Strategien entwickelt (Aggarwal 2015, S. 134). Ziel ist es unter anderem, ein Oszillieren der Werte um einen Zielwert und damit des ganzen Lernverfahrens zu vermeiden (Aggarwal 2015, S. 137f). Eine solche Strategie wird durch den Adam-Algorithmus umgesetzt (Aggarwal 2015, S. 140). Der Adam-Algorithmus ist weit verbreitet und verbindet die Vorteile verschiedener Strategien (Aggarwal 2015, S. 140).

Jedoch ist nicht nur die Gradientenstrategie ein wichtiger Mechanismus, um den Lernprozess zu beeinflussen. Auch die Lernrate entscheidet über die Konvergenz des Lernprozesses, in dem sie die Schrittweite einer Aktualisierung im Gradientenverfahren beeinflusst. Eine zu große Lernrate führt hierbei zu einer Oszillation des Lernprozesses, wohingegen ein zu kleiner Wert den Lernprozess stark verlangsamt und zudem die Gefahr birgt, mit dem Lernprozess früh in lokalen Minima zu verbleiben (Aggarwal 2015, S. 135). Als Ausweg aus dem Zielkonflikt zwischen schnellem Lernfortschritt und Konvergenz in eine möglichst gute Lösung bietet sich ein Lernratenabbau an. Hierbei wird zum Beispiel mithilfe der exponentiellen Abnahmefunktion die Lernrate sukzessive verkleinert (Aggarwal 2018, S. 135).

In der Formel 2-15 ist eine Gleichung zur Umsetzung eines Lernratenabbaus angegeben.

$$\alpha_t = \alpha_0 \cdot e^{-t \cdot k_{abba}} \quad \text{Formel 2-15}$$

$\alpha_t$  ist die Lernrate zum Zeitpunkt  $t$ .  $\alpha_0$  ist die Lernrate zum Zeitpunkt  $t = 0$  und  $k_{abba}$  ist die Abbaurate.

### 2.6.3 Klassifikation von Reinforcement Learning Verfahren

Innerhalb des Reinforcement Learning haben sich verschiedene Klassen von Verfahren herausgebildet (vgl. Abbildung 2-9). Es werden zunächst modellfreie und modellbasierte Verfahren unterschieden.

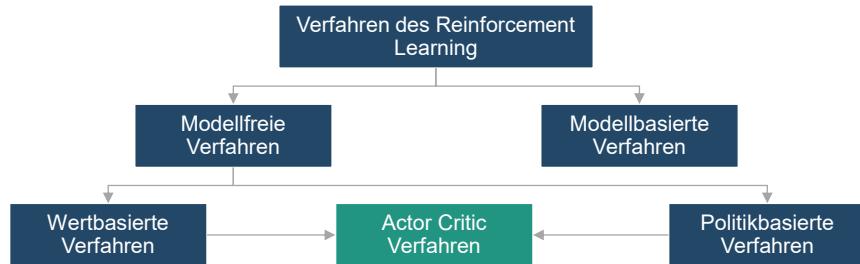


Abbildung 2-9: Klassifikation von Reinforcement Learning Verfahren

Von einem modellbasierten Verfahren wird gesprochen, wenn das Markow-Modell zur Bestimmung der Wertfunktion und optimalen Politik direkt verwendet werden kann (Rebala, Ravi & Churiwala 2019, S. 210). In diesem Fall kann von vornherein planerisch vorgegangen werden, also ohne mit der Umwelt interagieren zu müssen (Sutton & Barto 2018, S. 7; Zhang & Yu 2020, S. 127). Hierzu muss ein Modell der Umwelt in einer Weise vollständig vorliegen, dass es zur Optimierung der Politik unmittelbar eingesetzt werden kann (Zhang & Yu 2020, S. 127f). Demgegenüber stehen die modellfreien Verfahren. Bei diesen ist das Finden der optimalen Politik nur auf Basis des Ausprobierens verschiedener Aktionen und des Beobachtens der Veränderungen möglich (Zhang & Yu 2020, S. 127f). Die modellfreien Verfahren kommen in der Praxis häufiger vor als die modellbasierten Verfahren (Zhang & Yu 2020, S. 127f). Eine Ursache hierfür könnte der entfallende Aufwand für die Modellpflege bei den modellfreien Verfahren gegenüber den modellbasierten Verfahren sein (Plaat 2020, S. 56, 59; Zhang & Yu

2020, S. 128). Die modellfreien Verfahren werden wiederum unterschieden in politikbasierte Verfahren und wertbasierte Verfahren. Wertbasierte Verfahren zeichnen sich dadurch aus, dass versucht wird, den Ertrag einer Aktion zu approximieren (Sutton & Barto 2018, S. 321). Eine Aktion wird dann basierend auf der Wertfunktion gewählt. Bekannte Vertreter der wertbasierten Verfahren sind das Q-Learning und das Double Q-Learning. Der Begriff leitet sich aus der Zustand-Aktion-Wertfunktion ab, die mit „q“ bezeichnet wird. Beim Q-Learning und Double Q-Learning<sup>1</sup> wird versucht, diese Zustand-Aktion-Wertfunktion über die Zeit zu approximieren (Sutton & Barto 2018, S. 131). Demgegenüber bestimmen politikbasierte Verfahren eine Aktion direkt auf Basis des Zustandes als Input (Sutton & Barto 2018, S. 321). Ein bekannter Vertreter ist der REINFORCE Algorithmus (Sutton & Barto 2018, S. 326). Beide Verfahrensarten haben ihr Anwendungsgebiet, sodass kein Verfahren dominant besser als das andere ist (Sutton & Barto 2018, S. 323). Ein wichtiger Vorteil der politikbasierten Verfahren ist, dass sie mit kontinuierlichen Zustands- und Aktionsräumen umgehen können (Sutton & Barto 2018, S. 335). Demgegenüber sind wertbasierte Verfahren weniger anfällig gegen lokale Optima aufgrund ihrer besseren Stichprobeneffizienz (Zhang & Yu 2020, S. 129). Ein Hybrid aus wertbasiertem Verfahren und politikbasiertem Verfahren, welches versucht, die Vorteile aus beiden Verfahren zu vereinen, ist das Actor Critic Verfahren. Dieses wird detailliert im nächsten Abschnitt behandelt.

#### 2.6.4 Actor Critic Verfahren

Eine Kombination aus politikbasiertem und wertbasiertem Verfahren heißt Actor Critic Verfahren (Plaat 2020, S. 54f; Grondman & Busoniu et al. 2012, S. 1294). Hierbei werden zwei Funktionsapproximatoren als Reinforcement Learning Agent verwendet: Ein Politikapproximator (Actor) bestimmt die zu wählende Aktion in einem Umweltzustand und ein Wertapproximator (Critic) bestimmt den zu erwartenden Ertrag. Der Critic approximiert hierfür  $v_\pi(s = S_t)$ , indem er die Approximationsfunktion  $\hat{v}_\pi(s = S_t)$  bildet (Aggarwal 2018, S. 395). Dieser Wert wird einzig für das Training des Actor benötigt, sodass der Critic obsolet wird, sobald der Trainingsprozess abgeschlossen ist. Der Critic lernt über die Zeit aus den Beobachtungen, welcher Ertrag zu erwarten ist (Aggarwal 2018, S. 395). Im Ergebnis entsteht ein Verfahren, welches die Vorteile von politikbasierten und wertbasierten Verfahren vereint und durch die Kombination weitere Vorteile

---

<sup>1</sup> Double Q-Learning grenzt sich von Q-Learning ab, indem es zwei Approximationen der Zustand-Aktion-Wertfunktion nutzt, um ein Überschätzen von q zu vermeiden.

schafft (Arulkumaran & Deisenroth et al. 2017, S. 9; Ding & Huang et al. 2020, S. 92; Grondman & Busoniu et al. 2012, S. 1294). Das Actor Critic Verfahren ist in der Lage, sowohl mit kontinuierlichen Zustandsräumen, als auch mit kontinuierlichen Aktionsräumen umzugehen. Daher ist es auch geeignet, eine kontinuierliche Lieferzeit-Preisfunktion zu bestimmen, wie sie in Kapitel 2.3.1 bereits motiviert wurde. Der allgemeine Aufbau eines Actor Critic Modell ist in Abbildung 2-10 dargestellt.

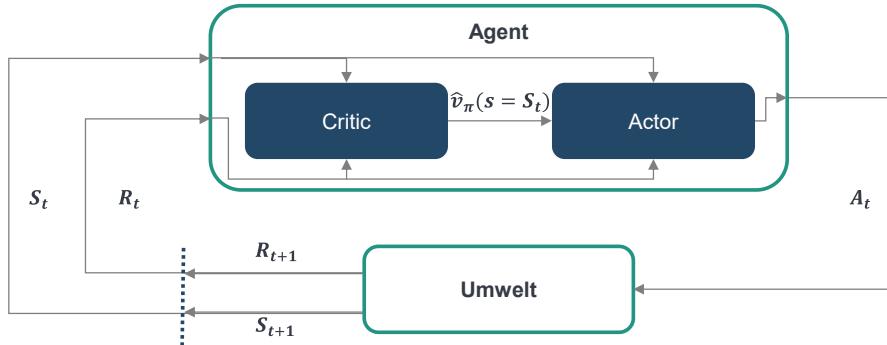


Abbildung 2-10: Actor Critic Modell in Anlehnung an Scharf, Helfenstein & Jäger (2021, S. 125)

Im allgemeinen Actor Critic Modell gibt der Critic einen Schätzwert  $\hat{v}_\pi$  für die Zustandswertfunktion aus. Folglich gibt der Critic stets eine einzelne Zahl aus. Die Ausgabe des Actor ist im Vergleich zum Critic komplexer: Um kontinuierliche Werte als Aktion zu erzeugen, werden diese nicht direkt ausgegeben, sondern aus einer vom Actor bestimmten Verteilung gezogen. Diese Verteilung kann beispielsweise eine Normalverteilung sein (Sutton & Barto 2018, S. 335). Zur Bestimmung einer Normalverteilung werden der Erwartungswert und die Standardabweichung benötigt. Diese werden von der Ausgabeschicht des Actor bestimmt (Sutton & Barto 2018, S. 335f). Für jeden festzulegenden Aktionswert ergeben sich somit zwei Werte in der Ausgabeschicht: Jeweils ein Mittelwert und eine Standardabweichung (Sutton & Barto 2018, S. 335). Dieses Vorgehen hat den wesentlichen Vorteil, dass der Actor den Grad der Exploration selbst festlegen kann. Je größer die Standardabweichung ist, desto mehr exploriert der Actor durch zufälliges Ziehen. Sobald eine gute Politik gefunden wurde, lässt sich die Standardabweichung auf null reduzieren, sodass der Erwartungswert gleich dem Ausgabewert ist.

Für das Training des Actor können folgende Gleichungen verwendet werden (Sutton & Barto 2018, S. 331f):

$$\theta_{t+1} = \theta_t + \alpha^\theta \delta_t \nabla \ln \pi(A_t | S_t, \vec{\theta}) \quad \text{Formel 2-16}$$

$$\delta_t = R_{t+1} + \gamma \hat{v}_\pi(S_{t+1}, \vec{w}) - \hat{v}_\pi(S_t, \vec{w}) \quad \text{Formel 2-17}$$

$\theta_t$  bezeichnet das einzelne Gewicht des Actor im neuronalen Netz zum Zeitpunkt  $t$ .  $\alpha^\theta$  ist die Lernrate des Actors und entspricht hier  $\alpha_t$  zum Zeitpunkt  $t$ .  $\delta_t$  ist der Vorteil (engl. Advantage) (Aggarwal 2018, S. 396). Die Berechnung des Advantage ist in Formel 2-17 zu sehen. Der Advantage gibt an, um wieviel die Aktion des Actor besser war als der Erwartungswert, den der Critic geschätzt hat. Der Advantage wird anstelle des reinen Erwartungswerts verwendet, weil dadurch die Belohnungen relativiert werden: Eine Aktion, die eine positive Belohnung erzielt, kann trotzdem unerwünscht sein, wenn zuvor mit einer anderen Aktion stets höhere Belohnungswerte erzielt werden konnten. Durch den Advantage setzt sich die bessere Aktion im Lernen schneller durch (Sutton & Barto 2018, S. 329, 330f). Für die Zustandswertfunktion, welche durch den Critic in Form eines neuronalen Netzes approximiert wird, wird hier die angepasste Form  $\hat{v}_\pi(S_{t+1}, \vec{w})$  verwendet. Hierdurch wird gezeigt, dass die Zustandswertfunktion von den Gewichtungsparametern  $\vec{w}$  beim Einsatz eines neuronalen Netzes abhängig ist. Der Critic bzw. das neuronale Netz des Critic kann gemäß folgender Lernfunktion trainiert werden (Sutton & Barto 2018, S. 332).

$$w_{t+1} = w_t + \alpha^w \delta_t \nabla \hat{v}(S_t, \vec{w}) \quad \text{Formel 2-18}$$

$w_t$  bezeichnet das einzelne Gewicht des Critic im neuronalen Netz zum Zeitpunkt  $t$ .  $\alpha^w$  ist die Lernrate des Actors und entspricht hier  $\alpha_t$  zum Zeitpunkt  $t$ .

### 3 Stand der Forschung

In diesem Kapitel wird der Stand der Forschung zum Thema dynamische Lieferzeit-Preisgestaltung in variantenreicher Produktion behandelt. Zunächst werden die Anforderungen an einen Lösungsansatz resümiert (Kapitel 3.1). Danach werden relevante Ansätze gruppiert nach ihrer zugrundeliegenden Methode vorgestellt (Kapitel 3.2). Aus den Anforderungen und den vorgestellten Ansätzen wird schließlich das Forschungsdefizit untersucht und abgeleitet (Kapitel 3.3).

Es wird kein Anspruch auf Vollständigkeit bei der Vorstellung der Ansätze erhoben. Ergänzend sei hier auf folgende Arbeiten und Sammelwerke verwiesen, die einen Überblick über das Themengebiet geben: (Pekgün 2010), (Gönsch 2017), (Kunath & Winkler 2019) und (Kubickova 2021).

#### 3.1 Anforderungen an den Lösungsansatz

Im Rahmen der Vorstellung der Grundlagen werden bereits verschiedene Anforderungen abgeleitet. Diese Anforderungen werden in Tabelle 3-1 zusammengefasst und referenziert:

*Tabelle 3-1: Übersicht über die Anforderungen an den Lösungsansatz*

Nr.	Kapitel	Anforderung
I	2.1.2	Dynamische Lieferzeit-Preisgestaltung in variantenreicher Produktion
II	2.1.2	Mehrere Produkte und Produktionsressourcen
III	2.2.1	Fristbasierte Ablaufplanung
IV	2.2.2	Berücksichtigung individueller, heterogener Kunden
V	2.2.2	Beidseitige Liefertermintreue
VI	2.3.1	Berücksichtigung von Informationsasymmetrie
VII	2.4.1	Experimentierbarkeit und Bewertbarkeit
VIII	2.4.1	Untersuchung der Stabilität
IX	2.5	Skalierbarkeit und Adaptivität der Methode

Nachfolgend werden nun verschiedene Lösungsansätze aus dem Stand der Forschung vorgestellt. Daran anschließend werden diese basierend auf den hier genannten Anforderungen hinsichtlich eines Forschungsdefizits bewertet (vgl. Kapitel 3.3).

### **3.2 Vorstellung der Lösungsansätze**

Die nachfolgend betrachteten Lösungsansätze werden zur besseren Übersichtlichkeit gemäß Kapitel 2.5 in Ansätze mit Reinforcement Learning, Ansätze mit analytischen Lösungsverfahren, Ansätze mit Entscheidungsbaumverfahren, Ansätze mit Heuristiken und sonstige Ansätze unterteilt. Die Zuordnung erfolgt anhand der dominanten Methode.

#### **3.2.1 Ansätze mit Reinforcement Learning**

Vengerov (2008) erforscht die Gewinnmaximierung für Dienstleister im Bereich der Stromnetzinfrastruktur durch eine dynamische Preispolitik unter Berücksichtigung des individuellen Kundennutzens und der erwarteten Verspätung von Aufträgen. Der Bedarf nach einer dynamischen Preispolitik besteht durch eine Vielzahl an Wettbewerbern, welche jeweils unterschiedliche Preispolitiken verfolgen. Der Autor nutzt hierzu einen Reinforcement Learning Ansatz in einem Modell mit Unsicherheit und mit konkurrierenden Unternehmen im Umfeld. Die Umweltzustände sind im Modell nur teilweise beobachtbar. Die Kunden werden mit individuellen Nutzenfunktionen modelliert. Der Autor kommt zu dem Ergebnis, dass das Unternehmen einen höheren Gewinn erzielen kann, wenn es neben einem Standardtarif auch einen Premiumtarif anbietet, sofern die Konkurrenz weiterhin statische Preise anbieten würde. (Vengerov 2008)

Arredondo & Martinez (2010) fokussieren ein Auftragsannahme-Problem in einer MTO-Fertigung zur Gewinnmaximierung. Im Modell gilt es zu entscheiden, ob ein Auftrag angenommen, verhandelt oder abgelehnt wird. Diese Entscheidung wird durch ein Q-Learning-Algorithmus getroffen. Der Algorithmus wird in verschiedenen Szenarien hinsichtlich Adaptivität positiv getestet. Der Ansatz zeigt eine signifikant bessere Performance gegenüber der Vergleichsheuristik von Hing, van Harten & Schuur (2007). Letztere selektiert Aufträge basierend auf dem Ertrag-Produkteinheitsquotient, der höher als ein fester Schwellwert sein muss. Die Autoren stellen fest, dass die Festlegung eines Ausbeute-Explorations-Verhältnisses eine Herausforderung beim Einsatz des Ansatzes ist. (Arredondo & Martinez 2010)

Li, Wang & Sawhney (2012) untersuchen den Einsatz von Q-Learning in einer MTO-Produktion zur Lösung einer Entscheidungsproblematik, die durch drei Aspekte charakterisiert ist: Das Annehmen oder Ablehnen von Aufträgen, das Festlegen eines Lie-

ferzeit-Preis-Tradeoffs und die Möglichkeit steigender Nachfrage gegenüber begrenzten Ressourcen. Ziel ist es, den Gewinn zu maximieren. Als Vergleichsansatz nutzen die Autoren zum einen eine FCFS-Annahmesteuerung und zum anderen eine Heuristik von Hing, van Harten & Schuur (2007). Hierbei zeigt sich die Dominanz des Q-Learning-Ansatzes. Darüber hinaus untersuchen die Autoren die Adaptivität des Ansatzes bei Veränderung von Nachfrage und Kapazität. Das Modell inkludiert lediglich eine Maschine. Die Autoren stellen fest, dass ihr Ansatz zwar die wenigsten Aufträge annimmt, dafür jedoch die profitabelsten. (Li, Wang & Sawhney 2012)

Hao, Yu & Wu (2013) entwickeln ein Modell der MTO-Fertigung für die dynamische Gestaltung von Preis und Lieferzeit. Als Algorithmus kommt ein Q-Learning zum Einsatz. Der Agent bestimmt für zufällige Aufträge ein Lieferzeit-Preis-Angebot unter Berücksichtigung von zur Verfügung stehenden Ressourcen, dem Ertrag des Auftrags und potenziell zukünftigen Aufträgen. Der Kunde entscheidet anschließend, ob er das Angebot annimmt oder ablehnt. Die Annahme bzw. Ablehnung ist dabei als Zufallsereignis modelliert, wobei die Wahrscheinlichkeit direkt von der Lieferzeit und dem Preis abhängt. Die Zielmetrik ist die Maximierung der durchschnittlichen Einnahmesumme. Als zusätzliches Ziel wird die Liefertreue aufgenommen. Die Autoren können zeigen, dass ihr Ansatz gegenüber statischen Lieferzeit-Preis-Angeboten sowohl höhere Einnahmen als auch eine höhere Liefertreue erreicht. (Hao, Yu & Wu 2013)

Lu, Hong & Zhang (2018) entwickeln einen Algorithmus auf Basis von Q-Learning, welcher im Bereich der Energiewirtschaft dynamisch den Endkundenpreis bestimmt. Die Autoren argumentieren die Wahl des Algorithmus mit der notwendigen Möglichkeit zur Adaption der Umweltveränderungen. Als beobachtbare Umweltzustände für das Q-Learning nutzen die Autoren das Zufriedenheitsniveau der Kunden, Nachfrageprofile und den Einkaufspreis des Unternehmens. In der Validierung kommen die Autoren zu dem Ergebnis, dass ihr Algorithmus zu einer Verbesserung auf Kunden- und Anbieterseite führt, da es zu niedrigeren Kosten und höheren Gewinnen kommt. Außerdem stellen sie fest, dass es mit ihrem Ansatz zu weniger Energielastspitzen kommt, wodurch die Zuverlässigkeit des Gesamtsystems steigt. Als weiteres Ergebnis der Anwendung sinkt der Gesamtenergieverbrauch. (Lu, Hong & Zhang 2018)

Bahrami, Chen & Wong (2021) forschen ebenfalls im Bereich der Energiewirtschaft. Sie entwickeln hier ein Actor Critic Modell, mit dem sie die Kundennachfrage bei Spitzenlasten über eine dynamische Erhöhung des Energiepreises nach hinten verschieben wollen. Der Kunde kann im Modell auswählen, wann er Energie beziehen möchte.

Seine Präferenzen werden über Beschwerdekosten repräsentiert. Diese sind dem agierenden Unternehmen jedoch nicht bekannt, sondern müssen aus dem Verhalten der Kunden antizipiert werden. In der Validierung können die Autoren zeigen, dass nicht nur weniger Spitzenlasten auf Seiten der Energieversorger auftreten, sondern auch Kunden insgesamt weniger Kosten tragen müssen. Das Actor Critic Modell wird außerdem mit Q-Learning und Double Q-Learning verglichen. Die Autoren kommen zu dem Schluss, dass ihr Actor Critic Modell die dominante Lösung ist, weil es schneller konvergiert als die Vergleichsmodelle. (Bahrami, Chen & Wong 2021)

Es sei zu diesem Kapitel noch zu ergänzen, dass es zahlreiche Ansätze mit Reinforcement Learning für eine dynamische Lieferzeit-Preisgestaltung aus dem Dienstleistungsbereich gibt (vgl. Seo & Chang et al. (2021), Bondoux & Nguyen et al. (2020) oder Gosavii, Bandla & Das (2002)). Exemplarisch sei hier die Bepreisung von Flugtickets genannt. Es könnte die Frage auftreten, warum diese Ansätze nicht betrachtet wurden. Der Grund dafür ist, dass die Rahmenbedingungen in der Dienstleistungswirtschaft von der Natur her grundlegend anders sind als die in der Produktion. Ein markanter Aspekt ist die Lagerbarkeit. Hinsichtlich der Ansätze aus der Energiewirtschaft lässt sich sagen, dass Energie einer Lagerbarkeit, zum Beispiel mithilfe von Akkumulatoren oder Pumpspeicherwerkten, nahe kommt. Große Kraftwerke, in welchen nutzbare Energie erzeugt wird, haben eine mit klassischer Produktion vergleichbare Natur. Energie unterscheidet sich damit von anderen Dienstleistungen. Daher wurden Ansätze aus der Energiewirtschaft in die Betrachtung eingeschlossen, auch wenn Energie kein klassisches Beispiel aus dem Bereich der Produktion sein mag.

### **3.2.2 Ansätze mit analytischen Lösungsverfahren**

Ahn, Gümüş & Kaminsky (2007) heben in ihrer Arbeit die zunehmende Schwierigkeit heraus, Marketing und Produktion zu vereinen: Durch die Beeinflussung der Nachfrage durch das Marketing und die resultierenden Schwankungen wird die Produktion unidirektional beeinflusst. Daher stellen die Autoren ein mathematisches Optimierungsmodell auf, welches den Erlös maximiert. Dieses enthält eine Mehrzahl an Kunden und Produktionskapazitäten. Jedoch nehmen die Autoren eine Mehrzahl an Annahmen auf, um das Modell analytisch lösbar zu machen. Hierzu zählt beispielsweise eine lineare Nachfragekurve. Die Autoren lösen dann das Modell für den einperiodischen Fall. Für den mehrperiodischen Fall nutzen die Autoren stattdessen eine selbstentwickelte Heuristik. Letztendlich können sie nachweisen, dass die Modellierung von Netzwerk und

Produktionsstandort im Rahmen einer dynamischen Lieferzeit-Preisgestaltung von Vorteil ist. (Ahn, Gümüş & Kaminsky 2007)

Ziel der Arbeit von Feng, Liu & Liu (2011) ist es, einen Ansatz zu entwickeln, der die Produktionsfähigkeit, die Kosten und das Kundenverhalten bei der Bestimmung eines Lieferzeit-Preis-Angebots berücksichtigt. Sie modellieren die Produktion mit einer FIFO-Steuerung. Basierend auf ihren Annahmen können sie eine optimale Strategie zur Gewinnmaximierung mathematisch nachweisen und das Gesamtproblem in eine Sequenz von drei Teilproblemen überführen: Zunächst wird entschieden, ob ein Auftrag angenommen wird, dann wird die Lieferzeit festgelegt und zuletzt der Preis abhängig von der Lieferzeit bestimmt. Im Ausblick schlagen die Autoren vor, das Modell unter anderem um Strafen für zu frühe Auslieferungen zu ergänzen. (Feng, Liu & Liu 2011)

Afèche (2013) untersucht Lieferzeit-Preis-Angebote zur effizienten Interaktion mit heterogenen Kunden, die unterschiedlich sensibel auf Preise und Lieferzeiten sind. Ziel ist es, den Erlös zu maximieren. Er modelliert die Produktion hierzu als M/M/1-Warteschlangenmodell. Im ersten Fall geht er von perfekter und vollständiger Information aus, sodass es dem Unternehmen gelingt, den Kunden optimale, individuelle Angebote hinsichtlich der Umsatzmaximierung zu machen. Im zweiten Fall herrschen imperfekte Informationen. Gegenüber der perfekten Informationslage verliert der Hersteller an Umsatz. Jedoch findet der Autor im Rahmen seiner Annahmen eine Lösung mit zwei Angeboten, die besser als ein Standardangebot ist. Hierbei ist ein Angebot durch einen Preisaufschlag für kürzere Lieferzeiten bestimmt. Die Arbeit impliziert, dass eine Be- trachtung des Kundennutzens kritisch für ein effizientes Leistungsangebot ist. (Afèche 2013)

Ata & Olsen (2013) entwickeln einen Ansatz, um Kunden individuelle Lieferzeit-Preis- menüs anzubieten. Es werden hierbei nur zwei Klassen von Kunden angenommen, welche unterschiedliche Lieferzeitpräferenzen haben. Ziel ist es, das optimale Menü zu finden, sodass der Erlös maximiert wird. Bei der Einplanung werden die Aufträge statt eines FCFS-Prinzips nach Umsatz priorisiert. Eine Besonderheit des Modells ist, dass Kunden eine S-Funktion zur Bestimmung der Verspätungskosten haben. Das bedeutet, dass die Verspätungskosten degressiv steigend sind. Die Autoren untersuchen ihr Modell teilweise durch Grenzwertuntersuchungen und teilweise durch einschränkende Annahmen, um Schlussfolgerungen ziehen zu können. (Ata & Olsen 2013)

Yu, Zhao & Sun (2013) untersuchen optimale Politiken in einem Monopol und Duopol hinsichtlich Preis, Investitionen und Lieferzeiten. Die zentrale Zielgröße ist der Gewinn. Die Produktion wird als M/M/1-Warteschlange modelliert. Sie nehmen hierzu an, dass es heterogene Kunden mit verschiedenen Sensitivitäten für Wartezeiten gibt. Außerdem herrscht in dem Modell vollständige und perfekte Information. Das fokale Unternehmen hat nur die Möglichkeit, entweder genau ein oder zwei Lieferzeit-Preis-Paare anzubieten. Im Duopol sind die Autoren in der Lage, im Rahmen ihrer Annahmen ein Gleichgewicht nachzuweisen. Die Autoren stellen fest, dass Duopolisten mit kürzeren Lieferzeiten nicht zwangsläufig höhere Gewinne erzielen. Jedoch ist es immer eine dominante Lösung, unterschiedliche Lieferzeiten anzubieten. (Yu, Zhao & Sun 2013)

Afèche & Pavlin (2016) nutzen ein M/M/1-Warteschlangenmodell zur Bestimmung eines gewinnoptimierenden Lieferzeit-Preisangebots. Sie nehmen heterogene, zeitsensitive Kunden an und gehen hier von einer Informationsasymmetrie aus. In diesem Modell untersuchen die Autoren vorteilhafte Strategien in Abhängigkeit der Lieferzeitsensitivität des Kunden und der vorhandenen Kapazität in der Produktion. Konkret gehen sie drei Fragen nach: Erstens, welcher von drei Kundentypen eher bedient werden sollte. Zweitens betrachten sie, ob und welchen Kundentypen exklusive Lieferzeit-Preismenüs angeboten werden sollten. Hierbei untersuchen sie auch, ob eine Zusammenfassung von Kundentypen unter einem Angebot sinnvoll ist. Drittens untersuchen sie, inwiefern das optimale Menü von Kapazität, Marktgröße und Präferenzen abhängig ist. Im Rahmen ihrer Untersuchung ist die Bestimmung des Angebots jedoch rein statisch. (Afèche & Pavlin 2016)

Albana, Frein & Hammami (2018) untersuchen eine gewinnoptimale Lieferzeitpolitik mit einem M/M/1-Warteschlangenmodell. Die Arbeit grenzt sich selbst gegenüber anderen Arbeiten mit M/M/1-Warteschlangenmodellen ab, indem variable Betriebskosten angenommen werden: Die Betriebskosten steigen mit kürzeren Lieferzeiten. Die Autoren modellieren die Produktion dergestalt, dass es für einen bestimmten Preis und eine bestimmte Kapazität eine minimale Lieferzeit gibt. Dies hängt mit der resultierenden Nachfrage zusammen. Um die minimale Lieferzeit weiter zu verkleinern, muss entweder die Nachfrage über den Preis beeinflusst oder weitere Kapazitäten beschafft werden. In der Lösung zeigt sich der generelle Zusammenhang, dass Lieferzeit-Angebote bei lieferzeitsensitiven Kunden kürzer und umgekehrt bei preissensitiven Kunden län-

ger werden sollten. Da die Betriebskosten bei kürzeren Lieferzeiten im Modell zunehmen, verkürzen sich die Lieferzeiten für lieferzeitsensitive Kunden weniger stark gegenüber den Vergleichsmodellen. (Albana, Frein & Hammami 2018)

Weitere Arbeiten mit Bezug zum Themengebiet finden sich von Wu & Kazaz et al. (2012), Feng & Zhang (2017) und Benioudakis, Burnetas & Ioannou (2021). Diese Arbeiten treffen spezifische Annahmen oder gelten nur in spezifischen Situationen, wie zum Beispiel die Zahlung einer fixen Eintrittsgebühr durch den Kunden, sodass die Lösungen in der allgemeinen Betrachtung variantenreicher Produktion eine untergeordnete Rolle spielen.

### **3.2.3 Ansätze mit Entscheidungsbaumverfahren**

Chaharsooghi, Honarvar & Modarres (2011) nutzen einen mehrstufigen Ansatz der stochastischen dynamischen Programmierung zur Bestimmung von Preis, Lieferzeit und Produktion für die MTO-Fertigung. Ziel ist es, den Gewinn zu maximieren. In ihrem Ansatz werden die Preise für jeweils einen Zeithorizont fixiert. Die Nachfrage ist hierbei vom Preis abhängig. Sie zeigen simulativ, dass die dynamischen Lieferzeit-Preis-Angebote vorteilhaft gegenüber statischen Lösungen sind. Als Ausblick schlagen die Autoren vor, zukünftig die Preise und Lieferzeiten kontinuierlich und dynamisch zu bestimmen. (Chaharsooghi, Honarvar & Modarres 2011)

Hao & Yu (2012) erforschen die Entscheidungssituation der Auftragsannahme in einer MTO-Fertigung und nutzen eine dynamische Programmierung als Algorithmus zur Entscheidungsfindung. Als zentraler Zielwert wird die Einnahmensumme verwendet. Die Aufträge der Kunden sind einzig über die Ankunftszeit, die Lieferzeit und den Preis definiert. Die Ankunftszeit wird als Poisson-Verteilung modelliert. Die restlichen Parameter sind gleichverteilt. Eine Verhandlung ist im Modell nicht vorgesehen, sodass entweder ein Auftrag angenommen oder abgelehnt wird. Wenn ein Auftrag nicht eingeplant werden kann, weil es nicht ausreichend Kapazität gibt, wird er automatisch abgelehnt. Die Autoren vergleichen den Algorithmus mit einer FCFS-Steuerung und zeigen, dass ihr Algorithmus besser abschneidet. Als Ausblick sehen die Autoren die Integration der Lieferzeit und der Produktion in die Entscheidungsproblematik. (Hao & Yu 2012)

Jin, Li & Sivakumar (2013) untersuchen die Problemstellung der Lieferzeitgestaltung bei vordefinierten Preisen mit dem Ziel, den Gewinn zu maximieren. Der Fokus der Arbeit liegt hierbei auf der Untersuchung der Recheneffizienz verschiedener Ansätze. Sie stellen hierzu ein gemischt-ganzzahliges Optimierungsproblem auf, in welchem sie

eine einzelne Produktionsressource und verschiedene lieferzeitsensitive Kundengruppen betrachten. Trotz der Vereinfachungen ist das Optimierungsproblem als NP-schwer einzustufen. Zur Lösung des Problems entwickeln die Autoren eine Heuristik und wenden zum Vergleich einen Branch-and-Bound-Algorithmus an, der das Finden einer optimalen Lösung garantiert. Jedoch wird die Lösung mithilfe des Branch-and-Bound-Algorithmus erwartungsgemäß nicht in polynomieller Zeit erreicht. Die Autoren merken an, dass in einem realen System der Preis ebenfalls Einfluss auf das Kundenverhalten hat und daher das Modell entsprechend angepasst werden müsste. (Jin, Li & Sivakumar 2013)

Guhlich, Fleischmann & Stolletz (2015) entwickeln eine stochastische dynamische Programmierung für ein Revenue-Management-Modell im Bereich der Variantenfertigung. Der Gewinn des fokalen Unternehmens soll maximiert werden, indem Aufträge angenommen oder abgelehnt werden. Außerdem wird eine Lieferzeit dynamisch bestimmt und dem Kunden angeboten. Der Kunde nimmt stets das Lieferzeit-Angebot an, jedoch werden Verspätungskosten des Kunden mit berücksichtigt. Die Preise für die Aufträge sind exogen vorgegeben und können nicht beeinflusst werden. Bei der Planung der Produktion werden zukünftige Aufträge antizipiert, sodass Aufträge später eingeplant werden, wenn kurzfristige Aufträge mit höherem Gewinn zu erwarten sind. Zur Terminierung der Aufträge entwickeln die Autoren eine Heuristik, die gegenüber der im Nachhinein bestimmten optimalen Lösung gute Näherungen liefert. (Guhlich, Fleischmann & Stolletz 2015)

Ouazene & Yalaoui et al. (2017) stellen fest, dass es für Industrieunternehmen essenziell ist, die Preisgestaltung mit der Produktion abzustimmen. Sie stellen ein mathematisches Optimierungsmodell auf, welches Kapazitäten, Lagerbestände und Nachfrage für ein Produkt abbildet. Die Nachfrage ist hierbei eine Funktion des Preises. Ziel ist die Gewinnoptimierung. Zur Lösung des resultierenden, nicht linearen, gemischt-ganzzahligen Optimierungsproblems nutzen die Autoren die sogenannte Outer-Approximation. Hierbei wird das Problem in zwei Teilprobleme geteilt, um obere und untere Grenzen zu finden. In der Validierung greifen die Autoren auf eine lineare Nachfragefunktion zurück, modellieren zwei Produkte und bilden vier verschiedene Szenarien. Eine tiefergehende Integration der Fertigungsabläufe in das Modell bleibt aus. (Ouazene & Yalaoui et al. 2017).

Garmdare, Lotfi & Honarvar (2018) nutzen ein nichtlineares, gemischt-ganzzahliges Optimierungsmodell im Bereich der MTO-Fertigung. Ziel ist es, den Gewinn mithilfe von

Lieferzeit, Preis und Planung zu optimieren. Im Rahmen des Modells modellieren die Autoren zwei konkurrierende Hersteller (Duopol) unter Berücksichtigung des aktuellen Auftrags- und Planungsstandes im Produktionssystem des fokalen Herstellers. Hierdurch sind sie in der Lage, den Einfluss des konkurrierenden Herstellers explizit zu untersuchen. Die Nachfrage wird in Abhängigkeit von Lieferzeit und Preis modelliert. Anhand einer Fallstudie können die Autoren nachweisen, dass eine Gewinnsteigerung durch ihr Optimierungsmodell möglich ist. In ihrer Untersuchung zeigen die Autoren außerdem, dass die Ergebnisse auch robust sind. (Garmdare, Lotfi & Honarvar 2018)

Öner-Közen & Minner (2018) untersuchen den Einfluss einer dynamischen Lieferzeit-Preis-Bestimmung auf den Gewinn und zentrale Produktionskennzahlen. Sie modellieren hierzu eine Fertigung mit einer Bedienstation und einer begrenzten Warteschlange als ein Markow-Entscheidungsproblem. Das fokale Unternehmen versucht, die Nachfrage über die Lieferzeit-Preis-Funktion zu beeinflussen. Kunden haben dementsprechend die Wahl, eine Lieferzeit-Preis-Kombination zu akzeptieren oder abzulehnen. Die heterogenen Kunden tun dies auf Basis ihrer individuellen Präferenzen. Im Falle einer Beauftragung wird der Auftrag verspätungsminimierend eingeplant. Die Lösung des Entscheidungsproblems erfolgt numerisch. Die Autoren stellen fest, dass die Einbeziehung des Produktionsstatus in die Bestimmung der Lieferzeit-Preis-Funktion zu einer höheren Effizienz des Systems führt, sofern es unterschiedlich sensitive Kunden gibt. (Öner-Közen & Minner 2018)

Esfandiari, Mashreghi & Emami (2019) entwickeln ein gemischt-ganzzahliges Optimierungsmodell, mit dem die Autoren die Auftragsannahme-, Auftragsreihenfolge- und Preissetzungsproblematik abbilden und lösen. Das zentrale Ziel ist die Gewinnmaximierung. Die Betrachtung wird beschränkt auf eine Mehrzahl von Maschinen, die alle parallel laufen und nicht gleich sind. Die Planung der Aufträge erfolgt auf Basis des vereinbarten Liefertermins, der geplanten Bearbeitungszeit, der Rüstzeit, planungsspezifischen Kosten und dem Produktpreis. Das Modell wurde als Erprobung bis zu einem Umfang von dreißig Aufträgen und sechs Maschinen getestet. Der Ansatz zeigt, dass eine integrierte Betrachtung von Produktpreis und Produktionssystem möglich und vorteilhaft ist. Jedoch herrscht im Modell vollständige und perfekte Information. Die Autoren geben als Ausblick an, dass eine Untersuchung des Problems mit Heuristiken sinnvoll wäre. (Esfandiari, Mashreghi & Emami 2019)

Auch Wang & Wang (2019) bilden das Problem der integrierten Preisgestaltung und Auftragsplanung als gemischt-ganzzahliges Optimierungsmodell ab. Sie nehmen hierbei eine diskrete Preis- und Nachfragefunktion an. Hierdurch vermeiden die Autoren eine nichtlineare Zielfunktion. Es soll hierbei der Erlös maximiert werden. Kunden und Produktion werden nicht im Detail modelliert. Die Autoren validieren das Modell anhand eines Anwendungsfalls mit drei Produkten und jeweils drei Preis-Nachfrage-Tupel. Das Beispielproblem wird in 400 Sekunden gelöst. Es konnte hierbei ein optimales Ergebnis gefunden werden, in dem jedoch nicht alle Aufträge angenommen wurden. Im Optimum wurde das günstigste Produkt gar nicht produziert, jedoch das teuerste auch nur in drei von fünf Anfragen. Nach eigener Aussage muss der Ansatz jedoch deutlich verbessert werden, um für Probleme auf industriellem Niveau einsetzbar zu sein. Die Autoren schlagen hierfür heuristische Ansätze und eine detailliertere Abbildung des Gesamtsystems vor. (Wang & Wang 2019)

Weitere Arbeiten im Themenbereich der Entscheidungsbaumverfahren sind Charnsirisakskul, Griffin & Keskinocak (2006) und So & Song (1998). Die Verfahren und Themenschwerpunkte der Lösungsansätze sind jedoch für das Vorhaben einer dynamischen Lieferzeit-Preisgestaltung im Kontext variantenreicher Produktion nur begrenzt relevant, da sie die Produktion an sich nur unzureichend abbilden. Die Ansätze werden daher nicht weiter beleuchtet.

### **3.2.4 Ansätze mit Heuristiken**

Çelik & Maglaras (2008) modellieren die Produktion eines MTO-Herstellers als Markow-Entscheidungsproblem mit mehreren Produkten und einer Maschine, bei dem der Hersteller Preis und Lieferzeit frei festlegen kann. Ziel ist die Gewinnmaximierung. Die Autoren heben drei Aspekte ihres Modells besonders hervor: Erstens den Einfluss von Lieferzeit und Preis auf die Nachfrage. Zweitens eine Beschaffung aus zwei Quellen, die es ermöglicht, die Lieferzeit zu bestimmten Kosten zu beschleunigen. Drittens das Zusammenspiel aus Produktionssteuerung und Nachfragebeeinflussung zur Umsatzgenerierung. Im Modell ist es möglich, die Bearbeitungsreihenfolge der Aufträge frei festzulegen. Die Kunden entscheiden über eine Bestellung auf Basis ihrer Präferenzen. Das bedeutet, sie erhalten auf eine Produktanfrage ein Lieferzeit-Preis-Angebot. Die Präferenzen der Kunden im Modell bestimmen die Autoren zuvor über eine Marktforschung. Als Lösungsverfahren kommt eine selbstentwickelte Heuristik zum Einsatz. Sie

---

vergleichen ihren Ansatz mit statischen Angeboten. Die Autoren untersuchen den Einfluss verschiedener Parameter ohne jedoch allgemeingültige Resultate aufzuzeigen. (Çelik & Maglaras 2008)

Jin (2012) untersucht Produktionsplanungsprobleme bei variablen Lieferzeit-Preiskombinationen. Es wird hierzu die Sicht eines fokalen Unternehmens eingenommen, welches eine Maschine mit bestimmter Kapazität und heterogene Kunden hat. Im Fokus steht die Erlösmaximierung. Es wird gezeigt, dass das resultierende Problem stets NP-hart ist, selbst wenn der Preis fixiert wird. Zudem wird eine Heuristik entwickelt, die eine nahezu optimale Lösung finden kann und in polynomieller Zeit zur Lösung kommt. Zusammenfassend wird festgestellt, dass Probleme mit bis zu zehn Bestellungen mit der Heuristik in vertretbarer Rechenzeit gelöst werden können. (Jin 2012)

Hafizoğlu, Gel & Keskinocak (2016) betrachten eine MTO-Fertigung, welche sowohl Termingeschäfte über Rahmenverträge als auch Kassageschäfte<sup>2</sup> (engl. spot deal) bedient. Die Rahmenverträge stellen eine Grundlast für das Unternehmen dar. Sie sind durch kurze Lieferzeiten und geringe Preise geprägt. Das Kassageschäft ist zur Auslastung von Restkapazitäten gedacht. Die Autoren entwickeln einen Markow-Entscheidungsprozess zur Dynamisierung von Preis und Lieferzeit für das Kassageschäft. Sie nehmen hierbei einen unendlichen Zeithorizont an und haben zum Ziel, den langfristigen durchschnittlichen Gewinn zu maximieren. Kunden können im Modell die Angebote annehmen oder ablehnen. Die Angebote selbst werden auf Grundlage des Zustands im Produktionssystem bestimmt. Zur Lösung des Modells wird eine Heuristik entwickelt. Die Autoren finden heraus, dass eine signifikante Gewinnsteigerung möglich ist, sofern die Kassakunden nicht extrem preis- oder lieferzeitsensitiv sind. Die Autoren merken an, dass das Ergebnis des Modells sehr stark von den angenommenen Eigenschaften der Kunden abhängt. (Hafizoğlu, Gel & Keskinocak 2016)

Liu, Lu & Qi (2020) motivieren ihre Arbeit über die zunehmende Notwendigkeit, Preisgestaltung, schwankende Liefertermine und Ablaufplanung in der Produktion integriert anzugehen. Hierbei gibt es durch die zunehmende Digitalisierung immer bessere Möglichkeiten, obgleich die Problemstellung eine rechentechnische Herausforderung bleibt. Die Autoren modellieren einen Hersteller mit einer Produktionsressource, bei dem zufällige Anfragen eingehen und eine statische Lieferzeit vorgegeben ist. Das Ziel ist es,

---

<sup>2</sup> Kassageschäfte bezeichnen Geschäfte, bei denen die Schließung und die Erfüllung eines Kaufvertrags im selben Zeitpunkt oder innerhalb weniger Tage erfolgen.

unter Unsicherheit einen Preis festzulegen, sodass es zu keinen Verspätungen kommt und der Gewinn maximiert wird. Zur Vereinfachung diskretisieren die Autoren das Problem in feste Preis-Wahrscheinlichkeiten-Tupel und eine finite Zahl an Bestellanfrage-wahrscheinlichkeiten. Es wird angenommen, dass die angefragte Lieferzeit immer länger als die Bearbeitungsdauer ist und die Produktionssteuerung sich nach der kürzesten Bearbeitungszeit richtet. Für dieses Problem entwickeln die Autoren eine Lösungs-heuristik. Nach eigener Aussage sollte das Modell für die Praxis weiter detailliert werden, zum Beispiel indem mehr als eine Produktionsressource betrachtet wird. (Liu, Lu & Qi 2020)

Baykasoğlu & Subulan et al. (2020) stellen ein dynamisches Verhandlungsmodell vor, das Lieferzeit-, Preis- und Kapazitätsallokationsentscheidungen über eine ereignisdiskrete Simulation simultan berücksichtigt. Die Autoren betrachten hierbei unterschiedliche Kundentypen hinsichtlich ihres Risikoverhaltens in Bezug auf den Servicelevel des Herstellers. Sie beschränken sich auf die Betrachtung einer einzelnen Engpassressource. Der Ansatz wird im Anwendungsfall mit FCFS und Available-To-Promise (ATP) verglichen. Bei letzterem handelt es sich um eine Herangehensweise zur Schätzung, ob ein Bedarf durch vorhandene Lager- und Produktionskapazitäten noch rechtzeitig gedeckt werden kann. ATP kann daher bei den fristbasierten Methoden eingesortiert werden. In der Arbeit kann gezeigt werden, dass der Umsatz mit dem vorgestellten Lösungsmodell signifikant erhöht werden kann. (Baykasoğlu & Subulan et al. 2020)

Eine weitere Arbeit im Bereich der heuristischen Verfahren existiert von ElHafsi (2000), in der jedoch der Fokus auf der kostenoptimalen Einplanung eines Auftrags liegt und die Preisgestaltung nur nebensächlich betrachtet wird.

### **3.2.5 Sonstige Ansätze zur dynamischen Lieferzeit-Preisgestaltung**

Akan, Ata & Olsen (2012) erarbeiten eine optimale Politik für dynamische Lieferzeit-Preis-Angebote. Ziel ist es, die Wohlfahrt zu maximieren, indem möglichst viele Kunden bedient und gleichzeitig die Lieferkosten begrenzt werden. Die Autoren betrachten hierbei heterogene Kunden, die unterschiedliche Präferenzen haben. Die Politik leiten die Autoren aus einem fluiden Modell ab. Das fluide Modell ist eine grobe Approximation, bei der ein diskretes in ein kontinuierliches Modell überführt wird. Die auftretenden, diskreten Warteschlangen werden hierbei als ein Fluid modelliert, welches kontinuierlich fließt. Das fluide Modell lässt sich dann analytisch lösen. (Akan, Ata & Olsen 2012)

Zhao, Stecke & Prasad (2012) stellen in der MTS-Fertigung ein einheitliches Lieferzeit-Preis-Angebot einem differenzierten Lieferzeit-Preis-Angebot gegenüber. Zur Untersuchung, welches Angebot für das fertigende Unternehmen das bessere wäre, modellieren sie preissensitive und lieferzeitsensitive Kunden. Im Rahmen des Modells allokalieren die Autoren die Kapazitäten jedoch strikt einem jeweiligen Lieferzeit-Preis-Modus. In der Folge gibt es eine Kapazität für kurze Lieferzeit-Angebote und eine Kapazität für niedrige Kosten-Angebote. Sie simulieren numerisch die Verwendung eines statischen Lieferzeit-Preispaars und eines statischen Lieferzeit-Preismenüs. Sie betrachten unter anderem den Einfluss von Kapazitätskosten, Servicegrad und Verspätungskosten. Im Ergebnis gibt es keine dominante Lösung im Rahmen der Annahmen. Stattdessen leiten die Autoren vier Theoreme ab, mit welchen eine situationsspezifische dominante Lösung bestimmt werden soll. Als Nachteil eines differenzierten Angebots im beschriebenen Modell sehen die Autoren höhere Kapazitätskosten. (Zhao, Stecke & Prasad 2012)

Li, Lin & Ye (2014) betrachten die Bestimmung von Preis und Lieferzeit unter Berücksichtigung risikoaverser Entscheider im fokalen Unternehmen. Sie analysieren hierbei numerisch den Zusammenhang zwischen Risikoaversion, Lieferzeit-Preiselastizität, Lagerhaltungskosten und Kapazitäten hinsichtlich der Sensitivität. Die Nachfrage wird in Abhängigkeit des Preises modelliert. Als ein Ergebnis finden die Autoren heraus, dass die versprochenen Lieferzeiten mit zunehmender Risikoaversion länger werden, um Verspätungsstrafen zu vermeiden. Für die Zukunft schlagen die Autoren vor, das Modell um Informationsasymmetrie zu erweitern. (Li, Lin & Ye 2014)

Ye, Sun & Li (2016) nutzen Spieltheorie zur Analyse der optimalen Machtverteilung zwischen Produktion und Marketing-Abteilung eines Unternehmens. Für nahezu identische Machverteilungen nutzen sie ein Nash-Spiel. Für große Machtasymmetrien modellieren sie ein Stackelberg-Spiel. Für die verschiedenen Spiele wird jeweils das Gleichgewicht hergeleitet und die Konsequenzen untersucht. Die Ergebnisse zeigen, dass höhere Gewinne erzielt werden können, wenn die Marketingabteilung mehr Macht gegenüber der Produktion hat, sofern es eine große Anzahl an Kunden gibt. Insgesamt sensibilisiert die Arbeit, wie wichtig eine gute Integration des Auftragsmanagements ist, um optimale Ergebnisse zu erzielen. (Ye, Sun & Li 2016)

Weitere Arbeiten befassen sich mit Einnahmeverteilungsmechanismen (Xiao, Yang & Shen 2011), mit der dynamischen Kapazitätsplanung (Kádár & Egri et al. 2018), der Zuordnung von Lieferzeit-Preis-Paketen mittels eines Diffusionsmodells (Plambeck

2004) und der Auftragsannahme abgebildet als Knapsack-Modell (Spengler, Rehkopf & Volling 2007). Die Ansätze sind bei der Beantwortung der Forschungsfrage von untergeordneter Relevanz und werden daher nicht weiter betrachtet.

### **3.3 Ableitung des Forschungsdefizits**

Im nächsten Schritt liegt der Schwerpunkt auf der Bewertung der Vielzahl an relevanten Ansätzen im Kontext der dynamischen Lieferzeit-Preisgestaltung für die variantenreiche Produktion. Hierzu findet sich in Tabelle 3-2 eine Übersicht der Arbeiten mit ihrer jeweiligen Erfüllung der Anforderungen. Anhand der Übersicht zeigt sich, dass keiner der vorgestellten Lösungsansätze alle Anforderungen erfüllen kann. Diese Schlussfolgerung wird nachfolgend diskutiert und begründet. Hierbei werden zuerst die Besonderheiten der jeweiligen Methodengruppe kurz hervorgehoben und dann die größten Defizite herausgearbeitet.

Lösungsansätze aus dem Bereich des Reinforcement Learning zeichnen sich durch eine hohe Adaptivität und Skalierbarkeit aus. Die Skalierbarkeit lässt sich auf die effizientere Handhabung von Komplexität gegenüber exakten Verfahren begründen. Entsprechend sind sie prinzipiell in der Lage, mit einem hohen Maß an Detail umzugehen, wie sich beispielsweise an der hohen Ausprägung im Bereich der Kundenmodellierung zeigt (vgl. Tabelle 3-2). Gleichzeitig sind die Ansätze nur bedingt in der Lage Informationsasymmetrie abzubauen. Stattdessen werden vorgefertigte Menüs oder Take-It-Or-Leave-It-Angebote betrachtet. Außerdem erfüllen die meisten Ansätze die Anforderung nach einer fristbasierten Ablaufplanung nicht. Beispielsweise verwenden mehr als die Hälfte der Lösungsansätze zur Steuerung der Produktion im Modell eine FCFS-Regel. Die FCFS-Steuerung nimmt dem Agenten das Potenzial, lukrative Aufträge einzuschließen, und kann zu vermeidbaren Verspätungen führen. Darüber hinaus fokussieren fast alle betrachteten Lösungsansätze Verspätungen und vernachlässigen die negativen Folgen einer verfrühten Fertigung. Nur drei Lösungsansätze sind in der variantenreichen Produktion anwendbar. Die Mehrheit der Lösungsansätze reduziert die Produktion auf ein Produkt, eine Station oder beides. Kein Ansatz untersucht die Stabilität der Lösung ausreichend. Fünf der sechs vorgestellten Ansätze aus dieser Gruppe fokussieren den Unternehmensgewinn als zentrale Kennzahl.

Lösungsansätze mit analytischen Lösungen zeichnen sich durch das Finden von Regeln durch Aufstellen und Umformen von Gleichungen aus. Diese sind zumeist recheneffizient und bieten unter den gegebenen Annahmen optimale Lösungen. Jedoch sind

die Lösungsansätze dadurch geprägt, dass sie unter stark einschränkenden, spezifischen Annahmen entstehen und daher keine gute Generalisierbarkeit bieten. Dies zeigt sich an den niedrigen Ausprägungen im Bereich Skalierbarkeit und Adaptivität. Sie sind nicht in der Lage, die Produktion geeignet abzubilden und dementsprechend modelliert kein Lösungsansatz komplexe Produktportfolios, Maschinen, beidseitige Liefertermin-treue und fristbasierte Ablaufplanungen vollständig. Alle Ansätze dieser Gruppe fokussieren den Gewinn oder den Erlös als zentrale Kennzahl.

Lösungsansätze aus dem Bereich der Entscheidungsbaumverfahren zeichnen sich durch das Finden optimaler Lösungen aus. Von den betrachteten Lösungsansätzen formulieren sieben den Gewinn als Optimierungsziel. Hervorzuheben ist, dass die Rechenzeit zur Findung optimaler Lösungen mit der Modellkomplexität exponentiell steigt. Dieser Nachteil wird auch von mehreren Autoren explizit nachgewiesen und hervorgehoben (Esfandiari, Mashreghi & Emami 2019; Jin 2012; Wang & Wang 2019). Die exponentielle Rechenzeit hängt damit zusammen, dass das zugrundeliegende Optimierungsproblem als NP-schwer zu klassifizieren ist. Hieraus folgen die niedrigen Ausprägungen bei der Skalierbarkeit und der Berücksichtigung von Informationsasymmetrie, welche die Komplexität des Modells steigern würden. Mehrere Autoren empfehlen, ihre Modelle mit mehr Details zu erweitern, um der industriellen Praxis näherzukommen. Entsprechend haben die Lösungsansätze niedrige Ausprägungen in den Anforderungen zu mehreren Produkten, mehreren Produktionsressourcen und mehreren heterogenen Kunden. Um dennoch zu Lösungen in annehmbarer Rechenzeit zu gelangen, wird mehrfach der Einsatz von Heuristiken motiviert. Zuletzt ist ein Nachteil der Entscheidungsbaumverfahren, dass sie wegen ihres exakt aufzustellenden Modells wenig adaptiv sind und daher mit Veränderungen nicht ohne manuelle Anpassung umgehen können (Juan & Keenan et al. 2021, S. 2; Domschke & Drexl et al. 2015, S. 134; Arulkumaran & Deisenroth et al. 2017, S. 6). Folglich erscheinen Lösungsansätze aus dem Bereich der Entscheidungsbaumverfahren nicht geeignet für eine dynamische Lieferzeit-Preisgestaltung.

Lösungsansätze mit Heuristiken sind verglichen mit den Entscheidungsbaumverfahren recheneffizienter und verzichten dafür auf die Garantie einer optimalen Lösung. Bei Heuristiken besteht eine Korrelation zwischen Generalität und Rechenzeit. Je problem-spezifischer die Heuristik entwickelt wurde, desto schneller kann sie ein Problem lösen. Dies zeigt sich exemplarisch an der Arbeit von Jin (2012). Durch geschickte Umformulierung wurde im Rahmen der Arbeit versucht, eine gute Kompromisslösung zu finden.

Jedoch stellt Jin (2012) fest, dass mit der Heuristik dennoch nur begrenzte Bestellumfänge gelöst werden können. Auffällig ist die niedrige Ausprägung bei der Berücksichtigung der Informationsasymmetrie, welche bei diesen Verfahren berücksichtigt werden könnte. Zuletzt sei noch angemerkt, dass drei von fünf Lösungsansätzen den Gewinn als Optimierungsziel formulieren. Basierend auf dieser Analyse ergibt sich das Gesamtergebnis, dass in dieser Gruppe kein Ansatz alle Anforderungen erfüllt.

Die verbleibenden Lösungsansätze lassen sich keiner der vorhergehenden Methodengruppen zuordnen. Sie haben jedoch gemeinsam, dass sie sich zumeist auf Nischenmodelle oder Teilaspekte fokussieren, die alleine nicht zur allgemeinen dynamischen Lieferzeit-Preisgestaltung in der variantenreichen Produktion anwendbar sind. Die Ergebnisse weisen zum Teil einen qualitativen Charakter auf und können nicht ohne umfangreiche Erweiterungen für die dynamische Lieferzeit-Preisgestaltung verwendet werden. Sie bieten aber grundlegende Erkenntnisse zu den Zusammenhängen einer Lieferzeit-Preisgestaltung an. Insgesamt kann somit kein Ansatz alle Anforderungen erfüllen.

Es lässt sich zusammenfassen, dass die bestehenden Lösungsansätze ein Defizit aufweisen, welches im Rahmen dieser Arbeit aufgelöst werden soll. Gleichzeitig zeigt sich, dass Reinforcement Learning im Bereich der dynamischen Lieferzeit-Preisgestaltung ein potentes Werkzeug ist. Es ist in der Lage, mit komplexen und großen Problemräumen umzugehen und diese zu adaptieren. Entsprechend haben Ansätze mit dieser Methode bei den Anforderungen den höchsten Erfüllungsgrad (vgl. Tabelle 3-2). Gleichzeitig finden sich hier Ansätze, die nicht auf die variantenreiche Produktion spezialisiert sind.

Tabelle 3-2: Überblick zum Forschungsdefizit

Anforderung:	I Dyn. Lieferzeit- Preisgestaltung in variantenreicher Produktion	II Mehrere Produkte und Produktionsressourcen	III Fristbasierte Ablaufplanung	IV Berücksichtigung individu- eller, heterogener Kunden	V Beidseitige Liefertermintreue	VI Berücksichtigung von Informationsasymmetrie	VII Experimentierbarkeit und Bewertbarkeit	VIII Untersuchung der Stabilität	IX Skalierbarkeit und Adaptivität der Methodik
<b>Ansätze mit Reinforcement Learning</b>									
Vengerov (2008)	○	○	●	●	○	○	●	○	●
Arredondo & Martinez (2010)	●	●	○	○	○	○	●	○	●
Li, Wang & Sawhney (2012)	●	●	●	●	●	●	●	○	●
Hao, Yu & Wu (2013)	●	○	○	○	●	●	●	○	●
Lu, Hong & Zhang (2018)	○	○	○	●	○	●	●	●	●
Bahrami, Chen & Wong (2021)	○	○	○	●	○	●	●	○	●
<b>Ansätze mit analytischen Lösungen</b>									
Ahn, Gümüş & Kaminsky (2007)	○	○	○	○	○	○	●	○	○
Feng, Liu & Liu (2011)	●	○	○	○	●	●	○	○	●
Afèche (2013)	●	○	●	●	○	○	●	●	●
Ata & Olsen (2013)	●	○	●	●	●	●	○	●	○
Yu, Zhao & Sun (2013)	●	○	○	●	○	○	●	●	○
Afèche & Pavlin (2016)	●	○	●	●	○	●	○	○	●
Albana, Frein & Hammami (2018)	●	○	○	○	○	○	●	●	●
<b>Ansätze mit Entscheidungsbaumverfahren</b>									
Chaharsooghi, Honarvar & Modarres (2011)	●	○	●	●	●	○	●	●	●
Hao & Yu (2012)	○	●	○	○	○	○	●	○	●
Jin, Li & Sivakumar (2013)	●	●	○	○	●	○	●	●	○
Guhlich, Fleischmann & Stölzle (2015)	●	○	●	●	●	○	●	●	●
Ouazene & Yalaoui et al. (2017)	●	●	○	○	○	●	●	●	○
Garmdare, Lofti & Honarvar (2018)	●	●	●	●	●	●	●	●	●
Öner-Közen & Minner (2018)	●	○	●	●	●	●	●	●	●
Esfandiar, Mashreghi & Emami (2019)	●	●	●	○	●	○	●	●	○
Wang & Wang (2019)	●	●	○	○	○	○	●	○	○
<b>Ansätze mit heuristischen Verfahren</b>									
Celik & Maglaras (2008)	●	○	●	●	○	●	●	●	●
Jin (2012)	●	○	●	●	●	○	●	○	●
Hafizoğlu, Gel & Keskinocak (2016)	●	○	○	●	●	●	○	○	●
Liu, Lu & Qi (2020)	●	●	●	●	●	●	●	●	●
Baykasoglu & Subulan et al. (2020)	●	●	●	●	●	●	●	●	●
<b>Sonstige Ansätze</b>									
Akan, Ata & Olsen (2012)	●	○	●	●	○	●	○	○	○
Zhao, Stecke & Prasad (2012)	●	○	○	●	○	●	●	●	●
Li, Lin & Ye (2014)	●	○	○	○	●	○	●	●	●
Ye, Sun & Li (2016)	●	○	○	○	●	●	●	●	○

## 4 Überblick des Lösungsansatzes

Ausgehend von der Motivation, der Forschungsfrage und den Forschungsdefiziten wird in diesem Kapitel der Lösungsansatz vorgestellt. In der Motivation wurde dargelegt, inwiefern eine dynamische Lieferzeit-Preisgestaltung trotz steigender Volatilität und unter Wahrung der Kundenzufriedenheit ein Baustein zur Steigerung der Profitabilität eines einzelnen Unternehmens darstellen kann. Um eine dynamische Lieferzeit-Preisgestaltung zu realisieren, muss zum einen die Netzwerkebene betrachtet werden, auf welcher Unternehmen interagieren. Zum anderen ist die Produktion am einzelnen Standort eines Unternehmens relevant. Das Auftragsmanagement bildet eine Brücke zwischen der Produktion am Standort und dem Netzwerk, insbesondere bei der Interaktion mit den Kunden. In der Angebotserstellung stellt es die Verbindung zwischen beiden her (Schuh, Brandenburg & Cuber 2012, S. 60f). Im Rahmen des Lösungsansatzes wird die Perspektive eines Herstellers eingenommen, der unter Informationsasymmetrie mit seinen Kunden interagiert. Es muss hierbei davon ausgegangen werden, dass sich den Kunden des Herstellers durch Wettbewerber Opportunitäten bieten, die der Hersteller nicht genau kennen kann. Die Zuliefererseite wird nicht explizit betrachtet, sondern über den Zustand der Produktion am Standort des Herstellers indirekt berücksichtigt.

Der nachfolgend präsentierte Lösungsansatz soll nun das gezeigte Defizit im Stand der Forschung auflösen und eine entsprechende dynamische Lieferzeit-Preisgestaltung im Kontext variantenreicher Produktion ermöglichen. Der Stand der Forschung unterstützt die These aus Kapitel 2.5, dass das Reinforcement Learning die passende Methodik zur Umsetzung der Lieferzeit-Preisgestaltung sein könnte. Auf dieser These fußt die zentrale Idee des Lösungsansatzes, einen Reinforcement Learning Agenten begrenzt variierbare Parameter einer herzuleitenden Lieferzeit-Preis-Funktion bestimmen zu lassen, aus welcher ein Kunde dann sein individuelles Lieferzeit-Preis-Paar wählen kann. Im Einzelnen setzt sich der Lösungsansatz aus fünf Kapiteln zusammen, die nachfolgend erläutert werden.

Das erste Kapitel (Kapitel 5.1) widmet sich zunächst dem Aufstellen eines zentralen Zielsystems. Das Zielsystem dient zum einen als Grundlage für die Belohnungsfunktion des Reinforcement Learning Agenten und erlaubt es zum anderen, die Zielerreichung des Ansatzes zu quantifizieren. Danach wird das Lösungsmodell konzeptioniert, welches auf das allgemeine Reinforcement Learning Modell aufbaut. Das Reinforcement Learning Modell besteht aus der Umwelt, dem Agenten sowie der Interaktion zwischen

beiden. Analog dazu umfassen zwei weitere Kapitel das Umwelt- bzw. das Agentenmodul.

Das Umweltmodul wird in Kapitel 5.2 behandelt. Es beschreibt die Struktur- und Ablauflogik der Umwelt, mit der der Agent interagiert. Die Umwelt wird in eine Netzwerkebene und eine Standortebene geteilt. In der Netzwerkebene werden sodann Modellkomponenten des Netzwerks, wie beispielsweise Kunden, abgebildet. Äquivalent dazu werden auf der Standortebene alle relevanten Modellkomponenten eines Produktionsstandorts dargestellt. Hier wäre eine Produktionsressource als Beispiel zu nennen.

Das Kapitel 5.3 widmet sich dem Agentenmodul. In diesem Modul wird als Erstes die Interaktion zwischen Agent und Umwelt bestehend aus beobachtbaren Zustandsparametern, Belohnung und Aktion, dargelegt. Als Zweites wird auf den Agenten selbst eingegangen. Dieser basiert auf mehreren neuronalen Netzen, für welche die Architektur und der Trainingsprozess zu bestimmen sind. Die Architektur des Agenten wird ausgehend von der zuvor definierten Interaktion mit der Umwelt abgeleitet. Danach wird erläutert, nach welcher Logik der Trainingsprozess abläuft, um die neuronalen Netze zur Bewältigung ihrer Aufgabe zu befähigen.

Das Kapitel 5.4 behandelt das Versuchsmodul. Das Versuchsmodul hat den Zweck, die Durchführung von Simulationsexperimenten zu ermöglichen, um das Potenzial des Ansatzes nachzuweisen. Das Versuchsmodul überführt hierzu das Umwelt- und Agentenmodul in ein ausführbares und experimentierbares Modell. Diese Einbettung wird als Erstes erläutert. Der Rest des Kapitels beschäftigt sich damit, wie der Potenzialnachweis für einen Anwendungsfall geführt werden kann. Im Rahmen dessen wird eine Hyperparameteroptimierung zur anwendungsfallspezifischen Parametrisierung des Agenten motiviert und entwickelt. Des Weiteren werden ein Vergleichssystem (Status Quo) und verschiedene Szenarien hergeleitet. Der Ansatz muss sich in den Szenarien gegen das Vergleichssystem behaupten, um das Potenzial dieses Ansatzes im Speziellen und der dynamischen Lieferzeit-Preisgestaltung im Allgemeinen nachzuweisen. Durch die vorherige Aufstellung des Zielsystems in Kapitel 5.1 wird es möglich, den Potenzialnachweis quantitativ anhand der Zielgrößen zu führen. Hierbei kommen Hypothesentests zum Einsatz, um eine statistische Absicherung zu schaffen. Die hierfür notwendigen Hypothesen werden ebenfalls in diesem Kapitel aufgestellt.

Das Kapitel 5.5 beinhaltet Erläuterungen zur Implementierung der drei Module in eine Software. Die Implementierung erfolgt komplett eigenständig und greift ausschließlich auf quelloffene, frei verfügbare Softwarepakete zurück. Im Zuge des Kapitels werden die Wahl der verschiedenen Softwarepakete und die Umsetzung der Implementierung begründet. Ausgehend von der Implementierung wird der Ablauf zur Nutzung des resultierenden Softwaresystems dargelegt. Zuletzt wird das Thema Verifikation behandelt, um darzulegen, wie die Korrektheit des Softwaresystems und damit der Ergebnisse gewährleistet werden soll.

## 5 Dynamische Lieferzeit-Preisgestaltung für die variantenreiche Produktion

Dieses Kapitel widmet sich der detaillierten Darstellung des Lösungsansatzes. Es ist gemäß den Erläuterungen in Kapitel 4 aufgebaut und unterteilt sich entsprechend in fünf Abschnitte.

### 5.1 Zielsystem und Reinforcement Learning Modell

In diesem Kapitel wird zunächst das Zielsystem des Lösungsmodells erläutert. Im zweiten Abschnitt wird dann, ausgehend vom allgemeinen Reinforcement Learning Modell, das Lösungsmodell konkretisiert.

#### 5.1.1 Erläuterung des Zielsystems

Im ersten Schritt steht das Zielsystem im Fokus. Das Zielsystem soll zwei Funktionen erfüllen. Die erste Funktion des Zielsystems ist die Ermöglichung einer differenzierten, quantitativen Bewertung des Lösungsansatzes. Zu diesem Zweck werden zentrale Kennzahlen aufgestellt, an welchen sich die Leistung des Lösungsansatzes messen lässt. Die zweite Funktion des Zielsystems ist es, dem Reinforcement Learning Agenten die Grundlage zur Bewertung seiner Aktionen zu geben, welche für den Lernprozess essenziell ist. Zwar bieten die Kennzahlen die Möglichkeit, verschiedene Zusammenhänge darzustellen. Sie sind jedoch nicht geeignet, um die Aktion des Agenten direkt zu bewerten und damit eine Lernrichtung vorzugeben – für das Lernen wird eine einzelne Zahl benötigt. Deshalb wird zusätzlich eine Zielfunktion definiert. Die Zielfunktion ermöglicht die Aggregation verschiedener Zielgrößen zu einer einzigen Größe. Da die verschiedenen Zielgrößen im Konflikt zueinander stehen können, bestimmt die Zielfunktion das Werteverhältnis der Zielgrößen.

##### 5.1.1.1 Einführung der Kennzahlen

Wie in der Einleitung der Arbeit motiviert wurde, soll eine dynamische Lieferzeit-Preisgestaltung in der Lage sein, die Profitabilität eines Herstellers zu steigern. Um die Steigerung der Profitabilität zeigen zu können, bietet sich der Gewinn als zentrale Kennzahl an. So fokussieren auch Ansätze im Stand der Forschung (Kapitel 3.2) den Gewinn als zentrale Kennzahl zur Entscheidung über das Potenzial einer Lösung.

## Unternehmensgewinn

Der Unternehmensgewinn aggregiert Ausgaben und Einnahmen zu einer eindimensionalen Zahl, die leicht zu vergleichen ist. In einer realen Gewinnrechnung müssen alle geldwerten Faktoren berücksichtigt werden, um den Gesamtgewinn eines Unternehmens korrekt abzubilden. Dazu müssen teilweise Aspekte eines Unternehmens abgebildet werden, die mit diesem Lösungsansatz nicht korrelieren, wie zum Beispiel Ausgaben für Versicherungen oder Kreditzinsen. Um die Komplexität der Gewinnrechnung zu limitieren, wird deshalb keine absolute, sondern nur eine Teilgewinnrechnung durchgeführt. Diese Teilgewinnrechnung genügt, um eine Aussage über die Änderung der Profitabilität unter Einsatz einer dynamischen Lieferzeit-Preisgestaltung im Vergleich zu einem Status Quo zu erhalten, wenn angenommen wird, dass alle nicht betrachteten gewinnrelevanten Faktoren konstant bleiben. Der Gewinn  $g$  wird nun in Anlehnung an Simon & Fassnacht (2019, S. 1) gemäß Formel 5-1 berechnet. Dabei ist der Gewinn zunächst von den Einnahmen aus dem Verkauf der Produkte abhängig. Die Einnahmen ergeben sich aus dem Preis  $p_{i,v}$  abzüglich der variablen Kostenarten  $k_{var}$  multipliziert mit der Menge  $m_{i,v}$ . Der Index  $i$  referenziert hierbei den einzelnen Kunden und der Index  $v$  den spezifischen Auftrag. Des Weiteren sind die relevanten, fixen Kosten  $k_{fix}$  abzuziehen. Die variablen Kosten  $k_{var}$  lassen sich weiter in auftragsspezifische Kosten  $k_{i,v}$ , Verspätungskosten  $k_{Verspätung,i,v}$  und Verfrühungskosten  $k_{Verfrühung,i,v}$  aufteilen (vgl. Formel 5-2). Der Vorteil dieser Aufteilung liegt darin, dass bei unbekannten spezifischen Kosten und Verkaufspreisen stattdessen mit Deckungsbeiträgen gerechnet werden kann (Thommen & Achleitner et al. 2020, S. 194, 388). Bei den Fixkosten werden lediglich Kapitalkosten  $k_{Kapital}$  und Instandhaltungskosten  $k_{Instand}$  berücksichtigt.

$$g = \sum_{i,v} (p_{i,v} - k_{var}) \cdot m_{i,v} - k_{fix} \quad \text{Formel 5-1}$$

$$g = \sum_{i,v} (p_{i,v} - k_{i,v} - k_{Verspätung,i,v} - k_{Verfrühung,i,v}) \cdot m_{i,v} - (k_{Kapital} + k_{Instand}) \quad \text{Formel 5-2}$$

Die Variablen der Gewinnrechnung werden zusätzlich in Tabelle 5-1 einzeln erläutert.

Tabelle 5-1: Erläuterung zu den Variablen der Gewinnrechnung

Parameter	Erklärung
Verkaufspreis $p_{i,v}$	Gibt den Verkaufspreis einer Einheit für einen Kunden $i$ bei einem Auftrag $v$ mit spezifischer Lieferzeit an.
Spezifische Kosten $k_{i,v}$	Auftragsspezifische Kosten der Auftragsbearbeitung – zum Beispiel Material- und Maschinenkosten.
Verspätungskosten $k_{Verspätung,i,v}$	Auftragsspezifische Vertragsstrafen bei zu später Lieferung.
Verfrühungskosten $k_{Verfrühung,i,v}$	Auftragsspezifische Lagerkosten beim Hersteller bzw. Kunden bei zu früher Herstellung.
Verkaufsmenge $m_{i,v}$	Gibt die Verkaufsmenge für einen spezifischen Kunden bei einem spezifischen Auftrag mit spezifischer Lieferzeit an.
Kapitalkosten $k_{Kapital}$	Kapitalkosten, welche für den Besitz von Stationen/Maschinen anfallen.
Instandhaltungskosten $k_{Instand}$	Kosten für die Instandhaltung von Stationen

Neben dem Unternehmensgewinn spielt auch die Kundenzufriedenheit eine wichtige Rolle. Eine Teilmotivation des Ansatzes ist es, den Kundenutzen besser oder zumindest gleich gut zu erfüllen, indem den individuellen Lieferzeit- und Preis-Präferenzen besser entsprochen werden kann. Da die Sicht des Herstellers eingenommen wird und Informationsasymmetrie herrscht, kann die Kundenzufriedenheit nicht direkt beobachtet werden. Daher kann die Kundenzufriedenheit nur indirekt anhand der Beauftragung und Vertragseinhaltung gemessen werden. Um zu beurteilen, inwiefern Lieferzeit- und Preis-Präferenzen entsprochen werden konnte, bietet sich eine Reihe von Kennzahlen an. Nachfolgend wird eine Auswahl an Kennzahlen einzeln motiviert und erläutert.

### Anzahl entgangener Aufträge

Eine dynamische Lieferzeit-Preisgestaltung kann zu einem höheren Gewinn führen. Jedoch ist es wichtig zu beachten, wie viele Kunden nach einer Anfrage keinen Auftrag erteilen. Diese Nichterteilung nach einer Anfrage wird im Rahmen dieser Arbeit als entgangener Auftrag angesehen. Andere Möglichkeiten, wie zum Beispiel die Ablehnung durch den Hersteller oder mögliche Aufträge von nicht-anfragenden Kunden werden explizit ausgeklammert. Viele entgangene Aufträge können daher auf viele unzufriedene Kunden hindeuten. Die Anzahl entgangener Aufträge errechnet sich über die

Summe aller Binärwerte  $h_{i,v}$ , welche 1 annehmen, wenn ein Kunde  $i$  für einen Auftrag  $v$  zusagt.

$$n_{aa} = \sum_{i,v} (1 - h_{i,v}) \quad \text{Formel 5-3}$$

### Summe der Lieferterminabweichungen

Ein wichtiger Aspekt der dynamischen Lieferzeit-Preisgestaltung ist die pünktliche Lieferung von Aufträgen. Hierzu existiert eine Vielzahl von Zielgrößen. Als erstes wird die Lieferterminabweichung betrachtet. Eine Besonderheit dieser Arbeit ist die Berücksichtigung von verfrühten Aufträgen. Auch die zu frühe Bearbeitung von Aufträgen ist nachteilig, zum Beispiel aufgrund von Lagerkosten, die entweder beim Kunden oder beim Hersteller anfallen. Die Lieferterminabweichung  $t_{la}$  ergibt sich aus der absoluten Summe von Verfrühtungszeit  $t_{verfrüht}$  und Verspätungszeit  $t_{verspätet}$ . Die Verfrühtungs- bzw. die Verspätungszeit werden aus der Differenz von realisierter Lieferzeit  $t_{Liefer,real,i,v}$  und vereinbarter Lieferzeit  $x_{i,v}$  berechnet. Diese beiden Werte sind einem Kunden  $i$  und einem Auftrag  $v$  zugeordnet.

$$\Delta x_{i,v} = t_{Liefer,real,i,v} - x_{i,v} \quad \text{Formel 5-4}$$

$$t_{verfrüht} = \sum_{i,v} |\Delta x_{i,v}| \text{ für } \Delta x_{i,v} < 0 \quad \text{Formel 5-5}$$

$$t_{verspätet} = \sum_{i,v} |\Delta x_{i,v}| \text{ für } \Delta x_{i,v} > 0 \quad \text{Formel 5-6}$$

$$t_{la} = t_{verfrüht} + t_{verspätet} \quad \text{Formel 5-7}$$

### Servicegrad

Eine weitere Kennzahl zur Bewertung der Pünktlichkeit von Lieferungen ist der Servicegrad. Für den Servicegrad gibt es verschiedene Definitionen. Hier wird der Definition gefolgt, nach der die Anzahl pünktlicher Lieferungen im Verhältnis zur Gesamtzahl aller Lieferungen angegeben wird (Arnold & Isermann et al. 2008, S. 244). Zusammen mit der Lieferterminabweichung lässt sich dann beispielsweise eine Aussage darüber treffen, ob einige wenige Aufträge eine hohe Lieferterminabweichung erreichen.

$$sg = \frac{\sum_{i,v} h_{i,v} \text{ für } \Delta x_{i,v} \leq 0}{\sum_{i,v} h_{i,v}} \quad \text{Formel 5-8}$$

## Anzahl Verspätungen

Der Servicegrad allein reicht nicht aus, um die Situation hinsichtlich pünktlicher Lieferungen ausreichend einschätzen zu können. Der Servicegrad ist eine Relativzahl und trifft keine Aussage über die Menge der Lieferungen. Deshalb wird außerdem die Anzahl verspäteter Aufträge ausgewertet.

$$n_{verspätet} = \sum_{i,v} h_{i,v} \text{ für } \Delta x_{i,v} > 0 \quad \text{Formel 5-9}$$

### 5.1.1.2 Aufstellen der Zielfunktion

Wie beschrieben, dient die Zielfunktion der Aggregation aller relevanten, zu bewertenden Faktoren zu einer Kennzahl, die wiederum die Grundlage für die spätere Belohnungsfunktion des Agenten darstellt. Im Zentrum der Zielfunktion steht zunächst der Unternehmensgewinn. Die Gewinnfunktion des Herstellers soll maximiert werden, indem eine individuelle Lieferzeit-Preis-Funktion für einen anfragenden Kunden parametrisiert wird. Dabei ist der Gewinn zunächst von den variablen Kosten und den Einnahmen je Produkteinheit abhängig. Entsprechend ergibt sich der erste Term  $(p_{i,v} - k_{var}) \cdot m_{i,v}$  (vgl. Formel 5-2). Es wird angenommen, dass nicht durchgeführte Aufträge negativ für das Geschäft des Unternehmens sind. Daher sollten entgangene Aufträge negativ in die Zielfunktion eingehen. Die Umrechnung von einem entgangenen Auftrag, welcher mit der binären Variable  $h_{i,v}$  dargestellt werden kann, in die Einheit der Zielfunktion erfolgt mithilfe eines Gewichtungsfaktors  $W_1$ , sodass sich der Term  $(1 - h_{i,v}) \cdot W_1$  ergibt. Auf der anderen Seite wird die Zielfunktion negativ durch fixe Kosten beeinflusst. Zur Vereinfachung werden hier die fixen Kosten aus Formel 5-2 wieder zu einer Variable  $k_{fix}$  zusammengefasst. Weiterhin wird im Rahmen des Auftrages ein fester Liefertermin zwischen Hersteller und Kunde vereinbart. Abweichungen vom Liefertermin  $t_{verfrührt}$  und  $t_{verspätet}$  gehen neben den variablen Kosten zusätzlich mithilfe der Gewichtungsfaktoren  $W_2$  und  $W_3$  negativ in die Gewinnfunktion ein. Hierdurch ist es möglich, Verfrühungen bzw. Verspätungen über den monetären Wert hinaus zu berücksichtigen. Es ergibt sich der Term  $-t_{verfrührt} \cdot W_2 - t_{verspätet} \cdot W_3$ . Hierbei geben  $W_2$  bzw.  $W_3$  an, wie stark die Verfrühungen bzw. Verspätungen eines Auftrages in die Zielfunktion eingehen sollen.

Der Vorteil gegenüber der direkten Nutzung von Verfrühungs- bzw. Verspätungskosten ist die Einstellbarkeit, wodurch nicht-monetäre Einflüsse berücksichtigt werden können. Aus den Überlegungen ergibt sich sodann die gesamte, zu maximierende Zielfunktion wie folgt:

$$\max z = \sum_{i,v} ((p_{i,v} - k_{var}) \cdot m_{i,v} - (1 - h_{i,v}) \cdot W_1) - k_{fix} - t_{verfrüh} \cdot W_2 - t_{verspätet} \cdot W_3 \quad \text{Formel 5-10}$$

### 5.1.2 Reinforcement Learning Lösungsmodell

Nachdem das Zielsystem des Lösungsansatzes dargelegt wurde, gilt es als nächstes das Reinforcement Learning Modell zu einem Lösungsmodell zu konkretisieren. Kernfunktion des Lösungsmodells ist die Umsetzung einer dynamischen Lieferzeit-Preisgestaltung. Das Lösungsmodell soll dies über eine kontinuierliche Lieferzeit-Preis-Funktion realisieren, aus der ein Kunde individuell die Lieferzeit und einen daraus resultierenden Preis wählen kann. Eine zentrale Idee des Ansatzes ist es, diese kontinuierliche Funktion nicht direkt durch den Agenten zu approximieren, sondern ein Funktionsgerüst zu entwerfen und den Agenten die kontinuierlichen Parameter der Funktion bestimmen zu lassen. Durch dieses Vorgehen wird die Komplexität des Agenten drastisch reduziert. Gleichzeitig werden die Stabilität und Kontrollierbarkeit des Verfahrens erhöht. Eine Kontrollierbarkeit des Verfahrens wird durch die Einstellbarkeit von Parametergrenzwerten möglich. Da es sich bei den Parametern einer kontinuierlichen Funktion um kontinuierliche Werte handelt, wird ein Reinforcement Learning Agent benötigt, welcher mit eben solchen kontinuierlichen Werten umgehen kann. Hier bietet sich das Actor Critic Verfahren an, welches die Vorteile von politikbasierten und wertbasierten Verfahren vereint (vgl. Kapitel 2.6.4).

Ausgehend vom allgemeinen Actor Critic Modell kann das Lösungsmodell wie in Abbildung 5-1 abgeleitet werden. Wichtig zu beachten ist, dass im Agenten bereits eine Trainings- und eine Aktionsperspektive unterschieden werden. Entsprechend kann die Trainingsperspektive des Agenten ausgeblendet werden, wenn er fertig trainiert ist.

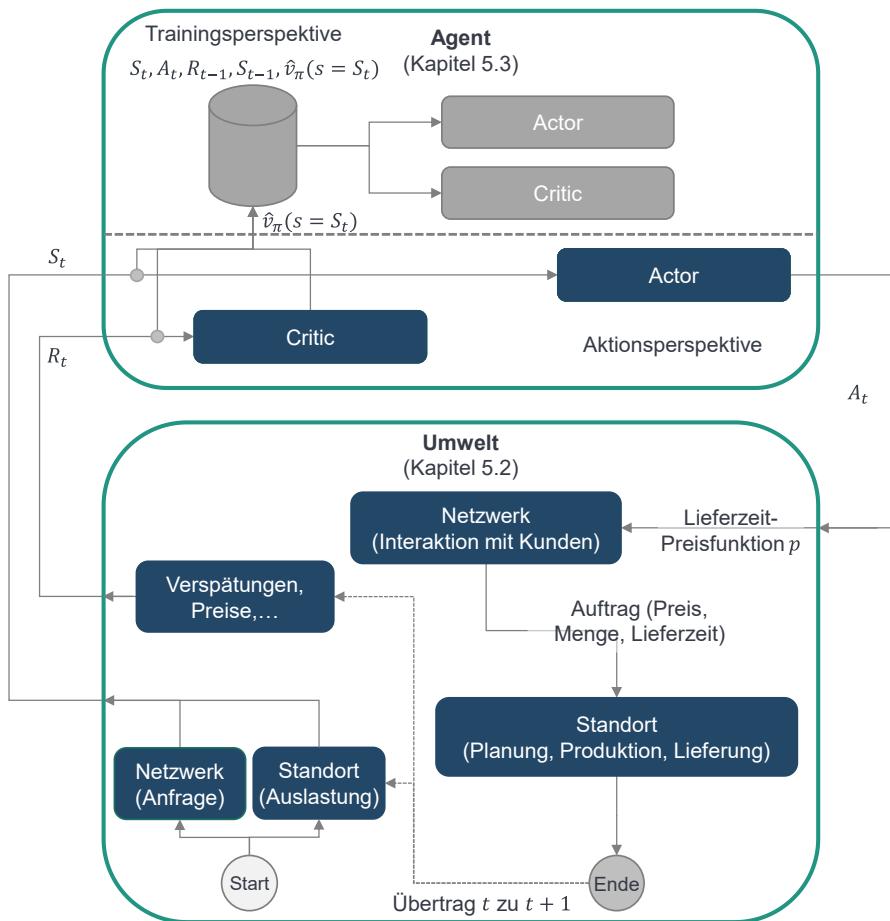


Abbildung 5-1: Überblick zum konkretisierten Lösungsmodell

Das Lösungsmodell ist in der Abbildung nicht als Strukturdiagramm, sondern als Flussdiagramm zu verstehen. Dies bedeutet, dass die Elemente in der Abbildung nach ihrer Aktivitätsreihenfolge dargestellt werden. Entsprechend lässt sich das Lösungsmodell Schritt für Schritt anhand der Darstellung durchgehen. Wichtig für das Verständnis ist, dass der gesamte Ablauf einen Kreislauf bildet. Beginnend unten links in der Darstellung entsteht eine neue Anfrage im Netzwerk. Gleichzeitig wird der aktuelle Zustand

der Produktion am Standort in Form der Auslastung ihrer Stationen erfasst. Beides zusammen bildet den Zustand der Umwelt  $S_t$ . Außerdem wird der Belohnungswert  $R_t$  der letzten Aktion, zum Beispiel auf Basis des zuletzt erzielten Preises und aufgetretener Verspätungen, ermittelt. Der Critic ermittelt nun den approximierten Zustandswert  $\hat{v}_\pi(s)$  auf Basis des Zustandes der Umwelt. Mithilfe des approximierten Zustandswerts, dem gespeicherten letzten Umweltzustand  $S_{t-1}$ , der letzten Aktion  $A_{t-1}$ , der erzielten Belohnung  $R_t$  und dem aktuellen Umweltzustand  $S_t$  können sodann Actor und Critic trainiert werden. Zeitlich unabhängig vom Training wählt der Actor basierend auf dem aktuellen Umweltzustand die nächste Aktion  $A_t$ . Diese besteht aus Funktionsparametern, mit denen die Lieferzeit-Preis-Funktion  $p$  parametrisiert wird. Ausgehend von dieser Lieferzeit-Preis-Funktion interagiert der Kunde nun mithilfe seiner individuellen Nutzenfunktion. Er beauftragt seine Bestellmenge zu einem individuellen Lieferzeit-Preis-Paar oder bestellt nicht.

Im Falle der Bestellung geht der Auftrag in den Standort ein, wird in das Produktionsystem eingeplant und bearbeitet. Es erfolgt ein Übertrag von  $t$  nach  $t + 1$ , weil sich die Umwelt durch den absolvierten Kreislauf nun, zum Beispiel hinsichtlich der Stationsauslastung, verändert. Wie die einzelnen Komponenten des Kreislaufs im Detail funktionieren, ist Inhalt der nachfolgenden Kapitel.

Beim Einsatz in einem realen Unternehmen als Umwelt würde der Reinforcement Learning Agent Daten zum Kunden und zur Produktion erhalten, um die Parameter der Lieferzeit-Preis-Funktion zu bestimmen. Im Rahmen dieser Arbeit wird der Agent jedoch nicht in einem realen Unternehmen eingesetzt, sondern mit einem Simulationsmodell kombiniert. Dies hat zwei Gründe: Erstens bietet sich eine Simulation an, um das reale System eines Unternehmens nicht negativ zu beeinflussen und trotzdem die Hypothesen zu validieren (vgl. Kapitel 2.4.1). Zweitens benötigt der Agent eine gewisse Trainingszeit bis er produktiv einsetzbar ist. Davor sind die Aktionen des Agenten nicht zielführend. Dieses Training kann im Rahmen der Simulation erfolgen. Entsprechend ist das Umweltmodul die Grundlage für ein experimentierbares Simulationsmodell.

## 5.2 Umweltmodul

Der Standort und das Netzwerk stellen im Kontext des Reinforcement Learning Modells die Umwelt dar, welche als experimentierbares Simulationsmodell abgebildet werden soll. Auf der Standortebene werden die physischen und planerischen Prozesse der Produktion einer Vielzahl an Produktvarianten betrachtet. Im Fokus der Arbeit steht die

variantenreiche Produktion, sodass der Produktionsstandort durch eine auftragsbezogene Produktion charakterisiert ist. Auf der Netzwerkebene wird eine beliebige Menge an Kunden explizit modelliert. Konkurrierende Hersteller und Zulieferer gehen in das Modell über den Marktpreis bzw. Kosten indirekt ein. Auf Basis dieser Überlegung kann das Umweltmodul strukturiert nach Standort- und Netzwerkebene in Modellkomponenten aufgeteilt werden.

Die Modellierung der Umwelt basiert auf der vom Verfasser der Arbeit betreuten Abschlussarbeit A\_Ochs (2021).

### 5.2.1 Netzwerkebene

Zur Herleitung der zu modellierenden Modellkomponenten auf der Netzwerkebene ist ein prozessorientiertes Vorgehen erforderlich. Hierbei wird der Prozessablauf Schritt für Schritt durchlaufen, um währenddessen die notwendigen Modellkomponenten zu identifizieren. In der nachfolgenden Beschreibung werden die Modellkomponenten jeweils hervorgehoben und danach in separaten Unterkapiteln beschrieben.

Auf der Netzwerkebene beginnt der Prozess damit, dass einer von mehreren Kunden (Modellkomponente: **Kunde**) in einem Netzwerk (Modellkomponente: **Netzwerk**) eine Nachfrage (Modellkomponente: **Nachfrage**) für ein oder mehrere Produkte (Modellkomponente: **Produkt**) in gewisser Menge hat. Der Hersteller bestimmt für das Produkt eine Lieferzeit-Preis-Funktion  $p_{i,v}(x_{i,v})$  pro Mengeneinheit auf Basis des Zustandes seines Produktionssystems und der Situation im Netzwerk. Der Kunde nutzt nun  $p_{i,v}(x_{i,v})$  in Verbindung mit seiner Nutzenfunktion  $u_i$ , um eine Lieferzeit  $x_{i,v}$  zu einem Preis  $p$  zu bestimmen, indem er seinen erwarteten Nutzen maximiert. Ist der Nutzen des Kunden unter Berücksichtigung der Opportunität oder absolut gesehen negativ, so bestellt er nicht. Im Falle einer Bestellung wird dann ein Kundenauftrag (Modellkomponente: **Kundenauftrag**) generiert, der in das Produktionssystem eingeplant wird.

Die vorhergehende Beschreibung der Netzwerkebene und eine Einordnung in die Umwelt des Lösungsmodells wird in Abbildung 5-2 dargestellt.

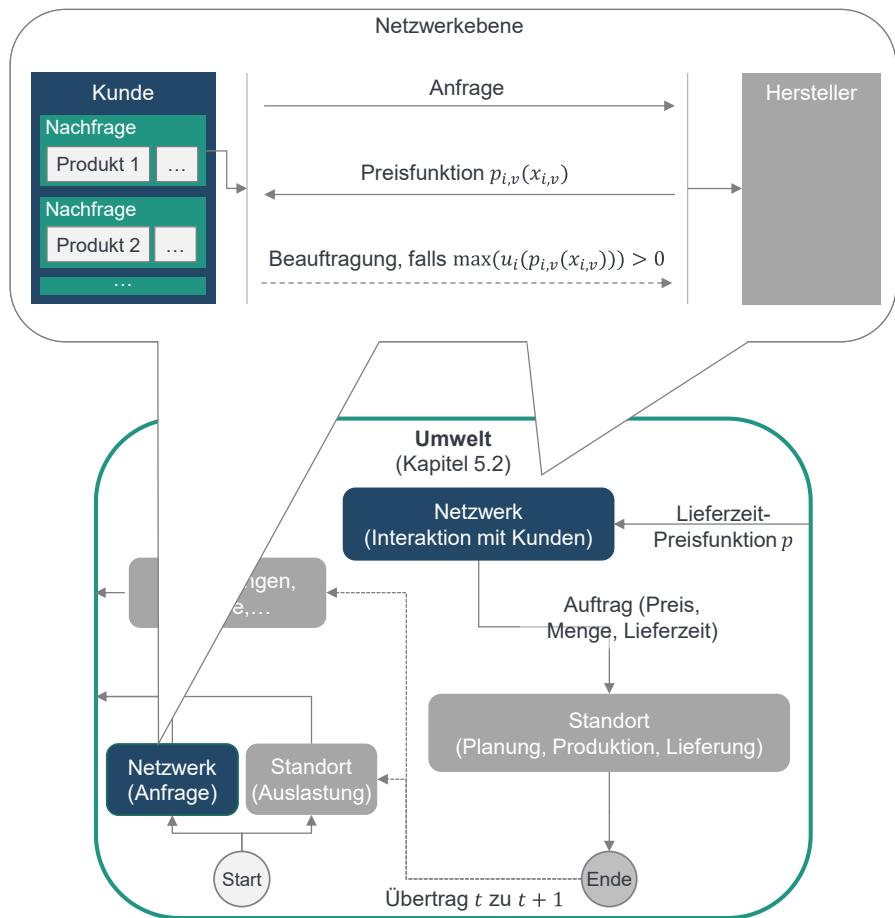


Abbildung 5-2: Überblick über Modellkomponenten und Interaktion auf Netzwerkebene mit Einordnung in die Umwelt des Lösungsmodells

### 5.2.1.1 Modellkomponente: Kunde

Die Kundenkomponente enthält zunächst eine Menge mit Nachfragen  $NF_{i,j}$ . Die Nachfrage ist eine eigene Modellkomponente und wird in Kapitel 5.2.1.2 erläutert. Neben der Produktnachfragemenge ist das Bestellverhalten ein wichtiger Bestandteil des Kunden. Das Bestellverhalten ergibt sich aus der Nutzenfunktion. Es wird angenommen, dass

die Nutzenfunktion bei der Integration der verschiedenen Faktoren einer additiven Verknüpfungsfunktion folgt.

Prinzipiell sind der Preis pro Einheit  $p_{i,v}$ , die auftragsspezifische Bestellmenge  $m_{i,v}$ , die Lieferzeit  $x_{i,v}$  und die Modellzeit  $t$  wichtige Faktoren für das Bestellverhalten. Es wird jedoch angenommen, dass jeder Kunde prinzipiell bereit ist, zum Selbstkostenpreis beim Hersteller zu kaufen, und mindestens mit der minimalen Durchlaufzeit eines Auftrags zufrieden ist. Der Fokus der Nutzenfunktion liegt daher einzig auf den Mehrkosten und der längeren Lieferzeit, die sich relativ zu diesem Minimum ergeben. Die Variablen für Preis und Lieferzeit sind daher immer als Differenz zu diesen beiden Basiswerten zu betrachten. Diese relative Definition der Nutzenfunktion führt später zu einem zentralen Vorteil, da sich hierdurch die Lieferzeit-Preisfunktion über Kunden und Produkte hinweg besser generalisieren lässt. Die Nutzenfunktion gibt weiterhin die Reihenfolge der Alternativen an, wobei alle Werte mit einem Nutzenwert kleiner null dem Bestellverzicht zugeordnet werden. Die Alternative, die ein Kunde wählen wird, ist die mit dem höchsten Nutzenwert. Die absoluten Werte haben neben der ordnenden Funktion keinen Aussagegehalt (vgl. Kapitel 2.3.2). Es wird noch einschränkend angenommen, dass Produktvarianten des Herstellers nicht im Produktpool des Herstellers substituierbar sind. Das bedeutet, der Kunde kann nicht auf ein günstigeres Ersatzprodukt ausweichen, wenn er mit dem Preis im ursprünglichen Angebot unzufrieden ist.

Es wird angenommen, dass alle genannten Faktoren der Nutzenfunktion (Produktpreis, Bestellmenge, ...) linear unabhängig voneinander sind. Zudem sei die Nutzenfunktion stetig und differenzierbar (Kolmar 2017, S. 167-169). Es wird zunächst analog zu Hillig (2006) angenommen, dass die Präferenzen eines Kunden über die Zeit eine limitierte Veränderung aufweisen, sodass mit kleinen Gradienten der Nutzenfunktion  $u_i$  über die Zeit zu rechnen ist. Formell ausgedrückt bedeutet dies:

$$\left| \frac{du_i(p_{i,v}, m_{i,v}, x_{i,v}, t)}{dt} \right| = \varepsilon \quad \text{Formel 5-11}$$

In Formel 5-11 hat  $\varepsilon$  einen sehr kleinen Wert im Verhältnis zur Frequenz der Anfrage. Daher wird  $t$  als Einfluss der Zeit auf die Nutzenfunktion in den nachfolgenden Formeln zunächst vernachlässigt.

Es wird weiter unterstellt, dass die herzuleitende Nutzenfunktion stets genau ein globales Maximum hat, welches die präferierte Bestellpolitik ist. Analog zu den Grundlagen

(vgl. Kapitel 2.3.2) wird eine negative zweite Ableitung im Verhältnis zur Lieferzeit angenommen, was zu einer degressiven Nutzenänderung führt. Es folgt:

$$\frac{d^2 u_i(p_{i,v}, m_{i,v}, x_{i,v})}{d^2 x_{i,v}} < 0 \quad \text{Formel 5-12}$$

Es wird weiter angenommen, dass es verfrühte und verspätete Lieferungen geben kann, die beide suboptimal sind. Um diesem Umstand Rechnung zu tragen, wird die Funktion dergestalt modelliert, dass der Nutzen quasilinear bis zum Hochpunkt steigt, zur Repräsentation von Lagerkosten bei verfrühter Lieferung, und dann steil abfällt, um Verspätungsfolgen abzubilden:

$$\left| \frac{d^2 u_i(p_{i,v}, m_{i,v}, x_{i,v}^+)}{d^2 x_{i,v}} \right| < \left| \frac{d^2 u_i(p_{i,v}, m_{i,v}, x_{i,v}^-)}{d^2 x_{i,v}} \right| < 0 \quad \text{Formel 5-13}$$

Hierbei repräsentiert  $x_{i,v}^-$  alle Werte kleiner und  $x_{i,v}^+$  alle Werte größer als der optimale Lieferzeitpunkt, also dem Hochpunkt der Nutzenfunktion.

Weiterhin wird angenommen, dass die Preisvariable negativ linear in die Nutzenfunktion eingeht und die Bestellmenge den Nutzen einer Bestellung linear skaliert. Für die Entscheidung über eine Bestellung ist die Menge daher unerheblich und kann für die weitere Untersuchung des Bestellverhaltens vernachlässigt werden. Darüber hinaus soll gelten, dass die Nutzenfunktion spezifisch für eine Produktvariante ist. Aus diesen Überlegungen lässt sich die folgende Nutzenfunktion aufstellen:

$$u_i(p_{i,v}, x_{i,v}) = -a_{i,v}e^{(x_{i,v} - b_{i,v})} + c_{i,v}(x_{i,v} - b_{i,v}) + d_{i,v} - p_{i,v} \quad \forall v \in V \quad \text{Formel 5-14}$$

Eine Kurve, die dieser Nutzenfunktion folgt, ist in Abbildung 5-3 gezeigt. Die neu eingeführten, vier freien Parameter  $a_{i,v}$ ,  $b_{i,v}$ ,  $c_{i,v}$  und  $d_{i,v}$  dienen hierbei zur kundenindividuellen Einstellung der Nutzenfunktion. So kann über  $a_{i,v}$  und  $c_{i,v}$  die rechtsseitige bzw. linksseitige Kurvensteilheit beeinflusst werden. Das liegt darin begründet, dass die von  $a_{i,v}$  gesteuerte  $e$ -Funktion für große Werte von  $x_{i,v}$  dominant wird (Bereich: Exponentielle Verspätungskosten). Auf der anderen Seite setzt sich für kleine Werte von  $x_{i,v}$  der lineare Anteil durch (Bereich: Lineare Verfrühungskosten), der über  $c_{i,v}$  beeinflusst wird. Mithilfe von  $a_{i,v}$  und  $c_{i,v}$  kann somit der Lieferzeitbereich eingestellt werden, welcher für den Kunden akzeptabel ist.  $b_{i,v}$  steuert den Kurvenhochpunkt, also den aus Kundensicht optimalen Lieferzeitpunkt in Tagen ab Bestelldatum, und  $d_{i,v}$  die Preisbereitschaft.

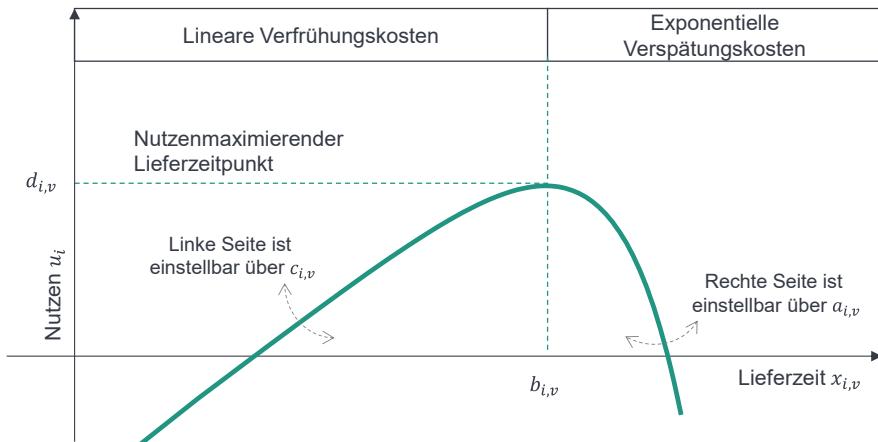


Abbildung 5-3: Zugrundeliegende Nutzenfunktion  $u_i$  des Kunden in Anlehnung an A\_Ochs (2021)

Durch eine Variation dieser vier Parameter können unterschiedliche Marktsituationen abgebildet werden. Stünde der Markt für Massenprodukte im Zentrum der Betrachtung, wären niedrige Preise gefordert. Entsprechend wäre  $d_{i,v}$  vergleichsweise klein. Angenommen eine Industrie mit kurzen Reaktionszeiten stünde im Zentrum der Betrachtung, so könnte das Anforderungsprofil über vergleichsweise kleine  $b_{i,v}$  abgebildet werden. Für eine Just-in-Sequence-Produktion, bei der das Ankommen der Teile im exakt richtigen Zeitpunkt in der richtigen Reihenfolge kritisch ist, ließen sich  $a_{i,v}$  und  $c_{i,v}$  so einstellen, dass nur ein kleines Lieferfenster akzeptiert wird.

Neben dieser Nutzenfunktion gibt es weitere mögliche Lösungen, die die aufgeführten Kriterien erfüllen. Wie sich nachfolgend zeigen wird, wurde die hier vorgestellte Funktion entwickelt, um besonders vorteilhafte Eigenschaften zu erzielen, beispielsweise hinsichtlich Stetigkeit und Differenzierbarkeit.

Für die Implementierung des Bestellverhaltens ist es unter anderem notwendig, die Maximalstellen der Nutzenfunktion zu bestimmen. Die Maximalstelle wird nämlich stets das Lieferzeit-Preis-Paar sein, welches der Kunde wählen wird. Für die Identifikation der Maximalstelle werden mithilfe der Methoden der Analysis die notwendige Bedingung, dass die erste Ableitung gleich null sein muss, und die hinreichende Bedingung,

dass die zweite Ableitung in den gefundenen Lösungen kleiner null sein muss, geprüft. Es gilt:

$$\frac{du_i(p_{i,v}, x_{i,v})}{dx_{i,v}} = u'_i(p_{i,v}, x_{i,v}) = -a_{i,v}e^{(x_{i,v}-b_{i,v})} + c_{i,v} \stackrel{!}{=} 0$$

$$\Rightarrow x_{i,v}^{opt} = \ln(c_{i,v}) - \ln(a_{i,v}) + b_{i,v}$$
Formel 5-15

$$\frac{d^2u_i(p_{i,v}, x_{i,v})}{d^2x_{i,v}} = u''_i(x_{i,v}) = -a_{i,v}e^{(x_{i,v}-b_{i,v})} \stackrel{!}{<} 0$$

mit  $x_{i,v} = x_{i,v}^{opt}$  folgt

$$\frac{c_{i,v}}{a_{i,v}} > 0$$
Formel 5-16

Es folgt, dass  $a_{i,v}$  und  $c_{i,v}$  immer positiv sein müssen, weil andernfalls der natürliche Logarithmus nicht definiert ist. Im Weiteren soll zur Vereinfachung gelten:

$$a_{i,v} := c_{i,v} \quad \text{Formel 5-17}$$

Hierdurch wird die Kurvensteilheit auf beiden Seiten der Maximalstelle allein durch  $a_{i,v}$  eingestellt, sodass diese Variable den akzeptablen Lieferzeitkorridor beeinflusst und  $c_{i,v}$  entfällt. Des Weiteren vereinfacht sich durch diese Definition das Finden der optimalen Lieferzeit zu:

$$x_{i,v}^{opt} = b_{i,v} \quad \text{Formel 5-18}$$

Wie in der Einleitung beschrieben, gilt, dass ein Kunde immer einen positiven Nutzen im Hochpunkt hat, sofern der Preis den Kosten des Herstellers für die Produktion des Produktes entspricht. Der Kunde würde also keine Anfrage stellen, wenn das Produkt unabhängig von der Lieferzeit zum Mindestpreis keinen positiven Nutzen stiften würde. Anders dargestellt bedeutet diese Annahme, dass der Hersteller prinzipiell in der Lage sein muss, zu einem konkurrenzfähigen Preis zu produzieren. Der Preisparameter ist jedoch eine relative Größe zum Selbstkostenpreis und daher nimmt  $p_{i,v}$  in dem Fall den Wert null an. Es folgt:

$$u_i(0, x_{i,v}^{opt}) = -a_{i,v}e^{(0)} + a_{i,v}(0) + d_{i,v} \stackrel{!}{>} 0 \quad \forall v \in V$$
Formel 5-19

Es lässt sich nun ableiten, dass  $d_{i,v}$  immer größer als  $a_{i,v}$  sein muss. Es gelte:

$$d_{i,v} = a_{i,v} + d_{i,v}^u \text{ mit } d_{i,v}^u > 0 \quad \text{Formel 5-20}$$

Durch die Einführung der Substitutionsvariablen  $d_{i,v}^u$  kann die Nutzenfunktion dergestalt umgeformt werden, dass die vorhergehende Bedingung stets erfüllt ist. Hieraus folgt die Nutzenfunktion als:

$$u_i(p_{i,v}, x_{i,v}) = -a_{i,v}e^{(x_{i,v} - b_{i,v})} + a_{i,v}(x_{i,v} - b_{i,v}) + a_{i,v} + d_{i,v}^u - p_{i,v} \quad \forall v \in V \quad \text{Formel 5-21}$$

$d_{i,v}^u$  kann nun als reiner Nutzenbeitrag des Produktes bereinigt vom Basispreis interpretiert werden. Die Gesamtnutzenfunktion ergibt sich nun zusammen mit der Bestellmenge  $m_{i,v}$  wie folgt:

$$m_{i,v}(-a_{i,v}e^{(x_{i,v} - b_{i,v})} + a_{i,v}(x_{i,v} - b_{i,v}) + a_{i,v} + d_{i,v}^u - p_{i,v}) \quad \forall v \in V \quad \text{Formel 5-22}$$

Die so aufgestellte Nutzenfunktion hat jedoch einen Nachteil: Sie bildet die Bereitschaft zur Zahlung von Lieferpreisaufschlägen als absolute Zahl ab. Das bedeutet, dass für ein sehr teures Gut die gleiche Bereitschaft zur Zahlung eines Aufschlags zum statischen Basispreis  $p^*$  wie für ein sehr günstiges Gut besteht. Dies deckt sich nicht mit dem realen Verhalten, bei dem die Preissensibilität abhängig von der Höhe der Kosten ist (vgl. hierzu Kahneman & Tversky (1979)). Deshalb wird die gesamte Nutzenfunktion des Kunden relativ zum Basispreis  $p^*$  definiert:

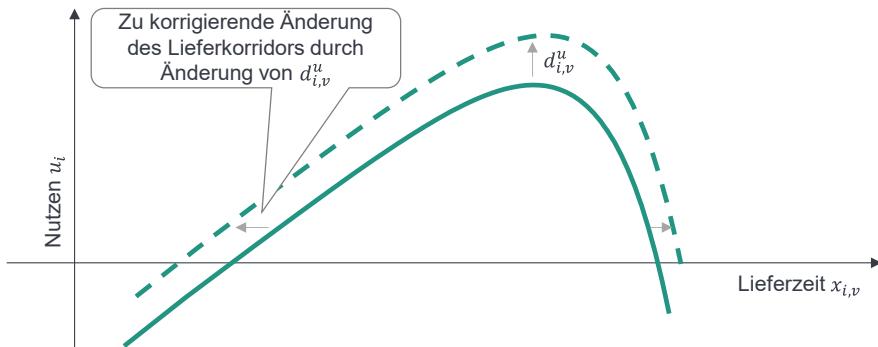
$$u_i^*(p_{i,v}, m_{i,v}, x_{i,v}) = \frac{m_{i,v}(-a_{i,v}e^{(x_{i,v} - b_{i,v})} + a_{i,v}(x_{i,v} - b_{i,v}) + a_{i,v} + d_{i,v}^u - p_{i,v})}{p^*} \quad \forall v \in V \quad \text{Formel 5-23}$$

Anschließend werden die Parameter nun in Relation zum statischen Preis  $p^*$  auf sinnvolle Wertebereiche begrenzt, in welchen die Funktion als Modell der Präferenz valide ist (vgl. Tabelle 5-2).

Tabelle 5-2: Grenzen der Nutzenfunktionsparameter

Parameter	Untere Grenze	Obere Grenze	Begründung
$\frac{a_{i,v}}{p^*}$	0,1	0,9	Begrenzung der Kurvensteilheit auf ein analytisch sinnvolles Maß. Außerhalb dieser Grenzen bildet der Kurvenverlauf des Graphen den Nutzen nicht mehr korrekt ab.
$b_{i,v}$	1	30	Es wird angenommen, dass Kunden nicht weiter als 30 Tage im Voraus bestellen. Eine anwendungsfallspezifische Anpassung ist möglich.
$\frac{d_{i,v}^u}{p^*}$	0	1,5	Es wird angenommen, dass die maximale, relative Zahlungsbereitschaft 150 % beträgt. Anwendungsfallspezifische Anpassungen sind möglich.

Des Weiteren können die Kunden nun klassifiziert werden. Die Steilheit der Kurve, welche über den Parameter  $a_{i,v}$  bestimmt wird, gibt die Lieferzeitsensitivität an. Es können entsprechend lieferzeitsensitive und nicht lieferzeitsensitive Kunden unterschieden werden. Parameter  $b_{i,v}$  bestimmt die Lage des Lieferzeitoptimums. Es können kurzfristig bestellende und langfristig bestellende Kunden unterschieden werden. Zuletzt können Kunden mit hoher und niedriger Preisbereitschaft modelliert werden. Die Preisbereitschaft kann jedoch mit der gegebenen Nutzenfunktion nicht allein durch den Parameter  $d_{i,v}^u$  isoliert eingestellt werden (vgl. Abbildung 5-4). Es wird zusätzlich der Parameter  $a_{i,v}$  benötigt, um die Nullstellen zu erhalten, denn eine Änderung der Preisbereitschaft bedeutet letztendlich, dass allein das Nicht-Bestellen als Alternative in der Rangliste der Möglichkeiten auf- oder absteigt. Der Lieferzeitkorridor sollte sich dadurch nicht ändern. Es können auf diese Art und Weise Kunden mit hoher und niedriger Preisbereitschaft unterschieden werden.

Abbildung 5-4: Effekt einer Änderung von  $d_{i,v}^u$ 

Die sich ergebenden acht Klassen, im Folgenden auch als Kundentypen bezeichnet, sind in Abbildung 5-5 mit den zugehörigen, typischen Funktionsverläufen dargestellt. Die Kundentypen helfen in der Erprobung des Ansatzes bei der Strukturierung und Analyse der Anwendungsfälle.

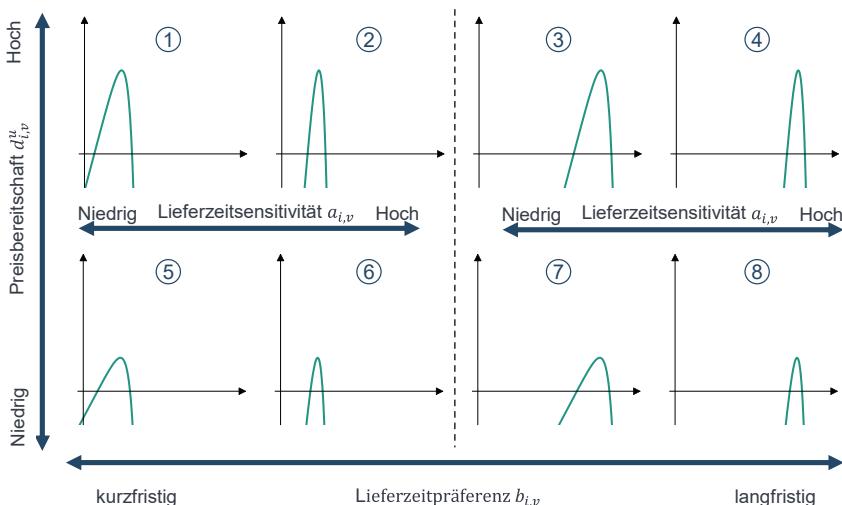


Abbildung 5-5: Darstellung der acht verschiedenen Kundentypen

Als Zusammenfassung der Kundenkomponente zeigt Tabelle 5-3 nochmal alle relevanten Variablen auf.

*Tabelle 5-3: Übersicht zur Kundenkomponente*

Variable	Erklärung
Nachfrage $NF_{i,j}$	s. Kapitel 5.2.1.2.
Nutzenfunktion $u_i^*$	Bildet zusammen mit den kundenspezifischen Parametern die Entscheidungslogik ab.
Steilheitsparameter $a_{i,v}$	Stellt die Flankensteilheit der Nutzenfunktion und damit den akzeptablen Lieferzeitkorridor ein.
Lieferzeitpräferenzparameter $b_{i,v}$	Stellt den Hochpunkt der Nutzenfunktion und damit den präferierten Lieferzeitpunkt ein.
Preisbereitschaftsparameter $d_{i,v}^u$	Stellt die Verschiebung der Funktion auf der y-Achse und damit die Preisbereitschaft ein.

### 5.2.1.2 Modellkomponente: Nachfrage

In der kundenzentrischen Produktnachfrageliste werden Bestellwahrscheinlichkeiten kundenspezifisch bestimmt, um ein Bestellverhalten abzubilden. Es wird die Bestellhäufigkeit  $BH_{i,j}$  und die produktsspezifische Bestellmenge  $BM_{i,j}$  einer Produktvariante  $j$  unterschieden. Die zugrundeliegenden Verteilungsfunktionen von Bestellmenge und -häufigkeit sind anwendungsfallspezifisch zu modellieren. Lassen sich die benötigten Informationen jedoch nicht beschaffen, so können Annahmen getroffen werden. Für die Bestellhäufigkeit wird in dem Fall angenommen, dass diese der Exponentialverteilung folgt (vgl. Formel 5-24). Diese Annahme ist konsistent mit dem Stand der Forschung (Bungartz & Zimmer et al. 2013, S. 37; Banks 2010, S. 209f).

$$F_{exp}(x_{exp}) = \begin{cases} 1 - e^{-\lambda_{exp} \cdot x_{exp}}, & x_{exp} \geq 0 \\ 0, & \text{sonst} \end{cases} \quad \text{Formel 5-24}$$

In der Exponentialverteilung gibt  $\lambda_{exp}$  die erwartete Anzahl der Bestellungen für ein Einheitsintervall an. Der Parameter  $x_{exp}$  bestimmt die obere Intervallgrenze und damit sozusagen die Wartezeit, bis zu der die Wahrscheinlichkeit  $F_{exp}(x_{exp})$  für das Eintreten von mindestens einer Bestellung zu berechnen ist. Ein typischer Kurvenverlauf der Exponentialverteilung ist in Anhang A2 zu finden.

Für die Bestellmenge wird bei mangelnder Informationslage eine Poissonverteilung zugrunde gelegt, was ebenfalls deckungsgleich mit dem Stand der Forschung ist (Banks

2010, S. 229f; Gutenschwager & Rabe et al. 2017, S. 103f). Durch die Poissonverteilung werden im Gegensatz zur klassischen Normalverteilung Werte kleiner 0 vermieden. Die Verteilung ist diskret, wodurch sie ideal zur Repräsentation diskreter Einheiten geeignet ist. Außerdem lässt sich die Poissonverteilung beliebig für andere Zeiteinheiten umrechnen, ohne die Verteilungsmetriken der ursprünglichen Zeitbasis zu verändern. Die Poissonverteilung ist wie folgt definiert:

$$F_{pois}(n_{pois}) = e^{-\lambda_{pois}} \sum_{k_{pois}=0}^{n_{pois}} \frac{\lambda_{pois}^{k_{pois}}}{k_{pois}!} \text{ mit } n_{pois} \in \mathbb{N} \quad \text{Formel 5-25}$$

Aus mathematischer Sicht ermittelt die Poissonverteilung die Wahrscheinlichkeit, dass bei einem Erwartungswert von  $\lambda_{pois}$  eine Bestellung bis zu  $n_{pois}$  Einheiten umfasst. Eine exemplarische Poissonverteilung ist in Anhang A3 zu sehen.

Tabelle 5-4 fasst die Nachfragekomponente nochmals zusammen.

*Tabelle 5-4: Übersicht zur Nachfragekomponente*

Variable	Erklärung
Bestellhäufigkeit $BH_{i,j}$	Modelliert die Häufigkeit einer Bestellung produktvariantenspezifisch pro Kunde. Ohne gegenläufige Informationen wird eine Exponentialverteilung angenommen.
Bestellmenge $BM_{i,j}$	Modelliert die Menge einer Bestellung produktvariantenspezifisch pro Kunde. Ohne gegenläufige Informationen wird eine Poissonverteilung angenommen.

### 5.2.1.3 Modellkomponente: Produkt(-variante)

Die Produktkomponente enthält die spezifischen Produktinformationen. Da die Unterscheidung von Produkt und Variante keinen Mehrwert bietet und daher unnötige Komplexität in das Modell bringen würde, wird nur von einem Produkt gesprochen. Zur Produktkomponente gehören eine eindeutige ID  $j$  und der Fertigungsplan  $F_{plan,j}$ , welcher in Kapitel 5.2.2.3 erläutert wird.

*Tabelle 5-5: Übersicht zur Produktkomponente*

Variable	Erklärung
Produkt $j$	Identifiziert ein Produkt eindeutig
Fertigungsplan $F_{plan,j}$	Referenz zum zugehörigen Fertigungsplan. S. Kapitel 5.2.2.3

### 5.2.1.4 Modellkomponente: Kundenauftrag

Für einen Auftrag wird zwischen einem Kundenauftrag und einem Produktionsauftrag unterschieden. Ein Kundenauftrag  $v$  entsteht auf der Netzwerkebene durch einen Kunden. Der Produktionsauftrag  $pa$  unterscheidet sich vom Kundenauftrag insofern, dass ein Produktionsauftrag stets nur ein spezifisches Produkt umfasst (vgl. Kapitel 5.2.2.3). Sollte in der Praxis also ein Kundenauftrag aus mehreren Produkten bestehen, so wird der Kundenauftrag in mehrere Produktionsaufträge gespalten. Durch die Aufspaltung können die unterschiedlichen Produkte eines Kundenauftrags unabhängig voneinander verarbeitet werden. Im Übrigen sind der Kundenauftrag und der Produktionsauftrag fast identisch. Eine Ausnahme hier ist die Referenz zu allen zugehörigen Produktionsaufträgen  $PA_v$ . Tabelle 5-6 fasst die Kundenauftragskomponenten nochmals zusammen.

*Tabelle 5-6: Übersicht zur Kundenauftragskomponente*

Variable	Erklärung
Auftraggeber $i$	Erlaubt die Zuordnung des Auftrags zum Kunden.
Bestellmengen $BM_{i,j}$	Bestimmt die zu fertigende Bestellmenge pro Produkt.
Produktionsaufträge $PA_v$	Gibt die Menge aller Produktionsaufträge an, die einem Kundenauftrag $v$ zugeordnet sind.
Bestelldatum $e_{datum}$	Zeitpunkt der Bestellung des Auftrags
Vereinbarter Lieferzeitpunkt $x_{i,v}$	Hält fest, bis wann der Auftrag fertiggestellt sein muss und dient als Grundlage der Produktionsplanung und der Ver-spätungsbewertung.
Geplante Lieferzeit $t_{Liefer,plan,i,v}$	Gibt die Planung zum Zeitpunkt der Abfrage wieder. Dieser Wert kann sich während der Ausführung ändern.
Tatsächliche Lieferzeit $t_{Liefer,real,i,v}$	Dieser Wert hält den Ist-Stand fest und erlaubt es, den Planungsfehler zu berechnen.
Vereinbarter Preis $p_{i,v}$	Hält den vereinbarten Preis pro Einheit für den Auftrag fest und erlaubt die Gewinnrechnung für Hersteller bzw. Nutzenrechnung für den Kunden.

### 5.2.1.5 Modellkomponente: Netzwerk

Die Netzwerkkomponente ist der Einstiegspunkt des Umweltmoduls. Sie soll das (externe) Netzwerk aus Herstellersicht repräsentieren. Innerhalb der Netzwerkkomponente kann eine variable Population an Kunden  $I$  mit individuellen Nutzenfunktionen und Produktnachfragelisten existieren. Konkurrenten und Zulieferer werden implizit berücksichtigt. Während der Ausführung des Versuchsmodells dient die Netzwerkkomponente als Plattform zur Auslösung kundenspezifischer Aufträge, zur Koordination der Verhandlung und zur Übergabe des bestätigten Auftrags an das Produktionssystem des Standorts  $S_{standort}$ . Tabelle 5-7 fasst die Variablen der Netzwerkkomponente zusammen.

*Tabelle 5-7: Übersicht zur Netzwerkkomponente*

Variable	Erklärung
Menge aller Kunden $I$	Referenz zur Menge aller Kunden. (vgl. Kapitel 5.2.1.1)
Standort $S_{standort}$	Referenz zum Standort. (Vgl. Kapitel 5.2.2.6)

### 5.2.2 Standortebene

Die relevanten Modellkomponenten auf der Standortebene werden analog zur Netzwerkebene aus dem Prozessablauf identifiziert. Hierzu werden die identifizierten Modellkomponenten im Text wieder hervorgehoben und nachfolgend einzeln erläutert. Einen Überblick gibt Abbildung 5-6.

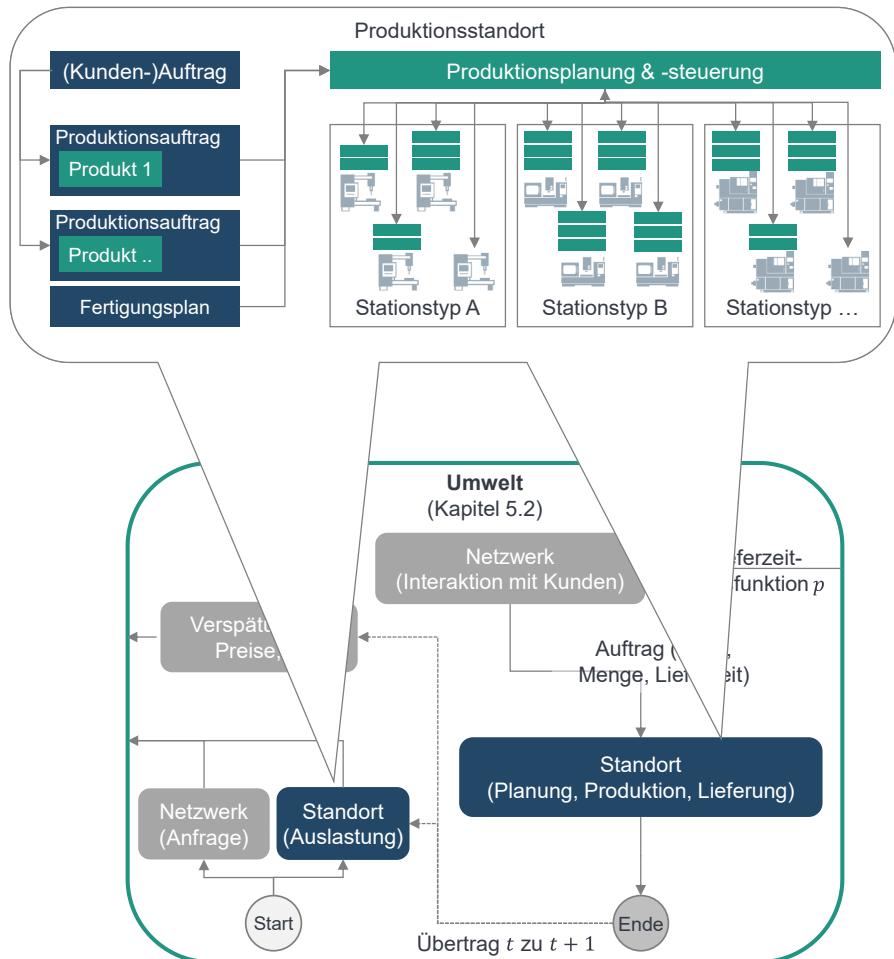


Abbildung 5-6: Übersicht über Modellkomponenten der Standortebene mit Einordnung in die Umwelt des Lösungsmodells

Der Ablauf auf der Standortebene beginnt mit dem Eingang eines angenommenen Auftrags in das Produktionssystem des Standorts (Modellkomponente: **Standort**). Im Auftrag des Kunden steht unter anderem beschrieben, welches Produkt zu welchem Zeitpunkt in welcher Menge hergestellt werden soll. Für eine einfachere Handhabung wird

dieser Kundenauftrag als Erstes in produktsspezifische Produktionsaufträge (Modellkomponente: **Produktionsauftrag**) aufgespalten. Die PPS (Modellkomponente: **Produktionsplanung und -steuerung**) plant die Produktionsaufträge in das Produktionssystem ein. Hierzu wird auf einen produktsspezifischen Fertigungsplan (Modellkomponente: **Fertigungsplan**) zurückgegriffen. Der Fertigungsplan beschreibt, welche Stationen (Modellkomponente: **Station**) bzw. Stationstypen (Modellkomponente: **Stationstyp**) für einen spezifischen Transformationsschritt eines spezifischen Produkts in Frage kommen. Eine Station kann beispielsweise ein Handarbeitsplatz, eine CNC-Maschine oder eine Roboterzelle sein. Die Liste an möglichen Stationen nutzt die PPS, um den Auftrag fristbasiert einzuplanen. Jede Station hat hierzu eine Warteschlange, in der die bisher eingeplanten Aufträge aufgereiht sind. Bei der Einplanung wird angenommen, dass ein Transformationsschritt eines Auftrags immer komplett auf einer Station abläuft. Ein Transformationsschritt eines spezifischen Auftrags kann also umgekehrt nicht anteilig auf einer Station und anteilig auf einer anderen Station durchgeführt werden. Sobald nun ein Auftrag als nächstes in der Warteschlange an der Reihe ist, wird er aus der Warteschlange entfernt und blockiert die Station für die Dauer des Transformationsschritts. Es wird hierbei angenommen, dass stets genug Material bereitsteht, um einen Transformationsschritt durchzuführen.

Nachdem der Transformationsschritt abgeschlossen ist, geht der Auftrag mit neuem Zustand zurück an die PPS zur Einplanung des nächsten Transformationsschritts mithilfe des Fertigungsplans. Hinsichtlich möglichen Ausschusses wird angenommen, dass mögliche Defekte jeweils direkt nach einem ursächlichen Transformationsschritt erkannt werden können und Ausschuss somit direkt ausgesteuert bzw. nachgearbeitet werden kann. Zu jedem Zeitpunkt können Maschinen ausfallen. Es wird hierbei angenommen, dass kein zusätzlicher Ausschuss entsteht, sondern die Maschine lediglich für eine gewisse Zeitdauer nicht nutzbar ist.

### 5.2.2.1 Modellkomponente: Station

Eine Station ist im Rahmen der Modellierung allgemein eine Instanz, die Produktionsfaktoren von einem Input in einen Output transformiert (vgl. Kapitel 2.1.1). Die Station führt hierzu einen Transformationsprozess aus. Für die Modellierung im Rahmen des Umweltmoduls ist vor allem die Zeitdauer, welche für den Transformationsprozess benötigt wird, zentral. Die Zeitdauer wird durch den Stationstyp  $S_{typ}$  bestimmt (vgl. Kapitel 5.2.2.2).

Jede Station hat eine eigene Warteschlange  $W_{schlange,o}$ , in der Zwischenprodukte mit ihren Aufträgen auf dem Weg zum Fertigprodukt auf ihre Transformation warten. Die Reihenfolge dieser Schlange ist von hoher Relevanz und wird über die Modellkomponente PPS bestimmt. Die Station selbst entnimmt stets den nächsten Auftrag der Warteschlange und bearbeitet diesen. Dieser Auftrag wird dann mit  $pa_{aktuell,o}$  bezeichnet. Es wird angenommen, dass die Warteschlangen stets ausreichend Kapazität zur Verfügung haben, um Aufträge einzurichten. Tabelle 5-8 fasst die Stationskomponente zusammen.

*Tabelle 5-8: Übersicht zur Stationskomponente*

Variable	Erklärung
Stationstyp $S_{typ}$	Der Stationstyp der Station.
Warteschlange $W_{schlange,o}$	Die Warteschlange an Aufträgen an einer betrachteten Station $o$ .
Aktueller Auftrag $pa_{aktuell,o}$	Der Produktionsauftrag $pa$ , der aktuell auf einer betrachteten Station $o$ in Bearbeitung ist.

### 5.2.2.2 Modellkomponente: Stationstyp

Ein Stationstyp ist ein Gruppierungselement für eine Menge an Stationen  $O_{styp}$  mit demselben Typ. Im Stationstyp werden Informationen festgehalten, welche allgemein für die Gruppe gültig sind. Diese betreffen den Transformationsprozess. So wird die Fertigungszeit  $fz_{j,q,styp}$  für einen spezifischen Transformationsschritt eines spezifischen Produkts in einem spezifischen Zustand festgehalten. Die Fertigungszeit ist zudem abhängig von der Produkt-Station-Kombination. Sie unterliegt zudem stochastischen Einflüssen und unterliegt somit einer gewissen zufallsbasierten Streuung. Für die Fertigungszeit wird deshalb allgemein eine Dreiecksverteilung angenommen (Gutenschwager & Rabe et al. 2017, S. 109), welche in Formel 5-26 dargestellt ist. Der Name der Verteilung kommt von der Dreiecksform, welche die Dichtefunktion, also die Ableitung der Verteilung, annimmt. Die Dreiecksverteilung hat bestimmte Vorteile, wie zum Beispiel eine klare obere und untere Grenze oder die einfache Erhebung durch Abschätzung (Gutenschwager & Rabe et al. 2017, S. 110), was in der Ablaufsteuerung Vorteile bietet.

$$F_{\Delta}(x_{\Delta}) = \begin{cases} 0, & \text{für } x_{\Delta} < a_{\Delta} \\ \frac{(x_{\Delta} - a_{\Delta})^2}{(b_{\Delta} - a_{\Delta})(c_{\Delta} - a_{\Delta})}, & \text{für } a_{\Delta} \leq x_{\Delta} \leq c_{\Delta} \\ 1 - \frac{(b_{\Delta} - x_{\Delta})^2}{(b_{\Delta} - a_{\Delta})(b_{\Delta} - c_{\Delta})}, & \text{für } c_{\Delta} < x_{\Delta} \leq b_{\Delta} \\ 1, & x_{\Delta} > b_{\Delta} \end{cases} \quad \text{Formel 5-26}$$

Die Dreiecksverteilung gibt die kumulierte Wahrscheinlichkeit  $F_{\Delta}(x_{\Delta})$  an, dass bis zu  $x_{\Delta}$  Minuten Fertigungszeit eintreten. Hierbei ist  $a_{\Delta}$  die niedrigste, mögliche Fertigungszeit.  $c_{\Delta}$  gibt die Fertigungszeit mit der höchsten Wahrscheinlichkeit an.  $b_{\Delta}$  entspricht der längsten, möglichen Fertigungszeit. Die Verteilungsfunktion ist exemplarisch in Anhang A4 dargestellt. Die Variablen des Stationstyps finden sich zusammengefasst in Tabelle 5-9.

*Tabelle 5-9: Übersicht zur Stationstypkomponente*

Variable	Erklärung
Stationsgruppe $O_{Styp}$	Gruppe an Stationen mit demselben Stationstyp
Fertigungszeit $fz_{j,q,Styp}$	Fertigungszeit von Produkt $j$ in Zustand $q$ auf dem Stationstyp $S_{typ}$

### 5.2.2.3 Modellkomponente: Produktionsauftrag

Ein Produktionsauftrag leitet sich immer aus einem Kundenauftrag ab und ermöglicht in erster Linie die Aufteilung eines Kundenauftrags in produktsspezifische Aufträge. Für einen Produktionsauftrag wird angenommen, dass dieser immer im Ganzen gefertigt wird. Das bedeutet, dass Produktionsaufträge nicht teilweise gefertigt und ausgeliefert werden können. Diese Annahme ist zulässig, sofern der vereinbarte Liefertermin für den ganzen Auftrag gilt. Besitzt ein Kundenauftrag in der Praxis mehrere, mengenspezifische Liefertermine, so kann der Kundenauftrag in weitere Produktionsaufträge mit verschiedenen Lieferterminen aufgespalten werden.

Ein Produktionsauftrag  $pa$  enthält zum größten Teil die gleichen Variablen wie ein Kundenauftrag (vgl. Tabelle 5-6). Statt einer Referenz zu den Produktionsaufträgen kommt jedoch eine Referenz zum übergeordneten Kundenauftrag  $v$  hinzu. Außerdem enthält ein Produktionsauftrag einen Auftragszustand  $q$ , welcher im nachfolgenden Kapitel näher erläutert wird. Tabelle 5-10 fasst die Variablen des Produktionsauftrags zusammen.

Tabelle 5-10: Übersicht zur Produktionsauftragskomponente

Variable	Erklärung
(Kunden-)Auftrag $v$	Referenz zum übergeordneten Kundenauftrag.
Auftragszustand $q$	Gibt den aktuellen Fertigungsstand eines Produktes/Auftrages wieder und steuert somit den Ablauf innerhalb der Fertigung. Vgl. Kapitel 5.2.2.3.

#### 5.2.2.4 Modellkomponente: Fertigungsplan

Der Fertigungsplan gibt einen Herstellungsprozess vor. Er erfüllt zwei Zwecke: Erstens kann ihm entnommen werden, welche Menge an Stationstypen  $MS_{typ}$  für welche Transformation eingesetzt werden kann und zweitens beschreibt er, mit welcher Wahrscheinlichkeit  $\{P_{q_0, q_1}, \dots\}$  ein nächster Zustand  $q$  eintritt. Als Beispiel sei ein unbearbeiteter Stahlzylinder genannt, der im nächsten Schritt rundgedreht werden soll. Der Fertigungsplan gibt vor, dass nur Drehmaschinen dafür in Frage kommen. Außerdem bestimmt er, mit welcher Wahrscheinlichkeit in dem Beispiel ein rundgedrehter Zylinder oder Ausschuss entstehen wird. Ein Produkt hat immer genau einen Fertigungsplan.

Die Fertigungsplankomponente enthält einen gerichteten, azyklischen Graphen, welcher beschreibt, wie ein Zwischenprodukt mit einer gewissen Wahrscheinlichkeit in den nächsten Zustand transformiert wird. Ein solcher Graph ist nachfolgend exemplarisch dargestellt (vgl. Beispiel in Abbildung 5-7).

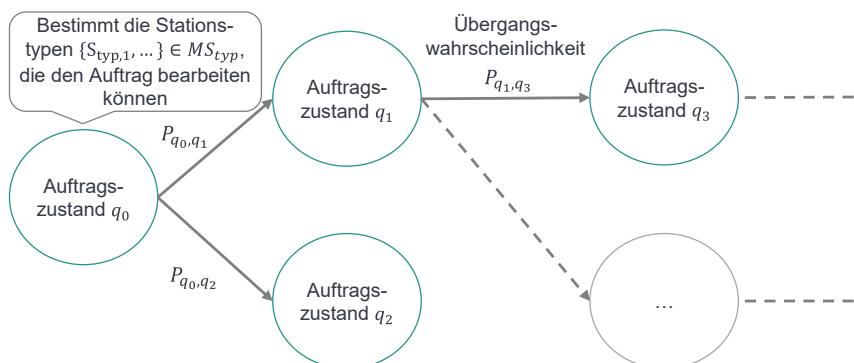


Abbildung 5-7: Beispieldarstellung eines azyklischen Graphen als Repräsentation eines Fertigungsplans

Der Fertigungsplan ist somit die Grundlage der Modellkomponente PPS, um eine geeignete Station für ein Zwischenprodukt in einem bestimmten Zustand auszuwählen. Hierzu wird mithilfe des aktuellen Produktionsauftragszustandes  $q$  und dem Graphen eine Liste aller möglichen Stationen erstellt, die den nächsten Transformationsprozess ausführen können. Die Variablen des Fertigungsplans sind in Tabelle 5-11 zusammengefasst.

*Tabelle 5-11: Übersicht zur Fertigungsplankomponente*

Variable	Erklärung
Stationstypen $\{S_{typ,1}, \dots\} \in MS_{typ}$ für $q$	Stationstypen, welche für die Bearbeitung eines Produktionsauftrags im Zustand $q$ möglich sind
Übergangswahrscheinlichkeiten $\{P_{q0,q1}, \dots\}$	Menge der Übergangswahrscheinlichkeiten, die den Ausgang eines Fertigungsprozesses bestimmen.

Zur Erstellung von Fertigungsplänen für das Versuchsmodul wurde eine quelloffene Software adaptiert, die öffentlich zugänglich ist (Stamer 2021a). Der Fertigungsplan basiert auf der vom Verfasser der Arbeit betreuten Abschlussarbeit A\_Werle (2021).

### 5.2.2.5 Modellkomponente: Produktionsplanung und -steuerung

Die PPS hat unter anderem die Aufgabe der Ablaufplanung. Diese spielt für die Zielerfüllung eine zentrale Rolle und wird daher fokussiert. Im Kontext dieser Arbeit gehen in das Design der Ablaufplanung drei Faktoren ein: Erstens soll die Ablaufplanung Produktionsaufträge in das Produktionssystem einplanen und terminieren, sodass alle Aufträge pünktlich erfüllt werden können. Zweitens muss die Ablaufplanung mit dem Lernprozess des Agenten kompatibel sein. Drittens muss die Ablaufplanung die Nivellierung der Auslastung unterstützen. Diese drei Faktoren werden nun anschließend kurz diskutiert und dann das Vorgehen zur Ablaufplanung dargelegt.

Zur Berücksichtigung des ersten Faktors, der verspätungsminimierenden Einplanung, muss es beispielsweise möglich sein, dass Aufträge von lieferzeitsensitiven Kunden mit hoher Preisbereitschaft im laufenden Betrieb vor Aufträge von nicht-lieferzeitsensitiven Kunden mit niedriger Preisbereitschaft gezogen werden können. Andernfalls kann die Priorisierung der sequenziell eintreffenden Aufträge nicht vorgenommen werden. Deshalb entfallen einfache Steuerungskonzepte wie FCFS. Vielmehr ist es nötig, eine frist-

basierte Steuerung zu nutzen, die die Verspätungsstrafen minimiert. Hierzu gibt es bereits verschiedene Konzepte, wie beispielsweise die „Minimum Slack“-Regel (vgl. Kapitel 2.2.1).

Für den zweiten Faktor muss die Ablaufplanung auch dem Lernprozess des Agenten Rechnung tragen. Arbeiten der Agent und die Steuerung am Ende schlecht zusammen, wird der Ansatz insgesamt schlecht abschneiden. Daher sollte eine Steuerung eingesetzt werden, die auch zum Agenten passt. Für den Agenten wird ein Reinforcement Learning Verfahren eingesetzt. Ein Faktor im Zielsystem dieser Arbeit ist, ob Verspätungen aufgetreten sind. Der Agent beinhaltet im Lernen auch zukünftige Auswirkungen seiner Aktionen durch die Lernfunktion. Dies lässt sich direkt der Formel 2-17 aus Kapitel 2.6.4 entnehmen.

$$\theta_{t+1} = \theta_t + \alpha^\theta (R_{t+1} + \gamma \hat{v}(S_{t+1}, w) - \hat{v}(S_t, w)) \nabla \ln \pi(A_t | S_t, \theta) \quad \text{Formel 2-17}$$

Durch die Diskontierung durch  $\gamma$  ist ersichtlich, dass der Lerneffekt stärker ist, wenn der Lernimpuls unmittelbar mit der verursachenden Aktion verknüpft ist. Außerdem kann die Rückführung von Belohnungen durch die verschiedenen Verzweigungen im Spielbaum zu Überlagerungen führen, die das „Erkennen“ schlechter Aktionen zusätzlich behindert. In der Folge konvergiert der Agent nicht oder nur langsam zum Optimum. Deshalb ist es vorteilhaft für die Stabilität des Verfahrens, wenn bewertbare Konsequenzen zeitlich nah mit Aktionen verknüpft sind. Übertragen auf die PPS bedeutet dies, dass komplexe Verfahren, die erst spät zu Situationen mit vielen Verspätungen führen und dann Kaskadeneffekte aufweisen, eher ungeeignet sind. Als Beispiel sei hier eine Reihe von Aufträgen genannt, die jeweils keinerlei Schlupf haben. Würde nun zum Beispiel ein neuer Auftrag gemäß einer Sortierung nach Liefertermin eingesteuert, gäbe es einen Kaskadeneffekt mit einer Verspätung in Höhe der Summe der Verspätungen aller Aufträge, die nach dem neuen Auftrag kämen. Dieser Effekt wird in Abbildung 5-8 gezeigt. Die Balkenlänge zeigt die Dauer eines Auftrags an. Der grüne Balken ist ein neu einzuplanender Auftrag. Das schwarze Stoppsignal zeigt die vereinbarte Lieferzeit an. Es zeigt sich, dass durch das Einplanen eines neuen Auftrags zwei bestehende sich verspäten. Bei einer Planung, die Verspätungen bestehender Aufträge blockiert, tritt dieser Effekt nicht auf (vgl. Abbildung 5-8 unten). Diesem Ansatz wird im Folgenden nachgegangen.

	Auftrag	Zeitachse							Dauer	Liefer-termin
		0	1	2	3	4	5	6		
Mit Kaskaden-effekt	1	████		∅	████		∅	████		2 3
	2 (neu)									2 4
	3				→████		∅	████		2 5 (+1)
	4					→████		∅		1 6 (+1)
Ohne Kaskaden-effekt	1	████		∅	████		∅	████		2 3
	2 (neu)									2 4
	3	████		∅	████		∅	████		2 5
	4				████		∅	████		1 6

Abbildung 5-8: Darstellung des Kaskadeneffekts bei der Einplanung von Aufträgen

Der dritte Faktor zielt auf die Nivellierung der Stationsauslastung ab. Hierbei sollen, falls möglich, sogar Stationen ganz eingespart werden. Eine Einsparung ist möglich, wenn es gelingt, mit wenigen, konstant hochausgelasteten Stationen alle Aufträge zu bedienen.

Basierend auf diesen Überlegungen bietet sich eine Ablaufplanung an, welche sich mit dem Ziel, Lieferverspätungen zu vermeiden, an der „Minimum Slack“-Regel orientiert (vgl. Kapitel 2.2.1). Um jedoch den Kaskadeneffekt zu unterbinden, wird die Regel eingeführt, dass ein Auftrag nur dann eingesortiert wird, sofern andere Aufträge hierdurch keine Verspätung erhalten. Stationeinsparungen werden durch folgende Regel ermöglicht: Unter den verfügbaren Stationen werden stets Schonstationen vorgesehen. Diese werden nur eingesetzt, wenn es ansonsten zu Lieferverspätungen käme. Ist der Agent in der Lage, das System so zu steuern, dass eine solche Schonstation nicht benötigt wird, so ist das Einsparungspotenzial dadurch unmittelbar ersichtlich.

Der auf diesen Prämissen entwickelte Algorithmus zur Ablaufplanung wird in Anhang A5 ausführlich beschrieben.

### 5.2.2.6 Modellkomponente: Standort

Der Standort umfasst alle anderen zuvor beschriebenen Modellkomponenten und regelt ihre Interaktion. Er nimmt den Auftrag  $v$  entgegen, teilt diese in Produktionsaufträge  $pa$  auf, identifiziert die relevanten Stationen aus der Menge aller Stationen  $O$  mithilfe

des zugehörigen Fertigungsplans und fragt die Produktionsplanung und –steuerung *PPS* an, den Auftrag auf einer der identifizierten Stationen einzuplanen. Alle Variablen der Standortkomponente sind in Tabelle 5-12 zusammengefasst.

*Tabelle 5-12: Übersicht zur Standortkomponente*

Variable	Erklärung
Auftrag $v$	S. Kapitel 5.2.1.4 und Kapitel 5.2.2.3
Stationen $O$	S. Kapitel 5.2.2.2
Produktionsplanung und -steuerung <i>PPS</i>	S. Kapitel 5.2.2.5

### 5.3 Agentenmodul

Nachdem das Umweltmodul erarbeitet wurde, steht nun der Reinforcement Learning Agent im Fokus. Ziel des Agenten ist es, eine sinnvolle Lieferzeit-Preis-Funktion für einen anfragenden Kunden festzulegen. Er hat hierzu die Möglichkeit, verschiedene Zustandsparameter unter Unsicherheit zu beobachten. Basierend auf der Beobachtung soll der Agent die Parameter einer Lieferzeit-Preis-Funktion festlegen. Diese Aktion führt zu einer Belohnung und einem neuen Umweltzustand. Die Belohnung wird dazu verwendet, das Verhalten des Agenten mithilfe eines Trainingsprozesses in die gewünschte Richtung zu bewegen. Dementsprechend wird in diesem Kapitel erklärt, wie die Belohnungsfunktion, der Umweltzustand und die Aktion zu bestimmen sind. Des Weiteren müssen auch der genaue Aufbau des Agenten und das Vorgehen zum Training dargelegt werden.

Für die weiteren Erläuterungen in diesem Kapitel werden nun zwei Sichtweisen eingenommen. In der ersten Sichtweise wird der Agent als Blackbox behandelt und entsprechend nur vom Agenten gesprochen. Dies erleichtert die Beschreibung des Belohnungswerts bzw. der Belohnungsfunktion (Kapitel 5.3.1), der Zustandsparameter (Kapitel 5.3.2) und der Aktion (Kapitel 5.3.3) an der Systemgrenze zum Agenten. Bei der zweiten Sichtweise wird diese Blackbox aufgebrochen und explizit von Actor und Critic gesprochen, wie es für das Lösungsmodell vorgesehen ist. Diese Sichtweise wird zum Verständnis der Architektur (Kapitel 5.3.4) und des Trainingsprozesses (Kapitel 5.3.5) notwendig. Die beiden Sichtweisen sind auch schematisch in Abbildung 5-9 dargestellt.

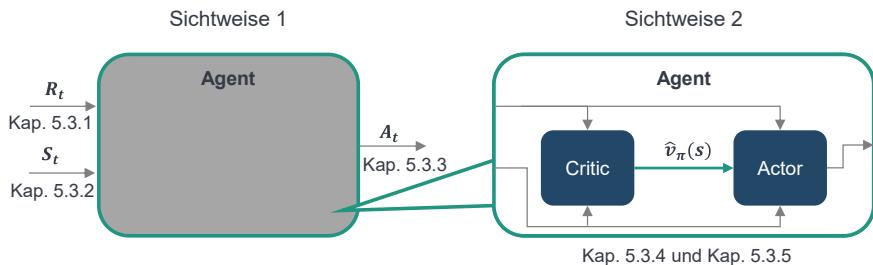


Abbildung 5-9: Darstellung der beiden Sichtweisen auf den Agenten

Das Agentenmodul basiert auf der vom Verfasser der Arbeit betreuten Abschlussarbeit A\_Henzi (2021).

### 5.3.1 Ableitung der Belohnungsfunktion für den Belohnungswert $R_t$

Der folgende Abschnitt befasst sich nun mit der Belohnungsfunktion und dem resultierenden Belohnungswert  $R_t$ .

Die Belohnungsfunktion beeinflusst das Verhalten des Agenten bei der Bestimmung einer Lieferzeit-Preis-Funktion. Es sollte entsprechend im Einklang mit dem Zielsystem stehen. Das Zielsystem (Kapitel 5.1) stellt somit die Grundlage für die Belohnungsfunktion dar. Das Wissen über die Terme der Zielfunktion reicht jedoch nicht aus, um die Belohnungsfunktion aufzustellen zu können. Das Reinforcement Learning ist durch eine zeitliche Dynamik geprägt, sodass es auch zu klären gilt, wann welche Belohnung anfällt.

Ein natürlicher erster Zeitpunkt für eine Belohnung ist der Abschluss einer Auftragsverhandlung. Die Anfrage des Kunden löst eine Aktionswahl des Agenten aus, die wiederum die Reaktion des Kunden hervorruft. Die Reaktion des Kunden beeinflusst unmittelbar die Gewinnfunktion. Die Beeinflussung ist zunächst durch ein Nicht-Bestellen oder ein Bestellen möglich. Im Fall einer Bestellung sind der generierte Umsatz und das durch die Ablaufplanung bereits geplante Auftragsende relevant. Beim geplanten Auftragsende werden Verfrühungen und Verspätungen berücksichtigt. Hier kommt nun eine Besonderheit des Lernprozesses zum Tragen: stark schwankende Belohnungen können zur Instabilität des Lernverfahrens führen. Diese Schwankungen kommen durch große Differenzen bei den Preisen und Bestellmengen einzelner Produkte zu-

stande. Um diesen Effekt auszugleichen, wird nicht direkt der generierte Umsatz, sondern der relative Preisaufschlag des Einzelprodukts als Belohnungswert verwendet (vgl. Kapitel 5.2.1.1):

$$p_{rel,i,v} = \frac{p_{i,v}}{p^*} \quad \text{Formel 5-27}$$

Abbildung 5-10 fasst die Berechnung des Belohnungswertes zum ersten Zeitpunkt zusammen.

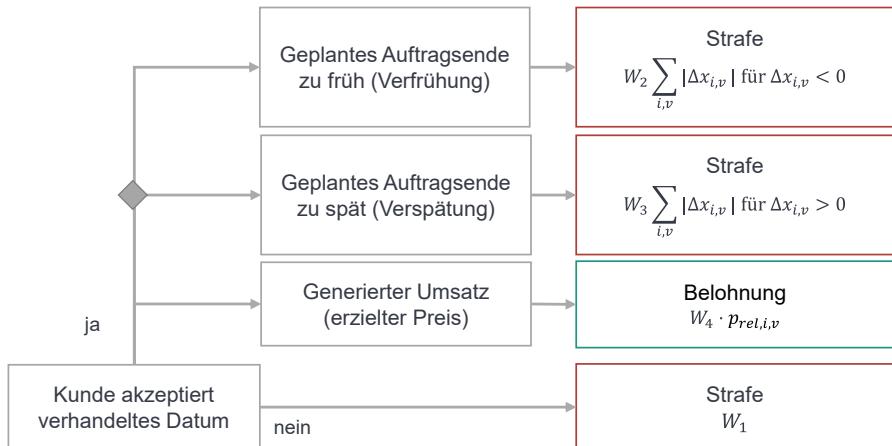


Abbildung 5-10: Belohnungswerte bei Beendigung einer Kundeninteraktion

Der zweite Zeitpunkt ergibt sich als Stützpunkt des Lernprozesses: Generell ist ein Ziel, Kapazitäten einzusparen und real abgeschlossene Aufträge pünktlich auszuliefern. Dies ließe sich theoretisch am Ende der Simulationszeit als Belohnung aggregiert ausschütten. In der Folge wäre jedoch der Trainingsprozess sehr ineffizient, da in die Summe alle vorherigen Aktionen hineinspielen. Daher wird als Stützzeitpunkt für Belohnungsausschüttungen das Ende eines (Simulations-)Tags gewählt. Nach jedem Tag in der Simulation wird dann geprüft, ob es verspätete Aufträge gibt.

Wenn keine verspäteten Aufträge existieren, kommt eine Belohnung für nicht genutzte Kapazitäten in Frage. Ob eine Kapazität zu einem gewissen Zeitpunkt genutzt wurde, wird mit der Variablen  $y_{o,t}$  repräsentiert, die entsprechend null oder eins annimmt. Wird eine Kapazität für einen Tag nicht benutzt, kommt sie für eine Einsparung in Frage. Die Belohnung für diese mögliche Einsparung ergibt sich aus dem Produkt von  $(1 - y_{o,t})$

über einen Zeitraum von einem Tag  $t \in T_{Tag}$ . Das Produkt nimmt nur dann eins an, wenn die Station zu keinem Zeitpunkt verwendet wurde. Durch diese Belohnung hat der Agent den Anreiz, unter Umständen bestimmte Aufträge abzulehnen oder stärker zu nivellieren, d.h. den Lieferzeitpunkt auf günstige Art und Weise zu beeinflussen. Die Berechnung zum zweiten Zeitpunkt wird in Abbildung 5-11 dargestellt.

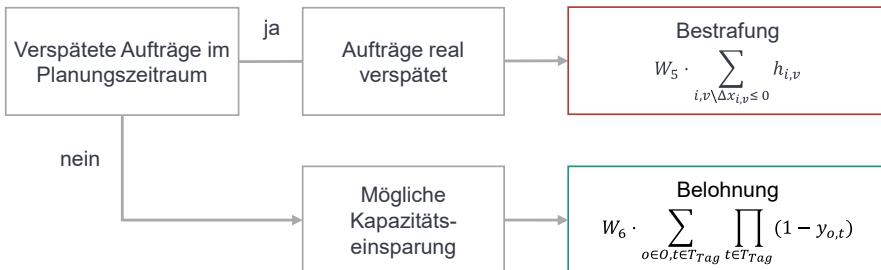


Abbildung 5-11: Belohnungswerte nach Beendigung eines Simulationstages

Um das Einsparen von Kapazitäten noch stärker anzureizen, wird als drittes ein Planungszeitraum von einem (Simulations-)Jahr definiert. Dieser schüttet nochmals eine Belohnung aus, wenn Kapazitäten konstant für ein Jahr  $t \in T_{Jahr}$  nicht genutzt wurden (vgl. Abbildung 5-12). In der Praxis könnte die Kapazität dann veräußert oder für andere Aufgaben eingeplant werden.

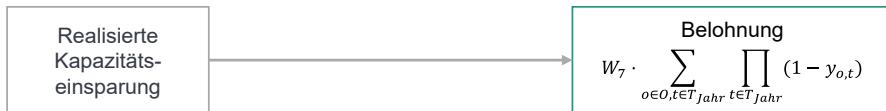


Abbildung 5-12: Belohnungswerte nach Beendigung eines Simulationsjahres

Auf diese Weise ergeben sich nun zusammenfassend drei Zeitpunkte mit unterschiedlichen Belohnungstermen. Die Verrechnung der Belohnungsterme in der zeitlichen Abfolge ist in der nachfolgenden Abbildung 5-13 dargestellt.

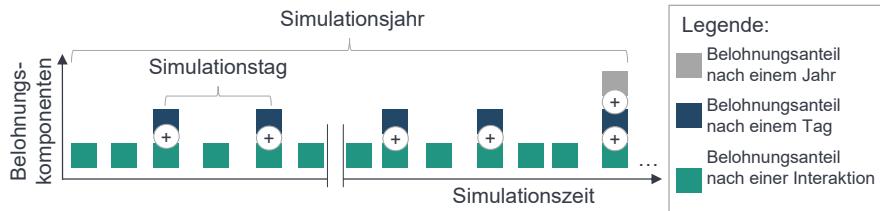


Abbildung 5-13: Berechnung der Belohnungswerte in der zeitlichen Abfolge

Wie den Erläuterungen zu den drei Belohnungszeitpunkten zu entnehmen ist, werden alle Terme mit Gewichtungsfaktoren versehen, sodass die Effekte relativ zueinander eingestellt werden können. Die genaue Höhe der Gewichtungsfaktoren ist empirisch anhand des Lernprozesses einzustellen. Wichtig ist hierbei, dass kein Term komplett aus der Beeinflussung des Lernprozesses heraus verschwindet. Als Beispiel sei hier eine Einstellung herangezogen, bei der der Agent die Strategie lernt, alle Aufträge zu vermeiden. Dies kann auf ein falsches Verhältnis zwischen Abweichungsstrafe, Nicht-Bestell-Strafe und Umsatzbelohnung zurückführbar sein. Die Bestimmung der Gewichtungsfaktoren erfolgt in Kapitel 5.4.2.

Um starken Schwankungen in den Belohnungswerten und damit einem instabilen Trainingsprozess des Agenten vorzubeugen, sollten die Belohnungswerte  $R_t$  auf eine einheitliche Skala von minus eins bis eins normiert werden. Eine Herausforderung ist hierbei, dass der höchste Absolutwert der Belohnung a priori nicht bekannt ist. Deshalb werden zur Laufzeit stetig die Belohnungswerte geprüft und der zum Zeitpunkt der Betrachtung maximale Wert als Basis der Normierung gewählt.

### 5.3.2 Beobachtbarer Zustand $S_t$ der Umwelt

Als nächstes wird nun auf den beobachtbaren (Umwelt-)Zustand  $S_t$ , bestehend aus mehreren Zustandsparametern, eingegangen. Die Zustandsparameter stellen die Entscheidungsgrundlage des Agenten dar.

Zur Bestimmung der beobachtbaren Zustandsparameter gilt zu berücksichtigen, dass nicht alle möglichen Parameter des Umweltmoduls auch für den Agenten als Zustandsparameter verwendet werden können. In der Praxis sind auch nicht alle Parameter für ein Unternehmen beobachtbar, wenngleich sie im Rahmen dieser Arbeit zur Abbildung der Modelllogik in Kapitel 5.1.2 aufgeführt wurden. So werden für die Kunden

Nutzenfunktionen festgelegt, die jedoch für den Agenten nicht direkt sichtbar sein dürfen, weil sie auch in der Realität nicht sichtbar sind. Auf der anderen Seite gibt es Zustände, die zwar prinzipiell dem Agenten sichtbar gemacht werden können, deren Offenbarung jedoch nicht zielführend ist, da hierdurch unnötige Komplexität entsteht.

Generell gilt, dass ein größerer Inputvektor mehr Neuronen auf der ersten Schicht des neuronalen Netzwerks voraussetzt und somit eine größere Komplexität verursacht. Eine größere Komplexität hemmt die Konvergenz des Lernprozesses. Gleichzeitig führt ein zu kleiner Inputvektor zu Informationsverlusten, wodurch es sein kann, dass der Agent nie einen Zusammenhang zwischen Zustand (zum Beispiel Stationsauslastung), Aktion (Lieferzeit-Preis-Funktion) und Belohnung (zum Beispiel Gewinn) herstellen wird. Daher ist die Wahl des Inputvektors, also der zu beobachtenden Zustandsparameter, ein Balanceakt, der empirisch erfolgt. Es gilt die Regel: So viel Komplexität wie nötig, so wenig Komplexität wie möglich.

Als erster Orientierungspunkt kann die bereits aufgestellte Zielfunktion dienen: Es sollten nur die Zustandsparameter sichtbar gemacht werden, die mit der Zielfunktion korrelieren. Hierbei gilt zu berücksichtigen, dass die Zielfunktion allein nicht zwangsläufig zur Identifikation aller relevanten Zustandsparameter führt. Sie bietet lediglich einen guten Ausgangspunkt. Für die effiziente Verarbeitung der Zustandsparameter durch den Agenten kann es zudem von Vorteil sein, diese zu transformieren. Eine Transformation kann helfen, den beobachtbaren Zustandsvektor in eine Form zu bringen, die die Konvergenz des Lernprozesses beschleunigen oder sogar erst ermöglichen kann. Welche Form ein Zustandsparameter haben sollte, lässt sich nur bedingt begründen und basiert auf Erfahrungswissen. Die Transformationen sollen sich an anerkannten Verfahren im Bereich des maschinellen Lernens orientieren (García, Luengo & Herrera 2015, S. 46-55; Aggarwal 2018, S. 127; Hancock & Khoshgoftaar 2020, S. 1-41).

Aus diesen Überlegungen wird folgendes Vorgehen abgeleitet:

- Isolation eines Zielfunktionsterms.
- Aufstellen und Diskussion von Hypothesen hinsichtlich korrelierender Zustände.
- Prüfung auf Sichtbarkeit in der Praxis.
- Überlegungen zur Transformation des Zustandsparameters.

Nachfolgend wird nun die Zielfunktion aus Kapitel 5.1.1 aufgegriffen, um die zu beobachtenden Zustandsparameter argumentativ abzuleiten. Zur Wiederholung ist die Zielfunktion nachfolgend nochmals aufgeführt und in einzelne Terme zerlegt.

$$\max z = \sum_{i,v} ((p_{i,v} - k_{var}) \cdot m_{i,v} - (1 - h_{i,v}) \cdot W_1) - k_{fix} - t_{verfr\ddot{u}ht} \cdot W_2 - t_{versp\ddot{a}tet} \cdot W_3$$

*Formel 5-10*

### 5.3.2.1 Erster Term der Zielfunktion

Der erste Term  $(p_{i,v} - k_{var}) \cdot m_{i,v}$  zeigt die Abhängigkeit des Gewinns zum einen von der Höhe der Lieferzeit-Preis-Funktion und zum anderen von der Beauftragung des Kunden. So könnte ein bestimmter Kunde einen bestimmten Preis noch akzeptieren, wohingegen ein anderer ihn ablehnen würde. Folglich sollte der einzelne Kunde anhand seiner ID  $i$  für den Agenten identifizierbar sein und als Zustandsparameter eingehen. Die Identifizierung sollte auch in der Praxis regelmäßig möglich sein, da beispielsweise für eine Beauftragung Kontakt-, Liefer- und Rechnungsdaten übermittelt werden müssen. Der zweite Faktor  $p_{i,v}$  wird durch die Aktion des Agenten festgelegt. Deshalb handelt es sich nicht um einen Zustandsparameter – er ist das Resultat der Aktion. Dieser Faktor wird folglich nicht weiter betrachtet. Der Kostenfaktor  $k_{var}$  ist ebenfalls zu vernachlässigen, da der Agent lediglich den Preisaufschlag ermittelt und der Kostenfaktor somit keinen Einfluss auf die Aktion hat.

### 5.3.2.2 Zweiter Term der Zielfunktion

Der zweite Term der Zielfunktion  $-(1 - h_{i,v}) \cdot W_1$  gibt an, ob ein Auftrag angenommen wurde. Für den Zustandsvektor ist dieser Term nicht weiter relevant, da er wiederum ein Resultat der Aktion des Agenten ist.

### 5.3.2.3 Dritter Term der Zielfunktion

Der dritte Term  $-k_{fix}$  beinhaltet Kapitalkosten  $k_{Kapital}$  und Instandhaltungskosten  $k_{Instand}$ . Kapitalkosten und Instandhaltungskosten sind abhängig von der Anzahl der benötigten Stationen zur Erfüllung der Aufträge. Die kostenoptimale Planung der Stationen ist nicht direkt durch den Agenten beeinflussbar, sondern wird durch die PPS bestimmt (vgl. Kapitel 5.2.2.5). Jedoch kann unter Umständen eine Station durch eine mittelfristige Reduktion der Auslastung dauerhaft eingespart werden. Hier offenbart sich eine Abhängigkeit zwischen den Kundenbeauftragungen über die Beeinflussung der

Stationsauslastung und -einsparung mit den Kapitalkosten. Diesem Umstand wird jedoch in der Belohnungsfunktion Rechnung getragen (vgl. Kapitel 5.1.1). Folglich ergibt sich kein beobachtbarer Zustandsparameter.

#### 5.3.2.4 Vierter und fünfter Term der Zielfunktion

Die letzten beiden Terme  $-t_{verfr\ddot{u}ht} \cdot W_2 - t_{versp\ddot{a}tet} \cdot W_3$  betreffen die Lieferzeitabweichung. Diese wird sowohl bei Verspätung als auch bei Verfrühung bestraft (vgl. Kapitel 5.2.1.1). Die Lieferzeitabweichung hängt von zwei Faktoren ab:

- Der Lieferzeit-Preis-Funktion in Kombination mit der Kundenpräferenz.
- Der Fähigkeit des Produktionssystems, einen Auftrag zu fertigen.

Da die Lieferzeit-Preis-Funktion die Ausgabe des Agenten ist, wird sie ausgeschlossen.

Die Fähigkeit des Produktionssystems, einen Auftrag zu fertigen, ist für die Erfüllbarkeit eines Auftrags relevant. Der Agent braucht folglich einen Input, der eine Einschätzung zur Möglichkeit der Erfüllung eines Auftrages über die Zeit erlaubt. Im Zentrum stehen hierbei die Stationsauslastungen im Produktionssystem. Es gibt mehrere Möglichkeiten, den Auslastungszustand im Produktionssystem darzustellen. Eine wichtige Einschränkung an dieser Stelle ist, dass sich der Inputvektor für das neuronale Netzwerk in der Länge nicht verändern darf. Gleichzeitig sind für einen Auftrag nicht alle Stationen relevant, sondern nur diejenigen, welche in der Lage sind, den Auftrag zu bearbeiten. Dies führt zu einem Konflikt, wenn die Länge des Inputvektors von der Anzahl in Frage kommender Stationen abhängt. Um diesen Konflikt aufzuheben, sollen zwei Möglichkeiten vorgestellt werden: Die erste Möglichkeit besteht darin, dem Agenten den Auftragsstand aller Stationen zu übergeben und ihm über einen separaten Binärvektor mitzuteilen, welche Station für den Auftrag relevant ist. So kann der Agent lernen, dass nur die Auftragsstände der relevanten Stationen für die Entscheidung zu berücksichtigen sind. Die zweite Möglichkeit liegt in der Repräsentation des Auftragsstands durch eine einzelne Zahl. Hierfür wird die gemäß Planung minimale Fertigungszeit für den betrachteten Auftrag vorgeschlagen. Letzteres senkt die Komplexität der benötigten Architektur des neuronalen Netzes, ohne an Aussagekraft zu verlieren. Daher wird nachfolgend die Aggregation des Produktionssystemzustandes auf eine Kennzahl weiter verfolgt. Hierzu wird eine Abfrage bei der PPS-Komponente durchgeführt, welche auf Basis des aktuellen Auslastungsstandes eine Fertigungszeit  $t_{fz,min}$  ausgibt. Für

stochastische Prozesse ist diese Zeit selbst ebenfalls stochastisch und dementsprechend mit einer gewissen Unsicherheit behaftet (vgl. Anhang A5).

### 5.3.2.5 Transformation und Zusammenfassung der Zustandsparameter

Die in den vorausgegangenen Abschnitten abgeleiteten Zustandsparameter, nämlich die Kundenidentifikation und die Stationsaustastung, können nicht direkt als Input für das neuronale Netzwerk verwendet werden. Stattdessen ist es notwendig, die Daten zu normalisieren (García, Luengo & Herrera 2015, 46-55). Hierfür bietet sich eine Min-Max-Normalisierung an (Aggarwal 2018, S. 127).

$$x_{skaliert} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad \text{Formel 5-28}$$

Als Abbildungsintervall für die Normierung wird der in der Literatur übliche Bereich [0,1] gewählt, welcher auch zum aktiven Bereich der Aktivierungsfunktionen passt. Offen bleibt, wie die Grenzen der Min-Max-Skalierung festgelegt werden. Dies muss für jeden Faktor einzeln überlegt werden.

Als erstes gilt zu klären, in welcher Form die Kundenidentifikation dem Agenten übergeben werden sollte. Als Kodierung für Klasseninformationen bzw. für kategorische Daten, zu denen auch die Kundenidentifikation gehört, hat sich im Stand der Forschung das sogenannte One-Hot-Encoding (deutsch: 1-aus-n-Code) durchgesetzt (Hancock & Khoshgoftaar 2020, S. 1-41). Der Vorteil liegt in der effizienten Darstellung für die spätere Verarbeitung durch den Agenten. Daher wird dem Stand der Forschung gefolgt und auch die Kundenidentifikation dergestalt transformiert. Der resultierende One-Hot-Binärvektor  $i_{one-hot}$  bedarf keiner weiteren Normalisierung.

Die Stationsauslastung in Form der minimal möglichen Fertigungszeit  $t_{fz,min}$  ist hingegen für die Normalisierung näher zu betrachten. Als Minimum existiert ein natürlicher Nullpunkt, da eine Fertigungszeit nie negativ werden kann. Für das Maximum muss hingegen eine Annahme getroffen werden. Es wird angenommen, dass das Maximum der Fertigungszeit nicht größer als die maximal vereinbare Lieferzeit ist. Diese Annahme kann falsch sein, sodass der Zustand einen Wert größer als eins annimmt. Dies ist jedoch für die korrekte Funktionsweise in gewissen Grenzen nicht kritisch.

Abschließend werden die zu beobachtenden Zustandsparameter nochmals tabellarisch zusammengefasst.

Tabelle 5-13: Überblick über die Zustandsparameter

Zustandsparameter	Beschreibung
Kundenidentifikation $i_{one-hot}$	Die Kundenidentifikation in One-Hot-kodierter-Darstellung
Stationsauslastung $t_{fz,min}$	Die [0,1] skalierte Stationsauslastung repräsentiert über das Minimum der Fertigungszeit für einen neuen Auftrag

### 5.3.3 Aktion $A_t$ des Agenten

Im dritten Schritt wird nun die Aktion des Agenten fokussiert. Die Aktion des Agenten parametrisiert die Lieferzeit-Preis-Funktion, aus welcher der Kunde wählen darf. Daher ist es sinnvoll, als Erstes die Lieferzeit-Preis-Funktion herzuleiten und dann die Parameter zu bestimmen, die durch den Agenten festgelegt werden.

Es wird angenommen, dass der Hersteller immer genau eine Lieferzeit hat, zu der er kostenminimal produzieren kann. Den Kostenverlauf bildet der Hersteller im Sinne einer Margenrechnung unmittelbar auf die Lieferzeit-Preis-Funktion ab, sodass der Zeitpunkt für eine kostenminimale Produktion mit dem Tiefpunkt der Lieferzeit-Preis-Funktion zusammenfällt. Weiter wird angenommen, dass die Lieferzeit-Preis-Funktion asymmetrisch ist. Sie fällt zunächst bis zum Minimum exponentiell ab und nimmt dann quasilinear zu. Damit wird den hohen Verspätungs- und Umlaufkosten Rechnung getragen, zu denen es bei kurzen Lieferfristen kommen kann. Der flache Verlauf bei langen Lieferzeiten soll zum einen Risiko- und Lagerkosten abbilden und den Kunden zum anderen einen Anreiz bieten, zum optimalen Punkt des Herstellers zu bestellen. Die Funktion, die diese Eigenschaften erfüllt, ist ähnlich zur Nutzenfunktion des Kunden (vgl. Formel 5-14) und ergibt sich wie folgt:

$$p_{i,v}(x_{i,v}, K_{i,v}, L_{i,v}, M_{i,v}, N_{i,v}) = K_{i,v} e^{(L_{i,v} - x_{i,v})} - M_{i,v} (L_{i,v} - x_{i,v}) + N_{i,v} \quad \text{Formel 5-29}$$

Die Parameter  $K_{i,v}$ ,  $L_{i,v}$ ,  $M_{i,v}$ ,  $N_{i,v}$  dienen hierbei der Justierung der Funktion und beeinflussen die Lieferzeit-Preis-Funktion analog zu  $a_{i,v}$ ,  $b_{i,v}$ ,  $c_{i,v}$  und  $d_{i,v}^u$ . Über  $K_{i,v}$  und  $M_{i,v}$  kann die Steilheit der Flanken eingestellt werden.  $N_{i,v}$  kann interpretiert werden als Verschiebung der gesamten Lieferzeit-Preis-Funktion auf ein anderes Preisniveau.  $L_{i,v}$  beeinflusst die für den Hersteller optimale Lieferzeit, also die Minimalstelle der Lieferzeit-Preis-Funktion. Weiter gilt analog zu den Ausführungen in Kapitel 5.2.1.1 für  $a_{i,v}$  und  $c_{i,v}$ , dass  $K_{i,v}$  und  $M_{i,v}$  zur Findung des Optimums nur positive Werte annehmen dürfen.

Hierzu werden wieder die erste und zweite Ableitung gebildet, um notwendiges sowie hinreichendes Kriterium zu prüfen. Es wird zudem wiederum  $K_{i,v}$  gleich  $M_{i,v}$  gesetzt, sodass sich das Optimum mit  $L_{i,v}$  leicht bestimmen lässt. Dies folgt der gleichen, mathematischen Logik wie Formel 5-18 und Formel 5-19.

$$p'_{i,v} = -K_{i,v}e^{(L_{i,v}-x_{i,v})} + M_{i,v} \stackrel{!}{=} 0 \Rightarrow x_{i,v}^{opt} = \ln(K_{i,v}) - \ln(M_{i,v}) + L_{i,v} \quad \text{Formel 5-30}$$

$$p''_{i,v} = K_{i,v}e^{(L_{i,v}-x_{i,v})} \stackrel{!}{<} 0 \Rightarrow p''_{i,v}(x_{i,v}^{opt}) = M_{i,v} > 0 \quad \text{Formel 5-31}$$

$$K_{i,v} := M_{i,v} \quad \text{Formel 5-32}$$

$$x_{i,v}^{opt} = L_{i,v} \quad \text{Formel 5-33}$$

Weiter soll gelten, dass das Minimum der Lieferzeit-Preis-Funktion immer größer oder gleich null ist. Aus der Bestimmung des Minimums in Verbindung mit dieser Annahme folgt, dass  $N_{i,v} + K_{i,v}$  immer größer als 0 sein muss. Daher wird  $N_{i,v}^l$  eingeführt und definiert (vgl. Formel 5-34).

$$N_{i,v}^l = N_{i,v} - K_{i,v} \text{ mit } N_{i,v}^l > 0 \quad \text{Formel 5-34}$$

$N_{i,v}^l$  kann hierbei als Abschöpfung der Zahlungsbereitschaft des Kunden interpretiert werden. Die Lieferzeit-Preis-Funktion ergibt sich somit als:

$$p_{i,v}(x_{i,v}, K_{i,v}, L_{i,v}, N_{i,v}^l) = K_{i,v}e^{(L_{i,v}-x_{i,v})} - K_{i,v}(L_{i,v} - x_{i,v}) - K_{i,v} + N_{i,v}^l \quad \text{Formel 5-35}$$

Über die Variierung der Parameter lässt sich die Lieferzeit-Preis-Funktion nun so einstellen, dass sowohl auf Kundencharakteristika als auch auf die Kapazitätsauslastung reagiert werden kann. Mit der so definierten Lieferzeit-Preis-Funktion lässt sich nun durch Einsetzen in die Nutzenfunktion des Kunden (vgl. Formel 5-23) folgende Nutzenfunktion darstellen:

$$\begin{aligned} u_i(p_{i,v}(x_{i,v}, K_{i,v}, L_{i,v}, N_{i,v}^l), m_{i,v}, x_{i,v}) &= \\ m_{i,v}(-a_{i,v}e^{(x_{i,v}-b_{i,v})} + a_{i,v}(x_{i,v} - b_{i,v}) + a_{i,v} + d_{i,v}^u) &= \\ -p_{i,v}(x_{i,v}, K_{i,v}, L_{i,v}, N_{i,v}^l) &= \\ m_{i,v}(-a_{i,v}e^{(x_{i,v}-b_{i,v})} + a_{i,v}(x_{i,v} - b_{i,v}) + a_{i,v} + d_{i,v}^u & \\ - (K_{i,v}e^{(L_{i,v}-x_{i,v})} - K_{i,v}(L_{i,v} - x_{i,v}) - K_{i,v} + N_{i,v}^l)) & \\ \forall v \in V \end{aligned} \quad \text{Formel 5-36}$$

Es wird deutlich, dass der Kundennutzen  $u_i$  abhängig von der angebotenen Lieferzeit-Preis-Funktion ist. Die sich aus der Kombination der Nutzenfunktion und der Lieferzeit-

Preis-Funktion ergebende Gesamtnutzenfunktion wird exemplarisch für eine Parametervariation in Abbildung 5-14 dargestellt.

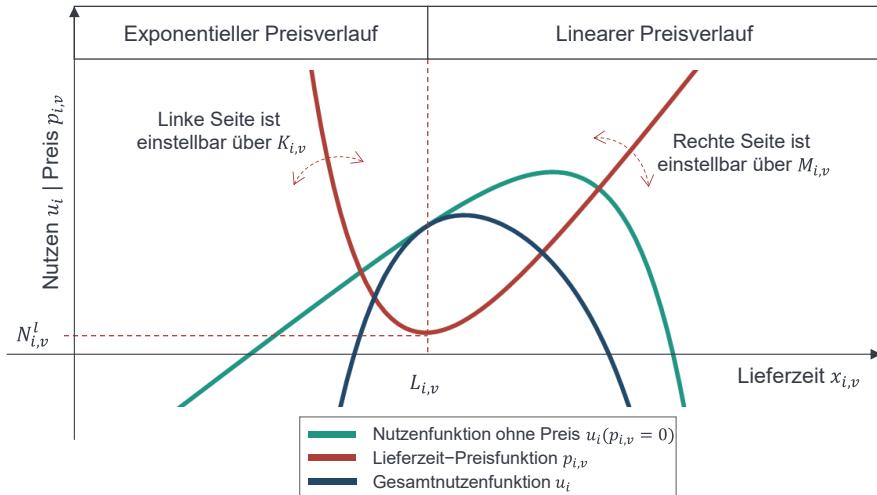


Abbildung 5-14: Nutzenfunktion in Kombination mit Lieferzeit-Preis-Funktion

Für die Modellierung des Kundenverhaltens werden die Nutzenfunktion  $u_i$ , die Lieferzeit-Preis-Funktion  $p_{i,v}$  und damit auch die darin enthaltenen Parameter

$a_{i,v}, b_{i,v}, d_{i,v}^u, K_{i,v}, L_{i,v}, N_{i,v}^l$  relativ zum Basispreis  $p^*$  verwendet. Der Grund hierfür liegt in der bereits erläuterten relativen Preisempfindlichkeit (vgl. Kapitel 5.2.1.1). Diese wiederum würde bei absoluten Parametern dem Agenten das Erlernen einer optimalen Politik erschweren, weil er für jedes Produkt und jeden Kunden mit unterschiedlichem Preis eigene Werte für die Parametrisierung finden müsste. Die Relativierung zum Basispreis hat den entscheidenden Vorteil, dass die relativen Parameter pro Kunde konstant werden. Je nach Ausprägung der Einzelwerte

$a_{i,v}, b_{i,v}, d_{i,v}^u, K_{i,v}, L_{i,v}, N_{i,v}^l$  kann die Nutzenfunktion, entgegen dem beabsichtigten Verhalten, unerwünschte Werte annehmen. Deshalb ist es notwendig, den zulässigen Werteräum auf ein sinnvolles Maß einzuschränken, in dem die Funktionen gültig sind. Die Grenzen für die Parameter  $a_{i,v}, b_{i,v}, d_{i,v}^u$  wurden bereits in Tabelle 5-2 hinreichend definiert. Analog dazu finden sich in der nachfolgenden Tabelle 5-14 die zulässigen Werte für die Parameter  $K_{i,v}, L_{i,v}$  und  $N_{i,v}^l$ .

Tabelle 5-14: Grenzen der Funktionsparameter

Parameter	Untere Grenze	Obere Grenze	Begründung
$\frac{K_{i,v}}{p^*}$	0,1	0,9	Begrenzung der Kurvensteilheit auf ein analytisch sinnvolles Maß. Analog zu $\frac{a_{i,v}}{p^*}$ .
$L_{i,v}$	1	30	Es wird angenommen, dass Kunden nicht weiter als 30 Tage im Voraus bestellen. Dementsprechend ist ein Lieferzeit-Preistiefpunkt außerhalb dieser Grenzen nicht notwendig. Analog zu $b_{i,v}$ .
$\frac{N_{i,v}^l}{p^*}$	0	1,5	Es wird angenommen, dass die maximale, relative Preisbereitschaft 150 % beträgt. Entsprechend sollte das Preismaximum nicht höher als 150% liegen. Anwendungsfallspezifische Anpassungen sind möglich. Analog zu $\frac{d_{i,v}^u}{p^*}$ .

Gemäß der First-Choice-Regel wählt ein Kunde nun immer die Alternative mit dem höchsten Nutzenbeitrag. Der Auftragsverzicht wird als Alternative mit einem Nutzenbeitrag von null angesehen. Zur Findung der Alternative mit dem höchsten Nutzenbeitrag müssen zunächst die möglichen Maximalstellen identifiziert werden.

Es werden hierzu mithilfe der Methoden der Analysis die Nullstellen der ersten Ableitung gesucht und das hinreichende Kriterium geprüft (Formel 5-37ff). Eine Herleitung von Formel 5-37 und Formel 5-38 findet sich in Anhang A6.

$$\frac{du_i}{dx_{i,v}} = u'_i = -a_{i,v}e^{(x_{i,v}-b_{i,v})} + a_{i,v} + K_{i,v}e^{(L_{i,v}-x_{i,v})} - K_{i,v} \stackrel{!}{=} 0 \quad \text{Formel 5-37}$$

$$x_{i,v,1/2}^{opt} = \ln \left( \frac{(a_{i,v} - K_{i,v}) \pm \sqrt{(a_{i,v} - K_{i,v})^2 + 4 \cdot a_{i,v} \cdot e^{-b_{i,v}} \cdot K_{i,v} \cdot e^{L_{i,v}}}}{2a_{i,v}e^{-b_{i,v}}} \right) \quad \text{Formel 5-38}$$

Als Zweites ist die hinreichende Bedingung in Form der zweiten Ableitung zu prüfen.

$$\frac{d^2u_i}{dx_{i,v}^2} = u''_i = -a_{i,v}e^{(x_{i,v}-b_{i,v})} - K_{i,v}e^{(l-x_{i,v})} \stackrel{!}{<} 0 \quad \text{Formel 5-39}$$

Zuletzt muss der sich ergebende Nutzwert größer als null sein.

$$\begin{aligned}
 u_i(p_{i,v}(x_{i,v}, K_{i,v}, L_{i,v}, N_{i,v}^l), m_{i,v}, x_{i,v}) \\
 = m_{i,v}(-a_{i,v}e^{(x_{i,v}-b_{i,v})} + a_{i,v}(x_{i,v} - b_{i,v}) + a_{i,v} + p^* \\
 + d_{i,v}^u - K_{i,v}e^{(L_{i,v}-x_{i,v})} + K_{i,v}(L_{i,v} - x_{i,v}) + K_{i,v} \\
 - N_{i,v}^l) \stackrel{!}{>} 0 \\
 \forall v \in V
 \end{aligned} \tag{Formel 5-40}$$

Der Kunde erteilt nun entsprechend dieser drei Gleichungen einen Auftrag oder lehnt ab.

Im Fall einer ersten Anfrage durch einen Kunden ist dessen Typ bzw. Nutzenfunktion dem Agenten unbekannt. Zu beachten ist, dass die in Kapitel 5.2.1.1 eingeführten, acht Kundentypen rein qualitativ sind und dem Verständnis der Umweltmodellierung dienen. Der Agent wird in der Lage sein, mit einem beliebigen Kunden umzugehen, solange seine Nutzenfunktion ungefähr der modellierten Charakteristik entspricht. Die Nutzenfunktion des Kunden bzw. den Kundentyp muss der Agent daher im Laufe der Interaktionen approximieren, um effizient Parameter der Lieferzeit-Preis-Funktion festzulegen. Bei der konkreten Bestimmung der Parameter  $L_{i,v}, K_{i,v}, N_{i,v}^l$  gilt es nun zu überlegen, welche Werte vom Agenten bestimmt werden sollen und inwiefern die Grenzen enger gelegt werden müssen. Zu beachten ist hier vor allem das Motiv, Intransparenz und Risiko in der Wahrnehmung des Kunden zu senken. Ein Hersteller könnte sich beispielsweise dazu entscheiden, den Parameter  $L_{i,v}$  zu fixieren, somit eine Art Lieferver sprechen geben und damit die wahrgenommene Zuverlässigkeit erhöhen (vgl. Kapitel 2.2.2). Ähnliches gilt für die Parameter  $K_{i,v}$  und  $N_{i,v}^l$ . Hier wäre es möglich, engere Grenzbereiche vorzugeben, um keine hohen Preisschwankungen (Parameter  $N_{i,v}^l$ ) oder enge Lieferzeitkorridore (Parameter  $K_{i,v}$ ) zu riskieren.

Welche Parameter genau durch den Agenten festgelegt werden, ist frei definierbar, weil es keine zwingenden Abhängigkeiten zwischen ihnen gibt. Dies zeigt sich auch im nachfolgenden Kapitel 5.3.4. Insofern hat ein Anwender aus technischer Sicht freien Spielraum, den zulässigen Wertebereich der Parameter zu definieren. Im Rahmen dieser Arbeit wird jedoch angenommen, dass der Agent  $L_{i,v}$  und  $N_{i,v}^l$  die Lieferzeit-Preis-Funktion frei bestimmen darf. Sie werden zu den Aktionsparametern des Agenten. Die Steigung  $K_{i,v}$  wird hingegen fest vorgegeben.

### 5.3.4 Konzeptionelle Architektur von Actor und Critic

Nachdem die Eingangsgrößen (Belohnungswert  $R_t$  und Zustandsparameter  $S_t$ ) sowie die Ausgangswerte (Aktion  $A_t$ ) beschrieben wurden, kann nun die Architektur des Agenten, bestehend aus mehreren neuronalen Netzen, abgeleitet werden. Die Architektur beeinflusst wesentlich die Leistung des Agenten bei der Bestimmung einer Lieferzeit-Preis-Funktion. Das Design der Architektur ist hierbei ein Kompromiss zwischen schnellen Lernzeiten und der Güte der gewählten Aktionen. Zu komplexe Architekturen benötigen eine zu große Rechenzeit, wohingegen eine zu simple Architektur nicht in der Lage sein wird, sinnvolle Lieferzeit-Preis-Funktionen zu bestimmen. Für das weitere Kapitel wird, wie eingangs erläutert, nun die zweite Sichtweise eingenommen und die Blackbox aufgelöst (vgl. Abbildung 5-15). Folglich wird zwischen Actor und Critic differenziert. Die Architekturen von Actor und Critic orientieren sich an den vorgestellten Grundlagen in Kapitel 2.6.1. Es kommen entsprechend vorwärtsgerichtete neuronale Netze zum Einsatz, deren Aufbau hier erläutert wird.

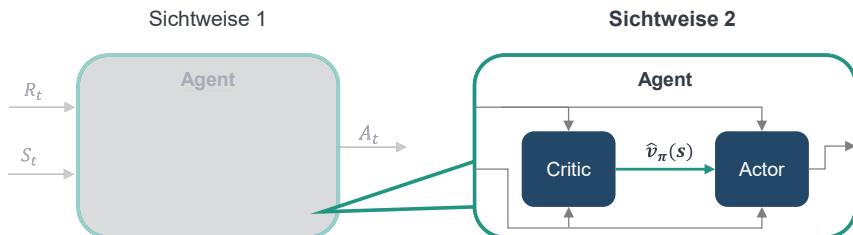


Abbildung 5-15: Einordnung des Kapitels 5.3.4 in die Sichtweisen

Anstatt alle Aktionsparameter durch ein einzelnes neuronales Netzwerk im Actor zu bestimmen, wird für jeden variierbaren Aktionsparameter, also  $L_{i,v}$  und  $N_{i,v}^l$ , ein getrenntes neuronales Netzwerk eingesetzt. Folgerichtig werden auch zwei neuronale Netze für den Critic vorgesehen. Durch die Nutzung von zwei neuronalen Netzen bei der Aktionswahl wird die Komplexität des Lernprozesses gesenkt und stattdessen diese Komplexität in die Architektur verschoben. Es könnte ein Nachteil entstehen, wenn die Parameter in komplexer Wechselwirkung zueinander stehen. In diesem Fall ist es jedoch möglich, die Parameter sequenziell zu bestimmen, ohne Nachteile in Kauf nehmen zu müssen. So lässt sich als Erstes der Lieferzeitiefpunkt mit  $L_{i,v}$  bestimmen und als Zweites der Preisfaktor  $N_{i,v}^l$ . Der Parameter  $K_{i,v}$ , welcher den Lieferzeitkorridor einstellt, soll, wie im vorherigen Kapitel festgelegt, extern vorgegeben sein.

Die neuronalen Netze des Actors sind jeweils nahezu gleich aufgebaut. Sie haben als Eingabeschicht die Anzahl an Zuständen  $|S_t|$ , die zuvor definiert wurden (vgl. Tabelle 5-13). Das Netzwerk, welches den Parameter  $N_{i,v}^l$  bestimmt, hat zusätzlich noch  $L_{i,v}$  als Eingangswert. Zwischen der Eingabe- und Ausgabeschicht liegen die Zwischenschichten. Die genaue Anzahl an Zwischenschichten und Neuronen pro Schicht wird in Abhängigkeit des Anwendungsfalls durch eine Hyperparameteroptimierung bestimmt (mehr dazu in Kapitel 5.4.2). Alle Neuronen einer Schicht sind generell mit allen Neuronen der vorherigen und nachfolgenden Schicht verbunden. Alle Neuronen der Zwischenschicht haben die „leakyReLU“ Aktivierungsfunktion. In der Ausgabeschicht können kontinuierliche Werte berechnet werden, indem die Ausgabeschicht des neuronalen Netzwerks einen Mittelwert  $\mu$  und eine Standardabweichung  $\sigma$  für eine Normalverteilung ausgibt. Durch Ziehen aus einer so definierten Verteilung wird dann der eigentliche Parameterwert bestimmt. Problematisch ist hierbei die Unbegrenztheit der Ausgabewerte. Die Verteilung kann bei beliebigen Mittelwerten und Standardabweichungen beliebige Werte ausgeben. Das ist sowohl für die Skalierung der Aktion auf die vorgefinierten Wertebereiche als auch für den Lernprozess nachteilig. Bei Letzterem kann das Problem explodierender Gradienten auftreten. Daher werden die Ausgabeneuronen mit der Tangens-Hyperbolicus-Funktion bzw. mit der Sigmoid-Funktion auf den Wertebereich von minus eins bis eins bzw. von null bis eins fixiert. Die resultierenden Werte können dann über vorgegebene Grenzwerte passend skaliert werden, gemäß folgender Formel:

$$a_{skaliert} = \frac{\phi^{-1}(x_{rnd}, \mu(\tanh(a_1)), \sigma(\text{sig}(a_2))) + 1}{2} \cdot (OG - UG) + UG \quad \text{Formel 5-41}$$

Die neuronalen Netze des Critics werden ebenfalls basierend auf den Grundlagen in Kapitel 2.5 aufgebaut. Analog zu den Actor Netzwerken haben die Critic Netzwerke eine Anzahl an Eingangsneuronen abhängig von den Zustandsparametern (vgl. Tabelle 5-13) und eine konfigurierbare Anzahl an Zwischenschichten. Auch hier wird wieder auf die „leakyReLU“ Aktivierungsfunktion zurückgegriffen. Die Ausgabeschicht eines Critic Netzwerks besitzt jedoch nur jeweils ein Neuron: Es bestimmt den Vorhersagewert der Zustandswertfunktion und approximiert somit  $v_n(s)$ , welches für das Training des Actor benötigt wird. Es wird wieder die Tangens-Hyperbolicus-Funktion für die Ausgabeschicht verwendet, um Belohnungswerte zwischen minus eins und eins zu erhalten.

Die resultierende Architektur des Agenten ist in Abbildung 5-16 dargestellt. Links im Bild sind die beiden Netzwerke des Critic zu sehen. Sie approximieren jeweils den Zustandswert  $\hat{v}_{\pi,L}(s)$  für die Politik von  $L_{i,v}$  bzw.  $\hat{v}_{\pi,N}(s)$  für die Politik von  $N_{i,v}$ . Auf der rechten Seite sind die beiden Netzwerke des Actor abgebildet, sowie die Abhängigkeit zwischen dem Netzwerk zur Bestimmung von  $N_{i,v}^l$  und dem Netzwerk zur Bestimmung von  $L_{i,v}$ . Der Parameter  $K_{i,v}$ , welcher konstant bestimmt wird, ist zur vollständigen Darstellung der Aktion ebenfalls aufgeführt.

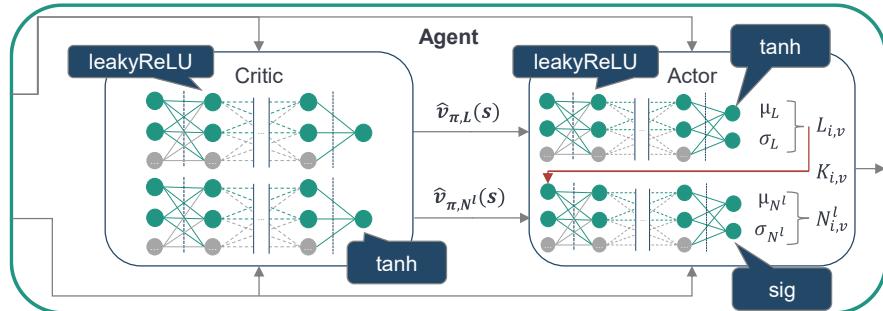


Abbildung 5-16: Architektur des Agenten

Für die Initialisierung der Netzwerke von Actor und Critic werden zufällige Gewichtungswerte aus einer Gleichverteilung gezogen. Diese Gleichverteilung wird auf den Wertebereich  $[-0,4; 0,4]$  begrenzt. Es handelt sich hierbei um einen empirischen Wert aus eigenen Untersuchungen zur Stabilität des Lernprozesses.

### 5.3.5 Trainingsprozess von Actor und Critic

Nachdem die Architektur der Netzwerke von Actor und Critic festgelegt ist, wird nun zuletzt der Trainingsprozess fokussiert. Der Trainingsprozess ist notwendig, um die zunächst zufällig bestimmten Neuronengewichte im Netzwerk dergestalt anzupassen, dass der Agent je nach Stationsauslastung und Kunde eine zielführende Lieferzeit-Preis-Funktion bestimmen kann.

Der Actor besteht in dieser Arbeit aus zwei neuronalen Netzen. Um den Trainingsprozess stabil zu halten, wird zunächst nur hinsichtlich des Lieferzeitparameters  $L_{i,v}$  trainiert. Sobald die Ergebnisse stabil sind, folgt die zweite Trainingsphase für den Preisparameter  $N_{i,v}^l$ . Parallel dazu wird stets das dazugehörige Netzwerk des Critic trainiert.

Für das Training von Actor und Critic selbst werden geläufige Methoden nach dem aktuellen Stand der Forschung angewendet (vgl. Kapitel 2.6.2). So wird der Adam-Algorithmus zur Aktualisierung der Neuronengewichte nach dem Gradientenverfahren verwendet (Aggarwal 2018, S. 140). Einzig die Verlustfunktionen für Actor und Critic sind problemspezifisch.

Für den Critic wird der Literatur folgend der quadratische Abstand aus gewähltem und realisiertem Wert als Verlustfunktion festgelegt (Aggarwal 2018, S. 7, 58). Die Verlustfunktion ist in Formel 5-42 dargestellt.

$$Loss_{critic} = \sum (\hat{v}_\pi(s) - v_\pi(s))^2 \quad Formel\ 5-42$$

Für den Actor gibt es keine Verlustfunktion im ursprünglichen Sinn, da kein Zielwert bestimmt werden kann. Die Leistung des Actor kann aber indirekt am Advantage  $\delta_t$  abgelesen werden (vgl. Kapitel 2.6.4). Wenn der Advantage langfristig gegen null geht, findet kein weiteres Training statt, wie den Gleichungen zur Aktualisierung der Netzwerkgewichte aus Formel 2-16 und Formel 2-17 in den Grundlagen zu entnehmen ist.

$$\theta_{t+1} = \theta_t + \alpha^\theta \delta_t \nabla \ln \pi(A_t | S_t, \vec{\theta}) \quad Formel\ 2-16$$

$$\delta_t = R_{t+1} + \gamma \hat{v}_\pi(S_{t+1}, \vec{w}) - \hat{v}_\pi(S_t, \vec{w}) \quad Formel\ 2-17$$

Die Lernrate des Trainingsprozesses wird nicht fixiert, sondern soll über die Zeit mithilfe einer exponentiellen Glättung abnehmen (vgl. Kapitel 2.6.2). Für die exponentielle Glättung werden drei Parameter benötigt: Startwert, Glättungsfaktor und Endwert. Diese Werte werden im Rahmen einer Hyperparameteroptimierung bestimmt (mehr dazu in Kapitel 5.4.2).

## 5.4 Versuchsmodul

Das Versuchsmodul hat den Zweck, die Durchführung von Simulationsexperimenten zu ermöglichen, um das Potenzial des Ansatzes in verschiedenen Anwendungsfällen nachzuweisen. Es erlaubt die Planung, Durchführung und Auswertung von Experimenten zur Untersuchung verschiedener Anwendungsfälle. Nachfolgend wird zunächst die Architektur und der Idealprozess anhand des Versuchsmoduls vorgestellt. Daran anschließend wird die Hyperparameteroptimierung zur Parametrisierung des Agenten behandelt. Zum Schluss wird auf die statistische Versuchsplanung zur Untersuchung der in der Einleitung aufgestellten Hypothesen eingegangen.

### 5.4.1 Funktionsweise und Aufbau des Versuchsmoduls

Das Versuchsmodul hat den Zweck, die Planung, Durchführung und Auswertung von Experimenten zu ermöglichen. Es überführt hierzu das Umweltmodul und Agentenmodul in ein experimentierbares Modell. Im Rahmen dessen erfüllt es verschiedene Funktionen, die einzeln im Ablauf erläutert und den Modellkomponenten zugeordnet werden. Eine Übersicht über das Versuchsmodul findet sich in Abbildung 5-17.

Um ein Experiment (Modellkomponente: **Experiment**) in bestimmter Konfiguration auszuführen, ist es notwendig, Konfigurationsdaten dem Modell zu übergeben. Das Versuchsmodul nimmt die Konfigurationsdaten entgegen und übersetzt diese (Modellkomponente: **Konfiguration**). Hierzu besitzt das Versuchsmodul eine serverbasierte Schnittstelle (Modellkomponente: **Server**), über die Daten empfangen und abgerufen werden können. Durch die serverbasierte Auslegung werden die Bedienung und Kontrolle von der ausführenden Hardware getrennt, was vorteilhaft für die Parallelisierung und Verteilung von Experimenten innerhalb einer Infrastruktur ist. Über die Schnittstelle ist es möglich, Experimente anzustoßen, Informationen zu laufenden bzw. gelaufenen Experimenten abzufragen und deren Ablauf zu steuern (Modellkomponente: **Experimentator**). Die abfragbaren Informationen und Steuerbefehle sind nicht in menschenlesbarer Form, weshalb eine (grafische) Übersetzungsschicht benötigt wird (mehr dazu in Kapitel 5.5.2). Um ein Experiment auszuführen, wird eine Simulationsengine benötigt (Modellkomponente: **Simulationsengine**). Diese implementiert und verwaltet die diskrete Ereignisliste.

Außerdem stellt sie ein Datums- und Zeitsystem sowie standardisierte Funktionen zur Buchung und Freigabe von Stationen bereit. Eine kritische Modellkomponente ist hierbei die Gewährleistung von Zufälligkeit (Modellkomponente: **Randomness**). Ein Rechner ist ein deterministisches System und hat üblicherweise keine eingebaute Quelle der Zufälligkeit. Stattdessen kann auf pseudozufällige Quellen, wie zum Beispiel die aktuelle Betriebszeit, zurückgegriffen werden. Hierbei besteht jedoch die Gefahr, in eine statistische Abhängigkeit zu geraten, zum Beispiel wenn mehrere Experimente immer exakt im selben Zeitrhythmus gestartet werden. Um dieses mögliche Risiko zu umgehen, wird der Dienst RANDOM.ORG verwendet, der Zufallszahlen einer Gleichverteilung aus atmosphärischem Rauschen bereitstellt<sup>3</sup>.

---

<sup>3</sup> Vgl. <https://www.random.org/> [24.04.2022]

Wird kein Startwert für die Zufälligkeit gegeben, so wird ein Zufallswert von diesem Dienst bezogen. Alle anderen Zufallszahlen basieren dann auf diesem Startwert. Dadurch wird gleichzeitig die exakte Reproduzierbarkeit jedes Experiments gewährleistet. Während einer Experimentausführung überwacht das Versuchsmodul den Experimentzustand, entnimmt hierzu Daten und stellt die Stabilität des Systems sicher (Modellkomponente: **Analytics**). Das Versuchsmodul dokumentiert die ausgewählten Daten des Experiments in einer Datenbank und verwaltet diese (Modellkomponente: **Datenbankverwaltung**). Außerdem übernimmt es die automatische Auswertung der Versuchsdaten zur Gewinnung der Zielgrößenwerte eines Experiments.

Es zeigt sich in der Abbildung 5-17, dass die zentrale Schnittstelle zwischen Versuchsmodul und Agent sowie Umwelt aus Softwaresicht die Engine ist (Modellkomponente: **Engine**), die auch genau zu diesem Zweck vorhanden ist. Unter der Engine ist eine Referenz auf die Modellkomponente Standort und die Modellkomponente Netzwerk aufgehängt, worunter wiederum alle weiteren Modellkomponenten des Umweltmoduls aufgehängt sind (vgl. Kapitel 5.1.2). Wie bereits motiviert, ist als zentrale Verbindung zwischen Standort und Netzwerk das Auftragsmanagement zu sehen, weshalb der Logik folgend die Modellkomponente Kundenauftrag die Verbindung zwischen beiden herstellt. Es sticht heraus, dass das Agentenmodul direkt an das Netzwerk und nicht an den Standort angebunden ist, was auf den ersten Blick entgegen der Intuition erscheint. Der Grund dafür liegt rein in der Implementierung: Die Aufträge eines Kunden entstehen auf der Netzwerkebene und werden dann sofort verhandelt, bevor ein bestätigter Auftrag in die Sphäre des Standorts übergeht. Durch die Referenz zum Agentenmodul in der Netzwerkkomponente wird die Schnittstelle zwischen Standort und Netzwerk aus Softwaresicht stark vereinfacht.

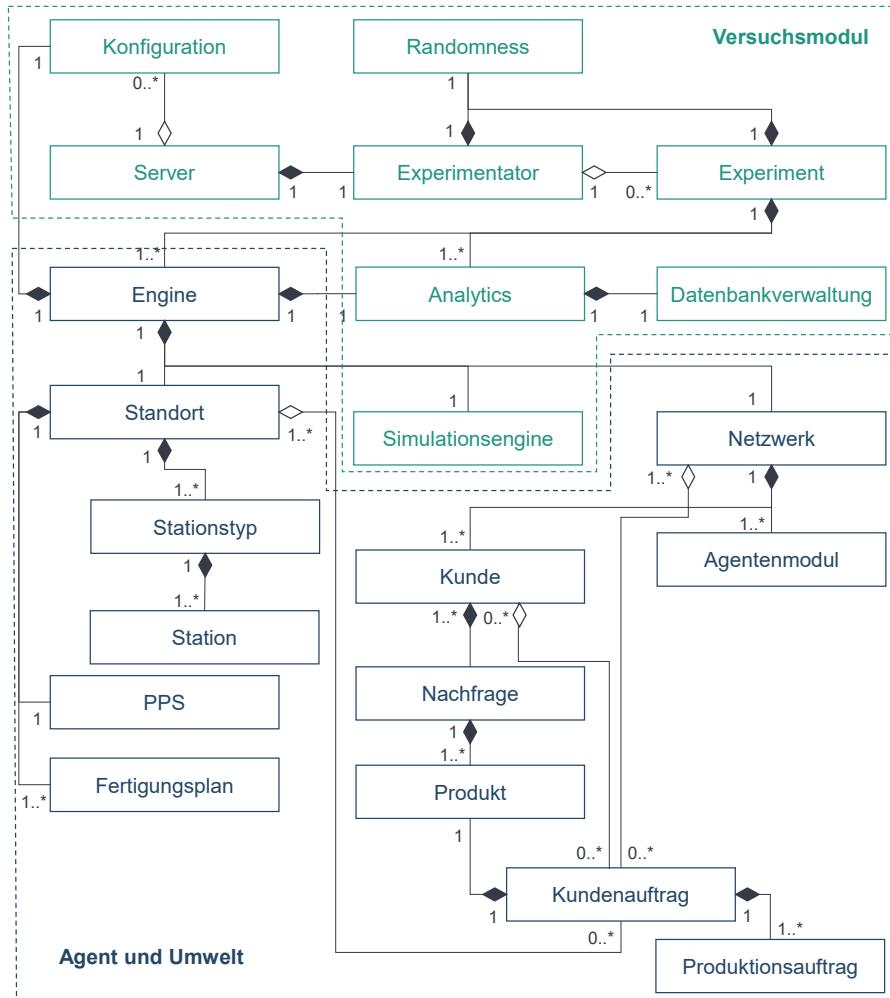


Abbildung 5-17: Überblick zum Aufbau des Versuchsmoduls

#### 5.4.2 Hyperparameteroptimierung

Nachdem die Funktionsweise des Versuchsmoduls allgemein dargelegt wurde, soll es nun darum gehen, wie das Versuchsmodul im Kontext der Arbeit zielgerichtet eingesetzt wird. Hier tritt im ersten Schritt die Hyperparameteroptimierung in den Fokus. Die

Hyperparameteroptimierung ist ein Verfahren, um Parameter, welche nicht durch den Lernprozess adaptiert werden, zielgerichtet einzustellen (Aggarwal 2015, S. 125). Die Wahl der richtigen Parameter ist für den erfolgreichen Einsatz von neuronalen Netzen als Agenten kritisch (Aggarwal 2018, S. 4-30). Sie bestimmen das Lernverhalten des neuronalen Netzwerks und damit die Modellgüte. Um die richtigen Parameterwerte für eine gute Approximation zu finden, bietet es sich an, diese zu variieren und über Experimente die erzielbare Leistung zu beurteilen. Die Hyperparameteroptimierung kann manuell oder automatisiert durchgeführt werden. In der Praxis zeigt sich, dass eine Kombination aus manueller und automatisierter Suche sinnvoll sein kann.

Eine rein automatisierte Suche kann an der Rechenzeit scheitern, die benötigt wird, um eine gute Lösung zu finden. Um diesen Umstand zu verdeutlichen, wird nachfolgend die Relation zwischen Versuchsplan und resultierender Rechenzeit untersucht. Die Dauer der gesamten Versuchsdurchführung  $T_{versuch}$  ist zunächst abhängig von der zu simulierenden Zeitdauer pro Einzelversuch  $T_{simmin}$ . Diese übersetzt sich in die reale Rechenzeit über einen Faktor  $f_{rechenzeit}$ , der wiederum von der Rechenleistung abhängig ist. Weiter multipliziert sich diese Rechenzeit mit der Anzahl an zu untersuchenden Konfigurationen  $n_{konf}$  und der Anzahl notwendiger Wiederholungen pro Konfiguration  $n_{wdh}$ . Letztere sind notwendig, um eine gewisse statistische Sicherheit bei einer nicht-deterministischen Simulation zu erhalten (vgl. Kapitel 2.4.1). Aus diesen Überlegungen kann die Dauer der Versuchsdurchführung wie folgt berechnet werden:

$$T_{versuch} = f_{rechenzeit} \cdot T_{simmin} \cdot n_{konf} \cdot n_{wdh} \quad \text{Formel 5-43}$$

Die Formel zeigt, dass die benötigte Rechenzeit für die gesamte Versuchsdurchführung sehr schnell impraktikable Werte annehmen kann. Deshalb ist eine manuelle Suche im ersten Schritt sinnvoll, um den Lösungsraum einzuschränken. Für die manuelle Suche werden Vorversuche und Erfahrungswissen eingesetzt. Im zweiten Schritt kann die automatische Suche dann innerhalb des Lösungsraums eine möglichst gute Lösung finden. Dieser zweite Schritt erfolgt anwendungsfallspezifisch.

Als Grundlage der Bewertung wird einzig der erzielte Belohnungswert herangezogen. Unter der Annahme, dass die Belohnungsfunktion korrekt aufgestellt ist, d.h. mit dem Zielsystem korreliert, und, dass der Agent eine gute Politik erlernt, bieten die Belohnungswerte eine gute Entscheidungsgrundlage. Um möglichen Fehlern bei der Konfigurationswahl vorzubeugen, werden im dritten Schritt die drei besten Konfigurationen

miteinander hinsichtlich der Zielgrößen (Kapitel 5.1.1.1) verglichen. Durch dieses Vorgehen zeigt sich auch, ob das System prinzipiell funktioniert, weil eine Konfiguration mit hohem Belohnungswert auch vielversprechende Werte in den Zielgrößen erreichen müsste. Aus den drei Konfigurationen wird die beste ausgewählt. Das beschriebene Vorgehen in drei Schritten wird auch in Abbildung 5-18 gezeigt.



Abbildung 5-18: Vorgehen bei der Hyperparameterbestimmung

Bei der Planung einer Hyperparameteroptimierung können Parallelen zur Versuchsplanung hergestellt werden (vgl. Kapitel 2.4.2). Es kann nach demselben Prinzip der Versuchsplan mit den Steuer- und Rauschfaktoren aufgestellt und so eine Struktur in die Planung gebracht werden. Analog werden nachfolgend variierbare und konstante Steuerfaktoren sowie variierbare und konstant bestimmte Rauschfaktoren identifiziert. Soweit sich die Wertebereiche der Faktoren für die später betrachteten Anwendungsfälle einschränken lassen, wird das bereits in diesem Kapitel getan.

Zu den variierbaren Steuerfaktoren können alle Parameter gezählt werden, die das neuronale Netzwerk beeinflussen. Bei diesen Parametern kann unterschieden werden zwischen solchen, die die Topologie des Netzwerks beeinflussen, und solchen, die das Verhalten des Lernprozesses einer Topologie beeinflussen.

Zunächst werden die topologischen Parameter betrachtet. Ausgangspunkt der Topologie ist ein vorwärts gerichtetes neuronales Netz. Die Anzahl der Zwischenschichten und die Anzahl der Neuronen pro Zwischenschicht sind hier wesentliche Steuerfaktoren. Sie beeinflussen die mögliche Dimensionalität der Approximation, wobei sowohl zu viel als auch zu wenig Dimensionalität für die Konvergenz des Lernprozesses problematisch ist. Bezuglich eines sinnvollen Wertebereichs müssen beide Parameter positiv ganzzahlig sein. Es konnte vom Verfasser empirisch bestimmt werden, dass ein Netzwerk mit ein oder zwei Schichten mit jeweils drei bis fünfzehn Neuronen zu guten Ergebnissen führt. Bei den Parametern des Lernprozesses ist die Lernrate als zentraler Steuerfaktor anzusehen. Sie sollte über Experimente anwendungsfallspezifisch optimiert werden. Für die Lernrate wurde der relevante Bereich auf das Intervall von 0,0001

bis 0,1 eingrenzt (Aggarwal 2018, S. 126). Der Glättungsfaktor der Lernrate ist ebenfalls ein für die Optimierung interessanter Parameter und wurde auf den Wertebereich von 0,0001 bis 0,001 bestimmt. Alle Parameter, die einem kontinuierlichen Wertebereich zuzuordnen sind, werden während der Hyperparameteroptimierung aus einer Gleichverteilung  $U$  gezogen. Die variierbaren Steuerfaktoren sind in Tabelle 5-15 zusammengefasst.

*Tabelle 5-15: Übersicht über die variierbaren Steuerfaktoren der Hyperparameteroptimierung inkl. Eingrenzung aus manueller Suche*

Steuerfaktoren	Actor	Critic
Anzahl der Zwischenschichten	{1,2}	{1,2}
Anzahl der Neuronen pro Schicht	{1,2, ..., 15}	{1,2, ..., 15}
Bereich der Lernrate	$U(0,0001; 0,1)$	$U(0,0001; 0,1)$
Faktor des Lernratenabbaus (exponentielle Glättung)	$U(0,0001; 0,001)$	$U(0,0001; 0,001)$

Neben den variierbaren Steuerfaktoren existieren Steuerfaktoren, die konstant gehalten werden. Da diese unabhängig vom Anwendungsfall festgelegt werden können, werden sie bereits in diesem Kapitel bestimmt. Die Bestimmung dieser Faktoren erfolgte ebenfalls basierend auf Vorversuchen und empirischem Wissen durch den Verfasser. Die Faktoren und ihre Werte lassen sich der nachfolgenden Tabelle entnehmen.

*Tabelle 5-16: Übersicht über konstant bestimmte Steuerfaktoren*

Name des konstant gehaltenen Parameters	Wert
Diskontierungsfaktor $\gamma$	0,95
Aktivierungsfunktion der Zwischenschichten	<i>leakyRelu</i>
Kundentypen	8

Auch die Strafen für abgelehnte Aufträge, Verfrühungen, geplante und realisierte Verspätungen sowie die Belohnungen für Preisaufschläge, mögliche und realisierte Kapazitätseinsparungen sind Steuerfaktoren, die jedoch anwendungsfallspezifisch festgelegt werden. Diese werden in Tabelle 5-17 festgehalten.

Tabelle 5-17: Übersicht über anwendungsfallspezifische Steuerfaktoren

Name des konstant gehaltenen Parameters
Strafe für einen abgelehnten Auftrag $W_1$
Strafe für Verfrühung $W_2$
Strafe für geplante Verspätung $W_3$
Belohnung für Preisaufschlag $W_4$
Strafe für realisierte Verspätung $W_5$
Belohnung für mögliche Kapazitätseinsparung (nach einem Tag) $W_6$
Belohnung für realisierte Kapazitätseinsparung (nach einem Jahr) $W_7$

Von den Steuerfaktoren abzugrenzen sind die Rauschfaktoren. Die Rauschfaktoren werden anwendungsfallspezifisch bestimmt. Im Rahmen der Simulation werden nicht alle Rauschfaktoren deterministisch am Anfang festgelegt. Stattdessen werden einige Rauschfaktoren über Zufallsverteilungen beschrieben. Zu den konstanten Rauschfaktoren zählen die Anzahl der Kunden und die Art und die Anzahl der Stationen. Zu den zufallsbasierten Rauschfaktoren zählen die Bestellhäufigkeit und die Bestellmenge bei Kundenanfragen. Außerdem wird auch die Dauer jedes Fertigungsprozesses als Verteilung abgebildet. Für die Rauschfaktoren können a priori keine Grenzwerte festgehalten werden. Sie werden erst im Rahmen der Erprobung anwendungsfallspezifisch bestimmt.

Tabelle 5-18: Übersicht über Rauschfaktoren

Konstante Rauschfaktoren	Zufallsbasierte Rauschfaktoren
Anzahl an Kunden $I$	Bestellhäufigkeit $BH_{i,j}$
Stationstypen $MS_{typ}$	Bestellmenge $BM_{i,j}$
Anzahl Stationen je Stationstyp $O_{styp}$	Fertigungszeit $fz_{j,q,styp}$

Zuletzt ist für die Hyperparameteroptimierung zu klären, welche Strategie bei der Wahl zu testender Konfigurationen verfolgt werden sollte. In dieser Arbeit wird eine zufallsbasierte Suchstrategie verfolgt, die in Anhang A7 begründet wird.

### 5.4.3 Hypothesen zum Testen des Ansatzes hinsichtlich des Potenzials

Das letzte Kapitel zum Versuchsmodul widmet sich dem Vorgehen zur Gegenüberstellung eines Vergleichssystems mit dem hier vorgestellten Lösungsmodell und dem aus dieser Gegenüberstellung ableitbaren Potenzial. Hierfür ist zunächst zu klären, wie das Vergleichssystem aussieht und wie dann das Potenzial zu quantifizieren ist.

Im ersten Schritt wird nun das Vergleichssystem hergeleitet. Ausgangspunkt beim Design des Vergleichssystems soll der in Kapitel 2.2.2 dargestellte Soll-Prozess sein. Dieser wird im Folgenden als Status Quo bezeichnet. Im Status Quo werden Preise nicht dynamisch festgelegt, sondern anhand von Listenpreisen ermittelt. Offen bleibt, wie die Lieferzeit bestimmt wird. Hier werden zwei Möglichkeiten unterschieden: Im ersten Fall wird die Lieferzeit zusammen mit dem Preis fest angeboten (Status Quo mit statischer Lieferzeit). Kunden haben die Möglichkeit, dieses Angebot anzunehmen oder abzulehnen. Im zweiten Fall wird nur ein Preis fest angeboten und es wird jede Lieferzeit akzeptiert, die der anfragende Kunde präferiert (Status Quo mit flexibler Lieferzeit). In beiden Fällen wird also den Grundlagen folgend von festen Listenpreisen ausgegangen. In einem Anwendungsfall kann immer nur ein System von beiden für eine gegebene Nachfrage im Einsatz sein. Mit dem Vergleichssystem mit statischer Lieferzeit käme es entsprechend zu vielen Verspätungen, wenn auf einmal eine flexible Lieferzeit angeboten werden würde. Umgekehrt gilt für das Vergleichssystem mit flexibler Lieferzeit, dass zu wenige Aufträge mit einem statischen Lieferzeitangebot erlangt werden würden. Zur Vollständigkeit werden jedoch stets beide Vergleichssysteme betrachtet. Für das jeweils im Anwendungsfall nicht passende Vergleichssystem wird die Nachfrage angepasst. In der Folge sind die Simulationsläufe von Status Quo mit statischer Lieferzeit und Agent nicht mit den Simulationsläufen von Status Quo mit flexibler Lieferzeit und Agent vergleichbar. Dieses Vorgehen wird auch in Tabelle 5-19 dargestellt.

Tabelle 5-19: Umgang mit verschiedenen Auftragslasten in Anwendungsfällen

	Vergleichssystem: Status Quo mit statischer Lieferzeit wird eingesetzt	Vergleichssystem: Status Quo mit flexibler Lie- ferzeit wird eingesetzt
Fall 1: Status Quo mit statischer Lieferzeit trifft auf Anwendungsfall zu	Keine Anpassung	Senkung der Nachfrage, um Anwendung zu ermöglichen
Fall 2: Status Quo mit flexibler Lieferzeit trifft auf Anwendungsfall zu	Erhöhung der Nachfrage, um Anwendung zu ermöglichen	Keine Anpassung

Für die Quantifizierung des Potenzials wird auf das bereits eingeführte Zielsystem zurückgegriffen (vgl. Kapitel 5.1.1). Als zentrale Zielgröße wurde im Rahmen des Zielsystems bereits der Gewinn eingeführt. Es soll gelten, dass der Ansatz genau dann ein höheres Potenzial als das Vergleichssystem aufweist, wenn der simulierte Gewinn statistisch-signifikant größer ist. Die Reduktion auf eine zentrale Kennzahl birgt jedoch gleichzeitig die Gefahr, bewertungsrelevante Aspekte durch die Aggregation zu vernachlässigen. Daher sollen sekundär auch die anderen bereits eingeführten Zielgrößen des Zielsystems in die Betrachtung eingeschlossen werden, um den Ansatz hinsichtlich Pareto-Optimalität zu untersuchen und das Potenzial nochmals differenzierter zu analysieren.

Für den Potenzialnachweis ist es besonders interessant, verschiedene Szenarien zu betrachten. Eine übergeordnete Hypothese der Arbeit ist, dass der Ansatz besser mit Volatilität bei gegebenen Produktionsmöglichkeiten umgehen kann als das Vergleichssystem. Ausgehend von der übergeordneten Hypothese werden dediziert verschiedene Szenarien betrachtet. Hier wird auf Änderungen der Kundenpräferenzen abgezielt. Ausgangspunkt ist stets eine Menge an heterogenen Kunden, die über die acht definierten Kundentypen repräsentiert werden (vgl. Kapitel 5.2.1.1).

Insgesamt werden drei Szenarien betrachtet:

- Szenario 1: Keine Änderung der Kundenpräferenzen  $a_{i,v}$ ,  $b_{i,v}$  und  $d_{i,v}^u$ .
- Szenario 2: Änderung der Kundenpräferenzen hin zu Kurzfristigkeit ( $b_{i,v} \rightarrow \min$ ).
- Szenario 3: Änderung der Kundenpräferenzen hin zu niedriger Preisbereitschaft ( $d_{i,v}^u \rightarrow \min$ ).

Für jedes einzelne Szenario soll der zentrale Potenzialnachweis anhand des Gewinns mithilfe eines Hypothesentests statistisch abgesichert werden. Hierfür bietet sich ein Mittelwertvergleich mit einem t-Test an. Es wird ein rechtsseitiger Test angewandt. Zudem werden nicht die Mittelwerte der Gewinne direkt miteinander verglichen, welcher größer ist, sondern der Mittelwert der Gewinndifferenz  $\mu_{\Delta G}$  mit einer Konstanten mit dem Wert null. Obwohl dieses Vorgehen nur wie ein Umformulieren derselben Aussage klingt, hat es einen gravierenden Vorteil. Dieser liegt darin, dass der Test aus statistischer Sicht nun sowohl für verbundene als auch für unverbundene Stichproben gilt (Sibbertsen & Lehne 2021, S. 366f). Es muss daher nicht diskutiert werden, ob es sich bei den Stichproben aus der Simulation um Daten aus der gleichen Quelle handelt (verbundene Stichprobe), wie es zum Beispiel bei Tests mit gleichen Probanden der Fall ist. Es können nun die Nullhypothese  $H_0$  und die Alternativhypothese  $H_1$  aufgestellt werden. Die Konstante als Vergleichswert sei zu null festgelegt, um eine signifikant positive Differenz nachzuweisen. Formal lassen sich die Hypothesen wie folgt aufstellen:

$$H_0: \mu_{\Delta G} \leq 0 \text{ und } H_1: \mu_{\Delta G} > 0$$

Formel 5-44

Die Hypothesen sind mit einem Signifikanzniveau von 5 % zu testen, wie es in der Literatur üblich ist (vgl. Kapitel 2.4.2).

## 5.5 Implementierung, Nutzung und Verifikation

Um eine computergestützte Experimentierbarkeit zu ermöglichen, müssen die Module des Kapitels 5 als Software implementiert werden. Dies ist Schwerpunkt des ersten Unterkapitels. Im zweiten Unterkapitel wird die Nutzung des Systems dargelegt. Im dritten Unterkapitel wird die Verifikation der Module und des Programmcodes behandelt, die für die Korrektheit des Ansatzes essenziell ist. Zuletzt wird die Nutzung des sich ergebenden Softwaresystems erläutert.

### 5.5.1 Implementierung

Die Implementierung der Modellkomponenten und Experimente leitet sich unmittelbar aus der Übersicht in Abbildung 5-17 ab und erfolgte in *TypeScript*. Die Implementierung basiert auf den vom Verfasser der Arbeit betreuten Abschlussarbeiten A\_Henzi (2021), A\_Werle (2021) und A\_Ochs (2021). Die resultierende, kompilierte *JavaScript*-Software wird mithilfe von *Node.js* ausgeführt. Daher ist sie beispielsweise mit *Windows 10*,

*Ubuntu 21.04* und *macOS 11* kompatibel. Es wurde bei der Entwicklung ausschließlich auf quelloffene, frei verfügbare Bibliotheken zurückgegriffen. Zu diesen zählt unter anderem *Tensorflow.js*, eine Bibliothek in *Javascript*, die das Implementieren von neuronalen Netzwerken erleichtert. Da bislang keine Implementierung in *Tensorflow.js* für das Reinforcement Learning existierte, wurden weite Teile der grundlegenden Implementierung im Rahmen dieser Arbeit angefertigt. Hierzu zählen

- die Belohnungsfunktion.
- die Handhabung der Zustandsdaten.
- der Aufbau und die Initialisierung der Neuronen in den Netzwerken gemäß der Architektur in Kapitel 5.3.4.
- die Orchestrierung der Zustandsaufnahme, der Aktionswahl und des Trainings von Actor und Critic.
- die allgemeine Speicherallokation und -freigabe für die Variablen<sup>4</sup>.

Eine weitere zentrale Bibliothek, auf die die Implementierung aufbaut, ist *SimJS*, eine freie Implementierung zur Durchführung von ereignisdiskreten Simulationen. Diese wurde erweitert und umgestaltet, sodass sie das Umweltmodul ausführen kann und in der Lage ist, responsiv sowie in parallelen Prozessen genutzt zu werden. Hierzu musste die blockierende, zentrale Ereignisschleife in eine rekursiv-asynchrone Architektur überführt werden. Ein wichtiger Aspekt einer Simulationsbibliothek ist die Erzeugung von Zufallsvariablen, die einer gewünschten Verteilung folgen, wie es zum Beispiel für die Modellierung der Nachfrage oder der Prozessdauer notwendig ist. Im Rahmen dieser Arbeit wurden bereits mehrere Verteilungen vorgestellt, die beispielsweise zur Simulation von Bestellhäufigkeiten oder -mengen verwendet werden sollen. Die vorgestellten Verteilungen sind jedoch in ihrer Normalform für eine Simulation nicht geeignet, da typischerweise keine Zufallszahlenquelle für jede beliebige Verteilung zur Verfügung steht. Stattdessen ist meist eine Zufallsvariable vorhanden, die einer Gleichverteilung folgt. So verhält es sich auch mit den Zufallszahlen, die von *RANDOM.org* bezogen

---

<sup>4</sup> Hinweis: In *NodeJS* wird die Speicherverwaltung automatisiert gehandhabt. *Tensorflow.js* ist jedoch hardwarenah in C implementiert und erfordert eine manuelle Speicherverwaltung, um Speicherlecks zu verhindern.

werden (vgl. Kapitel 5.4.1). Um nun eine Zufallsvariable, die der gewünschten Verteilung folgt, ausgehend von einer Gleichverteilung zu erzeugen, wurde deshalb auf die inverse Transformationsmethode zurückgegriffen (Gass & Fu 2013, S. 815).

Durch die Nutzung von quelloffenen, freien Bibliotheken ist die gesamte Software frei verfügbar und kann von jeder Person mit Zugang zu einem Rechner, der die Voraussetzungen zur Ausführung von *Node.js*-Applikationen erfüllt, ausgeführt und verwendet werden. Die Modelle sind eingebettet in das Versuchsmodul, welches wiederum als Serverapplikation ausgelegt ist. Hierdurch ist es leicht möglich, die Experimente auf diverse Infrastrukturen auszurollen und dementsprechend Berechnungen parallel durchzuführen. Zur einfachen Verteilung der Software wurde die gesamte Anwendung zudem in einen sog. Container gekapselt und kann entsprechend über eine Container-virtualisierungssoftware ausgeführt werden, ohne dass zusätzliche Installationsschritte erfolgen müssen. Die Implementierung wurde von Stamer (2021b) veröffentlicht und steht zum freien Download zur Verfügung.

Für die Dokumentation der anfallenden Daten wird zusätzlich zum implementierten System eine Datenbank benötigt. Dort werden unter anderem die Rohdaten und Auswertungsergebnisse eines Experiments durch das Versuchsmodul gespeichert. Für diese Aufgabe wurde eine *SQLite* Datenbank ausgewählt. Im Testaufbau erwies sich die *SQLite* Datenbank gegenüber dem Speichern der Daten in einzelnen Dateien und einer richtigen SQL-Datenbank als ausreichend leistungsstark, simpel und speichereffizient für die beschriebene Aufgabe.

Die resultierende Architektur ist schematisch in Abbildung 5-19 dargestellt. Eine detaillierte Argumentation zur Auswahl der Bibliotheken und der Implementierung im Allgemeinen findet sich in Anhang A8.

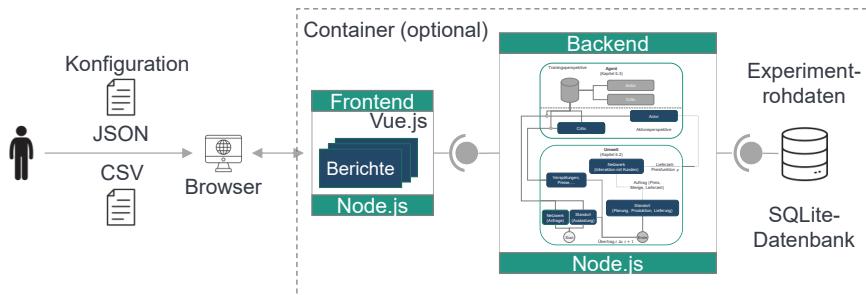


Abbildung 5-19: Schematische Darstellung der Architektur

### 5.5.2 Nutzung

Zur Parametrisierung und Nutzung des implementierten Systems müssen verschiedene Daten zusammengestellt werden. Hierzu wird auf die Dateiformate JSON und csv zurückgegriffen. Das csv-Format wird für Datensätze verwendet, die sich gut tabellarisch darstellen lassen. Hierzu zählen die Inputdaten für das Modell der Umwelt, welche sich zwischen den verschiedenen Testläufen im Rahmen eines Anwendungsfalls typischerweise nicht ändern.

Das JSON-Format wird für alle sonstigen Daten verwendet. Die wichtigste Datei im JSON-Format stellt die Experimentkonfigurationsdatei dar. Sie unterteilt sich in umgebungsspezifische und agentenspezifische Konfigurationen. Über die Konfigurationsdatei können alle vorgesehenen Parameter eingestellt werden (vgl. Kapitel 5.4.2). Die genauen Parameter und der detaillierte Aufbau der Experimentkonfigurationsdatei finden sich in Anhang A9. Die Software selbst besitzt eine Standardkonfiguration. Falls in der Experimentkonfigurationsdatei Parameter fehlen, so wird automatisch auf die Standardwerte zurückgegriffen. Die Experimentkonfigurationsdatei wird über die grafische Benutzerschnittstelle im Browser vor dem Starten eines Experiments hochgeladen. Das Format wird beim Hochladen automatisch auf Korrektheit geprüft. In der Benutzerschnittstelle können zu den gestarteten Experimenten verschiedene Berichte angezeigt und zur Laufzeit ausgewertet werden. Hierdurch sind eine schnelle Analyse des Experiments und ein frühzeitiges Abbrechen im Falle eines negativen Verlaufs möglich.

### 5.5.3 Verifikation

Aufgrund der Komplexität des Gesamtsystems muss bei der Implementierung davon ausgegangen werden, dass Umsetzungsfehler entstehen (VDI 3633 Blatt 1, S. 37). Daraus ist es notwendig, Verifikationsmethoden einzusetzen, um diese Umsetzungsfehler zu beseitigen und die Korrektheit des Systems abzusichern (Rabe, Spiekermann & Wenzel 2008, S. 93f). Im Rahmen der Arbeit werden mehrere Ansätze verfolgt: Zunächst wird die Implementierung in möglichst kleinen Funktionsinkrementen umgesetzt. Durch dieses Vorgehen wird der Test- und Fehlerbeseitigungsaufwand reduziert. Die Reduktion folgt aus dem minimierten Fehlerpotenzial und dem begrenzten Suchraum, welche auf die neu implementierten Funktionen reduzierbar sind. Durch systematisches Testen aller möglichen Sonderfälle im Code bzw. deren Absicherung wird sodann ein funktionierender Entwicklungsstand in den nächsten überführt. Um den Testaufwand weiter zu senken, werden automatisierte Testroutinen aufgebaut. Diese entstehen aus

den manuellen Tests der Funktionsinkremente und reduzieren das Risiko, dass neue Funktionen Quereinflüsse auf bestehende Funktionen haben. Die Tests können anhand der Allgemeinheit der Aussage unterteilt werden. So gibt es Tests, die eine Aussage zum korrekten Ablauf des Gesamtsystems geben. Diese helfen, die ganzheitliche Funktionalität zu prüfen und einen Überblick zu geben. Jedoch können hierbei kleinere Fehler, die die Ausführung nicht gefährden, aber zu Verzerrungen der Ergebnisse führen, übersehen werden. Hier kommen spezifische Tests ins Spiel, die einen kleinen Bereich des Gesamtsystems fokussieren.

Zusätzlich zu den Testroutinen wird der gesamte Ansatz im Rahmen der vom Bundesministerium für Bildung und Forschung (BMBF) öffentlich geförderten Projekte ReKoNeT und BaSys4SupplyQ diskutiert und das Modell auf Basis der Rückmeldung angepasst. Dabei soll die Gültigkeit des Modells und der Modellergebnisse aus Anwendersicht beurteilt werden.

## 6 Erprobung des Lösungsansatzes

Der Lösungsansatz zur dynamischen Lieferzeit-Preisgestaltung im Kontext variantenreicher Produktion wird anhand von zwei typischen Anwendungsfällen erprobt. Der erste Anwendungsfall betrifft den Bereich der Hochpräzisionsprodukte und zwar konkret die Fertigung und Kalibrierung von Hochpräzisionsgewichten. Er zeigt die Validität und die Möglichkeiten des Einsatzes des Lösungsansatzes auf. Der zweite Anwendungsfall kommt aus dem Bereich der Handhabungstechnik, konkret der Zerspanung von Pneumatikzylindern. Er hat den Zweck, die generelle Anwendbarkeit zu demonstrieren. Beide Anwendungsfälle zeichnen sich durch eine hohe Variantenzahl, hohe Volatilität und viele, heterogene Kunden aus, sodass sie für eine Erprobung des Lösungsansatzes geeignet sind. Die Durchführung und die Ergebnisse der Erprobung wurden mit Anwendungspartnern aus den Forschungsprojekten BMBF ReKoNeT und BMBF BaSys4SupplyQ diskutiert.

Die zwei Anwendungsfälle werden in den beiden folgenden Kapiteln behandelt. Dabei sind die beiden Kapitel in die Vorstellung des Anwendungsfalls, die Anwendung des Lösungsmodells und die Präsentation der Ergebnisse untergliedert.

### 6.1 Anwendung im Bereich der Hochpräzisionsgewichte

Der erste Anwendungsfall entstammt der Produktion von Hochpräzisionsprodukten, speziell von Hochpräzisionsgewichten. Das Besondere an dieser Produktart ist, dass die Produktion nicht durch die Fertigung der Gewichte maßgeblich charakterisiert ist, sondern durch die Erstkalibrierung bzw. Rekalibrierung der Gewichte. Die Kalibrierung der Gewichte nach Norm ist für eine Vielzahl von Anwendungsbereichen gesetzlich vorgeschrieben. Bei der Fertigung von Hochpräzisionsgewichten handelt es sich folglich um eine Mischform aus Dienstleistung und klassischer Produktion. Der Anwendungsfokus des Lösungsmodells liegt hierbei auf dem Kalibrierablauf. Als Beispielunternehmen wird hier ein deutscher Mittelständler zugrunde gelegt. Die Daten wurden durch Interviews im Rahmen der Forschungsprojekte BaSys4SupplyQ und ReKoNeT erhoben.

#### 6.1.1 Vorstellung des Anwendungsfalls

Der betrachtete Mittelständler hat ca. 50 Mitarbeitende. Er besitzt sowohl Merkmale des Auftragsfertigers als auch des Variantenfertigers: Er fertigt hochindividuelle Gewichte aber auch Gewichte, die der OIML Norm R 111-1 entsprechen (OIML R 111-1). Der

Fokus soll im Folgenden nur auf Gewichten dieser Norm liegen. Für den Anwendungsfall werden die Auftragsdaten von 2001 bis 2019 betrachtet. Allein in den Jahren 2018 und 2019 liegen ca. 29.000 Einträge für Kalibrierungen vor. Bei genauer Betrachtung des Nachfrageverlaufs lässt sich in der Auftragslage eine kontinuierliche Steigerung der Kalibriervorgänge pro Jahr beobachten. Ausgehend von dieser Beobachtung lässt sich der Nachfrageverlauf mithilfe einer Extrapolation forschreiben, um zu untersuchen, wie sich die Nachfrage weiterentwickeln könnte. Der bestehende Nachfrageverlauf von 2001 bis 2019, eine lineare Extrapolation und eine Extrapolation mit prozentualem Wachstum für die nächsten 10 Jahre sind in der nachfolgenden Abbildung 6-1 dargestellt. Die lineare Extrapolation bildet den Mittelwert der Steigung über die letzten fünf Jahre und schreibt diesen Mittelwert fort. Der Mittelwert beträgt rund 1632 Vorgänge pro Jahr. Eine lineare Extrapolation erscheint bei dem vergangenen Nachfrageverlauf eine konservative Schätzung. Würde stattdessen ausgehend von den historischen Daten ein Wachstum von 7 % pro Jahr unterstellt, käme es zu einer Verdopplung der Nachfrage in etwa 10 Jahren.

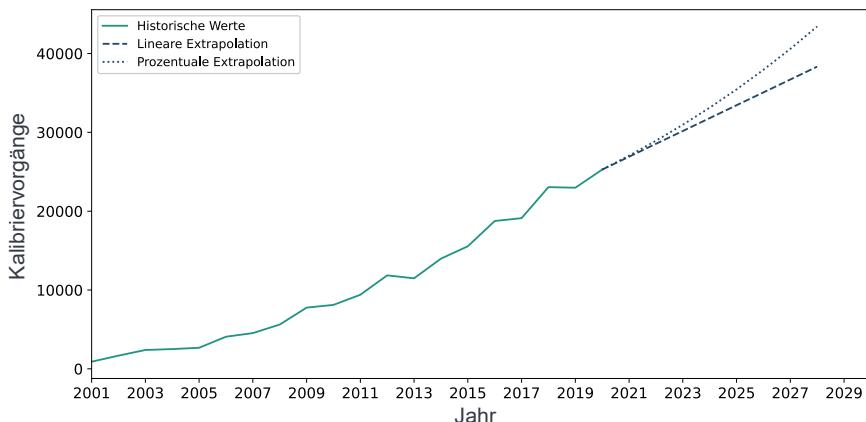


Abbildung 6-1: Nachfrageverlauf im ersten Anwendungsfall mit Extrapolationen

Für den Mittelständler stellt sich daher akut die Frage, wie in naher Zukunft die Nachfragerlast, die bis dahin weiter steigen könnte, möglichst mit den bestehenden Kapazitäten bedient werden kann.

### 6.1.1.1 Vorstellung der Produktvarianten

Im Fokus der Produktvorstellung stehen die Normgewichte. In Anhang A10 findet sich die Abbildung eines typischen Normgewichts. Die Norm beschreibt neun Toleranzklassen und dreißig Nominalwerte. Jedoch beinhaltet nicht jede Toleranzklasse alle Gewichte, sodass es insgesamt 201 zulässige Toleranz-Nominalgewicht-Paare gibt. Dies lässt sich auch der Tabelle in Anhang A11 entnehmen, die der Norm entstammt (OIML R 111-1, S. 24). Als weiterer Variantentreiber kann der zugrundeliegende Werkstoff angesehen werden. Je nach Toleranzklasse kommen nur bestimmte Werkstoffe aufgrund ihrer physikalischen und chemischen Eigenschaften für die Produktion der Normgewichte in Frage. Ebenfalls abhängig von der Toleranzklasse dürfen nur bestimmte Geometrien gefertigt werden (OIML R 111-1). Im Rahmen der Erprobung werden jedoch Werkstoff und Bauform vernachlässigt. Diese Vernachlässigung beeinflusst die Ergebnisse nur geringfügig, da die Prozesse, Preise und Kosten größtenteils toleranzklassenspezifisch sind.

Die Gewichte werden aktuell nach Listenpreisen verkauft. In Rücksprache mit den Experten aus der Industrie wird angenommen, dass zu jedem Einzelgewicht eine Deckungsbeitragsquote von ca. 15 % erwirtschaftet wird. Die Deckungsbeitragsquote ist nach Thommen & Achleitner et al. (2020, S. 194, 388) definiert. Durch die Deckungsbeitragsquote werden sämtliche variablen Kosten in der Fertigung adressiert. Tabelle 6-1 gibt eine Übersicht über die Verkaufspreise anhand der durchschnittlichen Einzelpreise nach Gewichtstyp.

*Tabelle 6-1: Überblick über durchschnittliche Verkaufspreise*

Gewichtstyp	Durchschnittlicher Verkaufspreis
E1	97 €
E2	34 €
F1	28 €
F2	46 €
M	39 €

### 6.1.1.2 Vorstellung des Kundenstamms

Bei der Modellierung der Kunden werden als Basis die acht Kundentypen zugrunde gelegt, die in Kapitel 5.2.1.1 dargelegt wurden. Zusätzlich liegen Informationen zur Zusammensetzung des Kundenstamms des betrachteten Mittelständlers vor, die berücksichtigt werden. Insgesamt hat der Mittelständler in den Jahren 2018 und 2019 ca. 600

verschiedene Kunden. Nach eigenen Angaben sind ca. 33,3 % dieser Kunden solche, die eine kurzfristige Präferenz haben, jedoch gleichzeitig bereit sind einen höheren Aufschlag für eine Expressanforderung zu zahlen. Die restlichen 66,6 % teilen sich auf in solche, die frühzeitig planen und das Gewicht zum geplanten Zeitpunkt exakt brauchen und solche, die keine engen Präferenzen zum Lieferzeitpunkt haben. Zu diesen 66,6 % gibt es keine gesicherten Informationen zur Preisbereitschaft, weshalb eine Gleichverteilung angenommen wird. Es ergeben sich die Kundentypen, wie in Tabelle 6-2 dargestellt.

*Tabelle 6-2: Überblick zu den Kundentypen*

Kunde	$a_{i,v}$	$b_{i,v}$	$d_{i,v}^u$	Zeitpräferenz	preissensitiv	Lieferzeit-sensitiv
1	0,9	2	1,5	kurzfristig		X
2	0,8	3	1,3	kurzfristig		X
3	0,6	5	0,8	langfristig		
4	0,5	7	0,6	langfristig	X	
5	0,4	9	0,6	langfristig	X	
6	0,7	3	0,9	kurzfristig		X
7	0,65	5	0,8	langfristig		
8	0,6	7	0,7	langfristig	X	

Auch die Nachfrage der Kunden lässt sich aus den historischen Daten bestimmen. Aus der Analyse ergab sich, dass über 83 % der Kunden entweder in einem einjährigen oder zweijährigen Rhythmus Gewichte zur Rekalibrierung bringen. Das Abweichungsfenster vom vorgegebenen Rhythmus lässt sich als Normalverteilung  $N$  mit einer Standardabweichung  $\sigma$  von 10 Tagen modellieren. Zum aktuellen Zeitpunkt können Kunden eine Expressbearbeitung aushandeln. Für die Expressbearbeitung werden üblicherweise keine Sondergebühren berechnet. Vielmehr ist das Aushandeln einer Expressbearbeitung eine Maßnahme zur Kundenbindung. Dieses Vorgehen ist ein Indiz, dass ein Status Quo mit flexibler Lieferzeit als Vergleichssystem der Praxis am nächsten kommt.

Wird ein Auftrag verspätet ausgeliefert, so greifen Verspätungskosten beispielsweise in Form von Vertragsstrafen. Solche Vertragsstrafen sind typischerweise von der Verzugszeit abhängig und durch einen Maximalwert beschränkt (Bohnstedt 2014, S. 161f). Eine generelle Höhe kann hingegen nicht pauschal festgelegt werden, da diese von

Fall zu Fall sehr verschieden sein kann. Für diesen Anwendungsfall liegen keine Informationen über Verspätungskosten vor. Es wird deshalb angenommen, dass die Vertragsstrafe durchschnittlich 25 % der Verkaufssumme pro Tag beträgt und bei 200 % das Maximum erreicht ist.

Die Bestimmung der Nachfragedaten basiert hierbei auf den vom Verfasser der Arbeit betreuten Abschlussarbeiten A\_Werle (2021) und A\_Lorenz (2021).

#### **6.1.1.3 Vorstellung des Produktionssystems am Standort**

Zur Kalibrierung von Gewichten werden nun die einzelnen Schritte erläutert. Eine Übersicht zu dem Prozess geben die Prozessdiagramme in Anhang A12 bis A16. Der Startpunkt des Prozesses ist der Eingang des Auftrags für ein zu (re-)kalibrierendes Gewicht. Das Gewicht muss sich zunächst in einer Akklimatisierungskammer an die Umgebungstemperatur anpassen. Dieser Schritt ist notwendig, weil der Wägewert des Gewichts von der Temperatur abhängt. Je nach Toleranzklasse muss das Gewicht länger akklimatisieren, um verarbeitet werden zu können. Nach der Akklimatisierung erfolgt eine Beschaffenheitsprüfung. Hierbei wird das Gewicht visuell auf Schäden und Verschleißerscheinungen untersucht. Als nächstes kann optional eine sog. „As-Found“-Kalibrierung erfolgen. Das bedeutet, dass Gewicht wird ohne Änderungen gemessen. Die Werte werden von einigen Kunden zur Dokumentation und Berechnung benötigt. Erst danach erfolgen die eigentliche Aufbereitung und Justage, sofern diese in der Beschaffenheitsprüfung oder der „As-Found“-Kalibrierung indiziert war. Muss das Gewicht nicht aufbereitet werden und war die „As-Found“-Kalibrierung erfolgreich, kann das Gewicht direkt zurückgesendet werden. Andernfalls erfolgt nun die Aufbereitung, Justage und erneute Kalibrierung. Ob ein Gewicht noch justiert und aufbereitet werden kann, ist abhängig von der Toleranzklasse. Bei der Aufbereitung wird Schmutz entfernt und es werden gegebenenfalls Lackschäden ausgebessert. Bei der Justage geht es zum einen um die magnetischen Eigenschaften und zum anderen um die Masse des Gewichts. Gewichte können sich über die Zeit magnetisieren oder durch mangelhafte Materialeigenschaften eine Magnetisierung aufweisen. Die Magnetisierbarkeit betrifft in den meisten Fällen die Gewichte der unteren Toleranzklassen. Im Fall unzulässiger magnetischer Eigenschaften wird versucht, das magnetische Feld des Gewichts zu entfernen, sodass es innerhalb der Toleranz liegt. Im zweiten Schritt wird versucht, die Masse des Gewichts im Bedarfsfall zu korrigieren. Hier lässt sich festhalten, dass die Masse durch Gebrauch des Gewichts üblicherweise abnimmt. Nur bei Gewichten der unteren

Klassen ist es möglich, durch Auffüllen einer Aussparung dem Gewicht wieder mehr Masse zu geben. Die Gewichte der höchsten Toleranzklasse müssen jedoch bei mangelnder Masse immer durch ein neues ersetzt werden, sofern sie die Norm erfüllen sollen. Verläuft die Aufbereitung, Justage und Rekalibrierung erfolgreich, so kann das Gewicht mit neuem Kalibrierschein zurückgesendet werden. Andernfalls ist mit dem Kunden abzustimmen, ob das alte Gewicht ohne Kalibrierschein zurückgesendet oder direkt durch ein neues, kalibriertes Gewicht ersetzt werden soll. Die Abstimmung mit Kunden geht nicht in die Lieferzeit mit ein. Zur Modellierung des Prozesses wurde für jede Toleranzklasse ein Prozessdiagramm entworfen (vgl. Anhang A12 bis A16). Mithilfe der einzelnen Prozessdiagramme und den Wahrscheinlichkeiten an den Übergängen können die Fertigungspläne für den zweiten Anwendungsfall erstellt werden. Ein Beispiel für solch einen Fertigungsplan findet sich im Anhang A17.

Zentral für den gesamten Prozess ist die Kalibrierung. Abhängig von der Toleranzklasse kommen verschiedene Stationen zum Einsatz. Insgesamt gibt es 32 Kalibrierstationen. Jede Station ist so individuell, dass sie ihren eigenen Stationstyp besitzt. Die Stationen unterscheiden sich in Ihrer Präzision und den Voraussetzungen. Einige Stationen prüfen Gewichte einzeln und müssen von Hand einzeln bestückt werden. Andere Stationen besitzen ein eigenes Magazin, aus dem die Station automatisiert bestückt werden kann. Sogenannte Komparator-Anlagen arbeiten mit vier oder sechs Gewichten auf einmal. Sie führen dann eine Dekadenmessung durch, bei der durch gegenseitiges Vergleichen die Gewichtsmasse mit sehr hoher Präzision bestimmt werden kann. Diese Nebenbedingungen gilt es in der Ablaufsteuerung für diesen Anwendungsfall individuell abzubilden. Die Kosten einer Station liegen zwischen ca. 25.000 € und 150.000 €. Für das Modell kommen zu den 32 Kalibrierstationen noch Akklimatisierung, Beschaffungsprüfungen, Aufbereitung, Justage und Versand als Stationen hinzu. Für das Modell werden also 37 Stationen in Summe abgebildet. Aus den heterogenen Stationstypen resultieren verschiedene Prozesszeiten, die in Anhang A18 exemplarisch für die am meisten gefertigte Toleranzklasse E2 zu finden sind. Im Anwendungsfall gilt ein Zwei-Schicht-Modell. Für etwaige Lagerhaltungskosten werden basierend auf den Angaben von Biedermann (2008, S. 43) 15 % pro Jahr bzw. 0,041 % pro Tag festgelegt.

Kosten neben den Stationskosten, Lagerhaltungskosten sowie variablen Kosten werden, sofern sie bereits in der Verkaufsmarge inkludiert sind, nicht explizit berechnet.

Die Erarbeitung des Produktionssystems basiert auf der vom Verfasser der Arbeit betreuten Abschlussarbeit A\_Werle (2021).

## 6.1.2 Exemplarische Anwendung und Ergebnisse

Nachdem der Anwendungsfall dargestellt wurde, kann die Anwendung des Lösungsmodells erfolgen. Wie im Kapitel 5.4 beschrieben, wird als Erstes eine Hyperparameteroptimierung durchgeführt, um eine Konfiguration zu finden. Mithilfe der aufgestellten Hypothesen wird dann das Lösungsmodell mit dem Vergleichssystem, bestehend aus Status Quo mit statischer Lieferzeit und Status Quo mit flexibler Lieferzeit, hinsichtlich des Potenzials verglichen (vgl. Tabelle 5-19).

### 6.1.2.1 Hyperparameteroptimierung

Wie in Kapitel 5.4.2 dargelegt, erfolgt die Hyperparameteroptimierung durch manuelle Suche, automatische Suche und manuelle Auswahl. Die manuelle Suche wurde bereits im Kapitel 5.4.2 allgemeingültig vorgenommen, sodass der Fokus dieses Kapitels auf der automatischen Suche und der nachgelagerten, manuellen Auswahl liegt.

Ziel ist es, eine Agentenkonfiguration zu finden, die möglichst gute Ergebnisse erzielt. Die Güte des Agenten wird bei den Versuchsdurchläufen am erzielten, durchschnittlichen Belohnungswert in der Bewertungsphase gemessen. Der erzielte, durchschnittliche Belohnungswert ergibt sich aus dem Mittelwert der erzielten Belohnungswerte über alle Szenarien und Wiederholungen. Für eine Wiederholung wird der Belohnungswert aus einer exponentiellen Glättung aller Belohnungswerte der Belohnungsfunktionen errechnet. Es werden pro Konfiguration drei Szenarien mit je zwei Wiederholungen geprüft. Auf diese Weise werden 60 Konfigurationen innerhalb des vorgegebenen Suchraums getestet. Es ergeben sich 360 Simulationsläufe. Sodann werden die drei besten Konfigurationen aus der automatischen Suche gegenübergestellt, hinsichtlich der definierten Zielgrößen miteinander verglichen und eine Konfiguration für die Prüfung der Hypothesen ausgewählt.

Für die automatische Suche gilt es zunächst, die für den Anwendungsfall spezifischen Steuer- und Rauschfaktoren zu bestimmen. Die spezifischen Steuerfaktoren beziehen sich auf das Verhältnis der verschiedenen Strafgewichtungen. Diese wurden zusammen mit dem Anwendungspartner festgelegt und sind in Tabelle 6-3 festgehalten.

Tabelle 6-3: Spezifische Steuerfaktoren im ersten Anwendungsfall

Name des konstant gehaltenen Parameters	Wert
Strafe für einen abgelehnten Auftrag $W_1$	30
Strafe für Verfrühung $W_2$	5
Strafe für geplante Verspätung $W_3$	75
Belohnung für Preisaufschlag $W_4$	$100 \cdot p_{rel,i,v}$
Strafe für realisierte Verspätung $W_5$	75
Belohnung für mögliche Kapazitätseinsparung $W_6$	5
Belohnung für realisierte Kapazitätseinsparung $W_7$	10

Für die konstant gehaltenen Rauschfaktoren müssen die Stationen und Kundentypen betrachtet werden. Wie aus der Vorstellung des Anwendungsfalls zu entnehmen ist, werden 37 Stationen mit jeweils eigenem Stationstyp untersucht. Die Kunden werden über die acht bekannten Typen repräsentiert und wurden bereits vorgestellt. Tabelle 6-4 fasst diese Rauschfaktoren zusammen.

Tabelle 6-4: Konstant gehaltene Rauschfaktoren im ersten Anwendungsfall

Name des konstant gehaltene Rauschfaktors	Wert
Anzahl Kunden(-typen) $I$	8
Anzahl der Stationstypen $MS_{typ}$	37
Anzahl Stationen je Stationstyp $O_{Styp}$	1

Für die zufallsbasierten Rauschfaktoren sind Fertigungszeiten sowie Bestellhäufigkeit und -menge zu bestimmen. Diese konnten aus den vorliegenden Auftragsdaten herausgearbeitet werden und ergeben sich aus Kapitel 6.1.1. Sie sind nochmal in Tabelle 6-5 übersichtlich zusammengefasst.

Tabelle 6-5: Zufallsbasierte Rauschfaktoren im ersten Anwendungsfall

Name des zufallsbasierten Rauschfaktors	Wert
Bestellhäufigkeit $BH_{i,j}$	$N(365; 10)$ oder $N(730; 10)$
Bestellmenge $BM_{i,j}$	Ein Gewichtssatz basierend auf den historischen Aufträgen
Fertigungszeiten $fz_{j,q,Styp}$	vgl. Anhang A18 $\pm 5\%$

Um repräsentative Ergebnisse zu erzielen, ist es notwendig eine Einschwingphase für die Simulation zu definieren (Gutenschwager & Rabe et al. 2017, S. 184-187). Andernfalls werden die Ergebnisse verzerrt, da zu Beginn das Produktionssystem komplett leer ist. Dies entspricht nicht dem vorgestellten Anwendungsfall, in dem kein Leerstand in der betrachteten Historie zu beobachten war. Außerdem benötigt der Agent bei einem zufallsbasierten Start Zeit, um die volle Funktionsfähigkeit über den Lernprozess zu erreichen. Die Bewertung des Agenten sollte also erst nach Abschließen einer Trainingsphase beginnen. Um eine Beeinflussung der Bewertungsphase durch Fehlentscheidungen aus der Lernphase zu vermeiden, werden vor der Bewertungsphase der Produktionsstandort und das Netzwerk nochmals komplett zurückgesetzt sowie neu eingeschwungen. Da im vorliegenden Anwendungsfall Aufträge auch im zweijährigen Rhythmus erteilt werden können, macht es Sinn, eine Dauer von mindestens zwei Jahren als Bewertungsphase einzustellen. Es ergibt sich folgender Ablauf für einen einzelnen Simulationslauf:

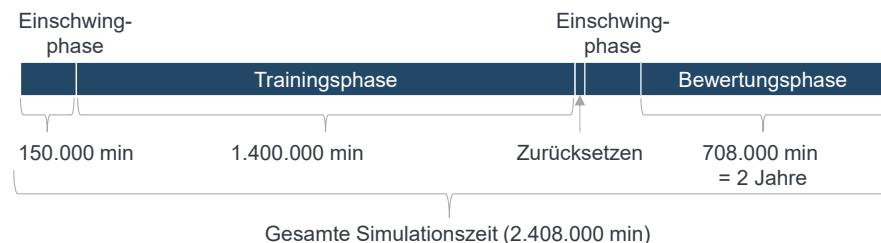


Abbildung 6-2: Übersicht über die Simulationsphasen im Anwendungsfall 1

Bei einer Parallelisierung auf acht Kernen benötigte die Suche für 60 Konfigurationen ungefähr 22 Stunden. Die Hyperparameteroptimierung basiert hierbei auf den vom Verfasser der Arbeit betreuten Abschlussarbeiten A\_Werle (2021) und A\_Henzi (2021). Die Konfiguration zur Durchführung der Zufallssuche ist Teil der frei verfügbaren Implementierung (Stamer 2021b).

Eine Übersicht über die erzielte Belohnungsverteilung der 60 Konfigurationen mit einer Hervorhebung der fünf besten Konfigurationen findet sich in der Abbildung 6-3.

Wie in Kapitel 5.4.2 beschrieben, werden aus der Zufallssuche mit 60 Konfigurationen die besten drei Konfigurationen ausgewählt und im Detail untersucht. Jede der drei Konfigurationen wird dann in jedem der drei Szenarien nochmals ausführlich getestet.

Hierbei wird jedes Simulationsexperiment zehnmal wiederholt, um eine bessere statistische Grundlage zu erhalten, als sie über die Zufallssuche rechenzeitbedingt möglich ist. Die Werte der zehn Wiederholungen werden gemittelt. Es ergeben sich folglich 90 Simulationsexperimente. Wichtig zu beachten ist, dass diese Versuche noch keine statistische Signifikanz aufweisen, sondern lediglich ein Indiz bieten, auf Basis dessen dann die Experimente für das Nachweisen der Hypothesen in Kapitel 6.1.2.2 erfolgen. Für die Bewertung der Ergebnisse wird auf die in Kapitel 5.1.1.1 eingeführten Zielgrößen zurückgegriffen.

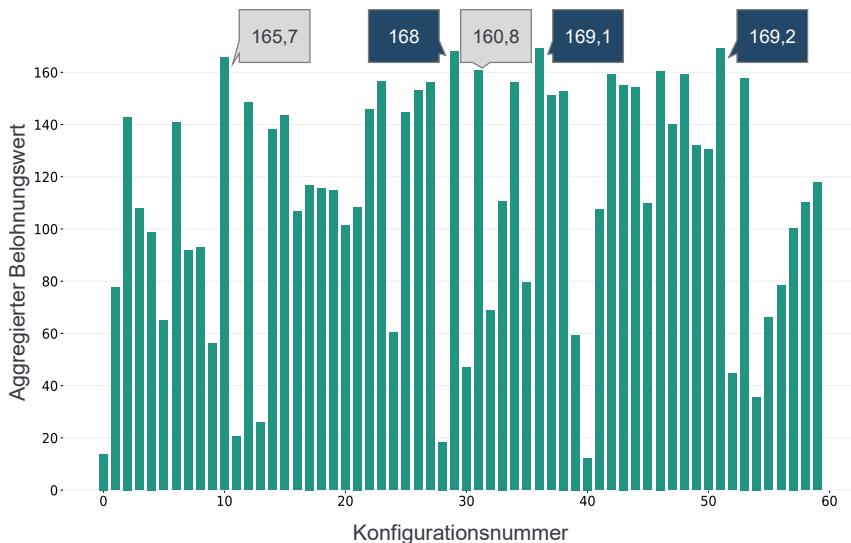


Abbildung 6-3: Verteilung der Belohnungswerte in der Hyperparameteroptimierung

Nachfolgend werden die Ergebnisse für jedes Szenario tabellarisch dargestellt. Es wird jeweils hervorgehoben, welche Konfiguration in welcher Zielgröße dominiert.

### Szenario 1: Statische, heterogene Kundenpräferenzen

Die Ergebnisse für das erste Szenario, in welchem heterogene Kunden mit statischen Präferenzen auftreten, ergeben sich wie folgt:

*Tabelle 6-6: Vergleich der Konfigurationen in Szenario 1 (Statische Präferenzen)*

	Konfiguration 1	Konfiguration 2	Konfiguration 3
Gewinn [€]	455627	311050	159086
Servicegrad [%]	57	77	94
Verspätungen [-]	2601,9	989,2	116,2
Entgangene Aufträge [-]	1416,7	2429,1	3986,3
Summe Liefertermin-abweichungen [min]	2,308941e+07	1,062651e+07	7,921745e+06

Aus der Auswertung zeigt sich, dass die Konfiguration 1 und 3 jeweils in den meisten Zielgrößen dominieren. Bei der zentralen Kennzahl Gewinn ist Konfiguration 1 an der Spitze, gefolgt von Konfiguration 2.

### **Szenario 2: Heterogene Kundenpräferenzen mit Drift zur Kurzfristigkeit**

Im Szenario mit Drift der Präferenzen zur Kurzfristigkeit stechen Konfiguration 2 und 3 hervor. Auffällig ist, dass die Verspätungen im Schnitt über die Szenarien nicht zugenommen haben. Der Servicegrad ist teilweise gestiegen. Alle Werte finden sich in der nachfolgenden Tabelle 6-7.

*Tabelle 6-7: Vergleich der Konfigurationen in Szenario 2 (Drift zu Kurzfristigkeit)*

	Konfiguration 1	Konfiguration 2	Konfiguration 3
Gewinn [€]	85464	247184	139104
Servicegrad [%]	0,79	0,82	0,89
Verspätungen [-]	378,6	723,9	309,7
Entgangene Aufträge [-]	5376,8	3046,5	3681,2
Summe Liefertermin-abweichungen [min]	3,874193e+06	8,117962e+06	7,506984e+06

### **Szenario 3: Heterogene Kunden mit Drift zu geringerer Preisbereitschaft**

Im dritten Szenario sind Konfiguration 1 und 2 dominant. Konfiguration 2 kann hier abermals den höchsten Gewinn erzielen.

Tabelle 6-8: Vergleich der Konfigurationen in Szenario 3 (Sinkende Preisbereitschaft)

	Konfiguration 1	Konfiguration 2	Konfiguration 3
Gewinn [€]	85215	204432	112257
Servicegrad [%]	0,88	0,84	0,93
Verspätungen [-]	124,2	592,8	184,9
Entgangene Aufträge [-]	5479,3	3442,3	3839,1
Summe Liefertermin-abweichungen [min]	4,525037e+06	7,985408e+06	8,527759e+06

Die Ergebnisse zeigen, dass die Konfigurationen jeweils nahe beieinander liegen. So kann jede Konfiguration sich zweimal hervortun. Jedoch hat einzig Konfiguration 2 in zwei Szenarien den höchsten Gewinn erzielt. Daher wird diese Konfiguration weiter verwendet. Die Konfiguration 2 ist Teil der frei verfügbaren Implementierung (Stamer 2021b).

### 6.1.2.2 Potenzial im Vergleich zum Status Quo

Mit der gefundenen Konfiguration aus dem vorherigen Kapitel soll nun die Steigerung des Unternehmensgewinns auf Signifikanz untersucht werden. Das Vorgehen basiert hierbei wesentlich auf den Ausführungen in Kapitel 2.4.2.

Für jedes Szenario werden 30 Zufallsversuche mit und ohne Agent durchgeführt, um den zentralen Grenzwertsatz zu erfüllen und einen t-Test durchführen zu können. Beide Simulationsläufe erhalten jeweils den gleichen Zufallsstartwert, um eine Vergleichbarkeit der Läufe sicher zu stellen. Es gilt dann zu prüfen, ob die Nullhypothese  $H_0$  verworfen werden kann, sodass der Erwartungswert der Gewinndifferenz  $\mu_{\Delta G}$  kleiner oder gleich null ist. Die Alternativhypothese  $H_1$  formuliert entsprechend das Gegenteil.

$$H_0: \mu_{\Delta G} \leq 0 \text{ und } H_1: \mu_{\Delta G} > 0$$

Formel 5-44

Die Potenzialermittlung basiert auf der vom Verfasser der Arbeit betreuten Abschlussarbeit A\_Werle (2021).

### Szenario 1: Statische, heterogene Kundenpräferenzen

Für das erste Szenario ergeben sich die Gewinnwerte von Status Quo mit statischer Lieferzeit und Agent wie folgt:

*Tabelle 6-9: Ergebnisse – statische Lieferzeit – Szenario 1 – Anwendungsfall 1*

	$\bar{G}$ [€]	$\sigma_g$ [€]	min [€]	max [€]
Agent	306.661,19	21.926,14	246.564,19	344.748,67
Statische Lieferzeit	111.861,15	558,00	109.714,11	112.507,29

Basierend auf diesen Werten kann der Hypothesentest wie folgt durchgeführt werden:

*Tabelle 6-10: Hypothesentest geg. statische Lieferzeit – Szenario 1 – Anwendungsfall 1*

	$\bar{G}$ [€]	$\hat{\sigma}_{\Delta G}$ [€]	$n_{stichprobe}$	$T^*$	$t_{29,95\%}$	$H_0$
Differenzwerte	194.800,04	22.002,97	30	48,49	1,70	abgelehnt

Die Nullhypothese kann folglich verworfen werden, weil  $T^* > t_{29,95\%}$  gilt. Stattdessen wird die Alternativhypothese angenommen, dass der Gewinn signifikant durch den Einsatz des Lieferzeit-Preisgestaltungsansatzes gesteigert werden kann. Das Konfidenzintervall beträgt hierbei [186.583; 203.017]. Wie aus Anhang A19 ersichtlich, ist der Servicegrad des Agenten hoch, jedoch ist er im Status Quo mit statischer Lieferzeit nochmals im Mittel besser. Gleiches gilt für die Verspätungen. Der Agent konnte mehr Aufträge gewinnen als der Status Quo.

Für die Simulationsläufe für den Vergleich von Status Quo mit flexibler Lieferzeit mit dem Agenten werden folgende Gewinnwerte erreicht:

*Tabelle 6-11: Ergebnisse – flexible Lieferzeit – Szenario 1 – Anwendungsfall 1*

	$\bar{G}$ [€]	$\sigma_g$ [€]	min [€]	max [€]
Agent	258.448,71	59.819,19	56.177,91	324.099,76
Flexible Lieferzeit	82.125,79	98.976,16	-178.261,87	176.991,04

Es lässt sich wiederum der Hypothesentest durchführen:

*Tabelle 6-12: Hypothesentest geg. flexible Lieferzeit – Szenario 1 – Anwendungsfall 1*

	$\bar{A}G$ [€]	$\hat{\sigma}_{\Delta G}$ [€]	$n_{stichprobe}$	$T^*$	$t_{29,95\%}$	$H_0$
Differenzwerte	176.322,92	96.618,85	30	10,00	1,70	abgelehnt

Auch hier wird die Nullhypothese verworfen und der Agent kann sich gegen den Status Quo durchsetzen. Das Konfidenzintervall ist durch den Bereich [140.244, 212.401] definiert. Anhang A20 zeigt die weiteren Kenngrößen. Der Agent konnte hier einen deutlich besseren Servicegrad erreichen und auch absolut weniger Verspätungen erzeugen. Dafür wurden insgesamt weniger Aufträge angenommen.

### **Szenario 2: Heterogene Kunden mit Drift zur Kurzfristigkeit**

Das Vorgehen läuft nun analog zum ersten Szenario startend mit den Simulationsergebnissen für den Gewinn von Agent und Status Quo mit statischer Lieferzeit ab:

*Tabelle 6-13: Ergebnisse – statische Lieferzeit – Szenario 2 – Anwendungsfall 1*

	$\bar{G}$ [€]	$\sigma_g$ [€]	min [€]	max [€]
Agent	194.988,34	86.272,54	35.706,28	338.031,01
Statische Lieferzeit	138.900,84	824,86	136.601,62	140.104,90

Der Hypothesentest ergibt sich wie folgt:

*Tabelle 6-14: Hypothesentest geg. statische Lieferzeit – Szenario 2 – Anwendungsfall 1*

	$\bar{A}G$ [€]	$\hat{\sigma}_{\Delta G}$ [€]	$n_{stichprobe}$	$T^*$	$t_{29,95\%}$	$H_0$
Differenzwerte	56.087,49	86.624,22	30	3,55	1,70	abgelehnt

Der Agent konnte sich auch im Szenario 2 durchsetzen. Das Konfidenzintervall beträgt [23.741; 88.434]. Die übrigen statistischen Werte finden sich in Anhang A21. Es zeigt sich, dass der Agent diesmal beim Servicegrad gleichauf mit dem Status Quo liegt. Teilweise kann er diesen sogar übertreffen. Dafür liegt der Agent bei den entgangenen Aufträgen hinten.

Für den Simulationslauf des Status Quo mit flexibler Lieferzeit ergeben sich folgende Gewinnwerte:

*Tabelle 6-15: Ergebnisse – flexible Lieferzeit – Szenario 2 – Anwendungsfall 1*

	$\bar{G}$ [€]	$\sigma_g$ [€]	min [€]	max [€]
Agent	204.609,30	90.047,59	30.655,84	349.135,75
Flexible Lieferzeit	117.381,76	70.564,66	-85.375,61	194.036,42

Der zugehörige Hypothesentest ist in Tabelle 6-16 zu sehen.

*Tabelle 6-16: Hypothesentest geg. flexible Lieferzeit – Szenario 2 – Anwendungsfall 1*

	$\bar{D}\bar{G}$ [€]	$\hat{\sigma}_{DG}$ [€]	$n_{stichprobe}$	$T^*$	$t_{29,95\%}$	$H_0$
Differenzwerte	87.227,54	120.415,76	30	3,97	1,70	abgelehnt

Die Nullhypothese wird wiederum verworfen. Das Konfidenzintervall liegt im Bereich [42.263; 132.192]. Die weiteren Kenngrößen sind in Anhang A22 aufgeführt. Die sich darstellende Situation ist hierbei vergleichbar mit Szenario 1, in welchem die Kundenpräferenzen statisch sind.

### **Szenario 3: Heterogene Kunden mit Drift zu geringerer Preisbereitschaft**

Im dritten und letzten Szenario wurden folgende Gewinnwerte im Vergleich mit dem Status Quo mit statischer Lieferzeit erreicht:

*Tabelle 6-17: Ergebnisse – statische Lieferzeit – Szenario 3 – Anwendungsfall 1*

	$\bar{G}$ [€]	$\sigma_g$ [€]	min [€]	max [€]
Agent	201.815,9	72.171,97	20.108,61	305.769,81
Statische Lieferzeit	111.825,42	580,47	109.544,95	112.624,44

Mit diesen Werten stellt sich der Hypothesentest gemäß Tabelle 6-18 dar.

*Tabelle 6-18: Hypothesentest geg. statische Lieferzeit – Szenario 3 – Anwendungsfall 1*

	$\bar{D}\bar{G}$ [€]	$\hat{\sigma}_{DG}$ [€]	$n_{stichprobe}$	$T^*$	$t_{29,95\%}$	$H_0$
Differenzwerte	89.990,48	72.211,23	30	6,83	1,70	abgelehnt

Somit konnte der Agent auch im dritten Szenario einen statistisch signifikant höheren Gewinn mit einem Konfidenzintervall [63.026; 116.955] erzielen. Anhang A23 zeigt die

weiteren Kenngrößen. Hier schneidet der Agent sowohl in den Verspätungen als auch im Servicegrad schlechter ab. Jedoch ist die Lieferterminabweichung in Minuten geringer als im Vergleichssystem.

Für den Vergleich mit dem Status Quo mit flexibler Lieferzeit werden folgende Gewinnwerte erzielt.

*Tabelle 6-19: Ergebnisse – flexible Lieferzeit – Szenario 3 – Anwendungsfall 1*

	$\bar{G}$ [€]	$\sigma_g$ [€]	min [€]	max [€]
Agent	188.302,03	83.899,12	12.841,27	289.759,84
Flexible Lieferzeit	109.244,12	69.184,08	-131.519,23	195.901,14

Tabelle 6-20 zeigt den zugehörigen Hypothesentest.

*Tabelle 6-20: Hypothesentest geg. flexible Lieferzeit – Szenario 3 – Anwendungsfall 1*

	$\bar{A}\bar{G}$ [€]	$\hat{\sigma}_{\Delta G}$ [€]	$n_{stichprobe}$	$T^*$	$t_{29,95\%}$	$H_0$
Differenzwerte	79.057,90	75.862,80	30	5,71	1,70	abgelehnt

Es zeigt sich, dass der Agent in allen drei Szenarien und mit allen Vergleichssystemen besser abschneidet. Das Konfidenzintervall beträgt hier [50.730; 107.386]. Die weiteren Kenngrößen sind in Anhang A24 dargestellt. Der Agent schneidet hier deutlich besser ab als das Vergleichssystem, außer bei den entgangenen Aufträgen, wie in den anderen drei Szenarien auch. Auffällig ist bei diesem Szenario die deutlich größere Standardabweichung im Servicegrad. Diese deutet auf eine gewisse Instabilität der Simulationsläufe hin.

## Zusammenfassung

Aus den Versuchsläufen ergibt sich die folgende Übersicht, welche zeigt, in welchen Fällen der Agent besser und in welchen er schlechter war als das jeweilige Vergleichssystem.

*Tabelle 6-21: Darstellung der Verbesserung zum Status Quo durch Mittelwertvergleich*

Legende: Agent ist...	Szenario 1		Szenario 2		Szenario 3	
	Statische Lieferzeit	Flexible Lieferzeit	Statische Lieferzeit	Flexible Lieferzeit	Statische Lieferzeit	Flexible Lieferzeit
besser	✓	✓	✓	✓	✓	✓
gleich	—	✓	—	✓	—	✓
schlechter als Vergleichssystem	✗	—	✗	—	✗	—
Gewinn	✓	✓	✓	✓	✓	✓
Servicegrad	✗	✓	✗	✓	✗	✓
Verspätungen	✗	✓	✓	✓	✗	✓
Entgangene Aufträge	✓	✗	✗	✗	✗	✗
Liefertermin-abweichungen	✗	✓	✓	✓	✓	✓

Es zeigt sich insgesamt ein gemischtes Bild. Da der Agent in der Lage ist, stets einen höheren Gewinn zu erzielen, funktioniert die Belohnungsfunktion korrekt. Trotz hohem Gewinn und Servicegrad kann der Agent sich jedoch nicht immer in allen Punkten gegen das Vergleichssystem durchsetzen. Auffällig ist die Anzahl entgangener Aufträge bei gleichzeitig hohem Gewinn. Diese deutet darauf hin, dass der Agent in der Lage ist, Kunden mit hoher Preisbereitschaft und passender Lieferzeitpräferenz zu bevorzugen. Bisher nicht explizit betrachtet wurde das Potenzial zur Kapazitätsnivellierung, welches ebenfalls in der Einleitung motiviert wurde. Ein Teil der Hypothese aus der Einleitung besagt, dass dieselbe Menge an Aufträgen durch die geschickte Verhandlung der Lieferzeiten mit weniger Kapazitäten erfüllt werden kann, ohne den Servicegrad wesentlich zu senken. Diese Untersuchung soll Schwerpunkt des nachfolgenden Kapitels sein.

### 6.1.2.3 Erweiterung der Betrachtung auf Kapazitätsentscheidungen

Im Rahmen einer erweiterten Betrachtung soll nun untersucht werden, inwiefern der Ansatz in der Lage ist, Kapazitäten einzusparen bzw. mit Kapazitätsänderungen umzugehen, und wo sich Kippunkte ergeben. Hierbei zeigt sich ein großer Vorteil des Lösungsmodells. Es bietet die Möglichkeit, den variierbaren Betrachtungsraum auf die Nachfrage und die eingesetzten Kapazitäten auszudehnen, um somit Gewinnschwellen zwischen verschiedenen Standortkonfigurationen zu untersuchen. Die Ergebnisse lassen sich zudem dem Status Quo mit statischer Lieferzeit zum Vergleich gegenüberstellen.

Für die Variierung der Nachfrage wird ein Korridor aufgespannt, welcher vom letzten Datenpunkt 2020 ausgehend eine Steigerung um 100 % bzw. ein Sinken um 50 % bis 2025 abbildet (vgl. Abbildung 6-4).

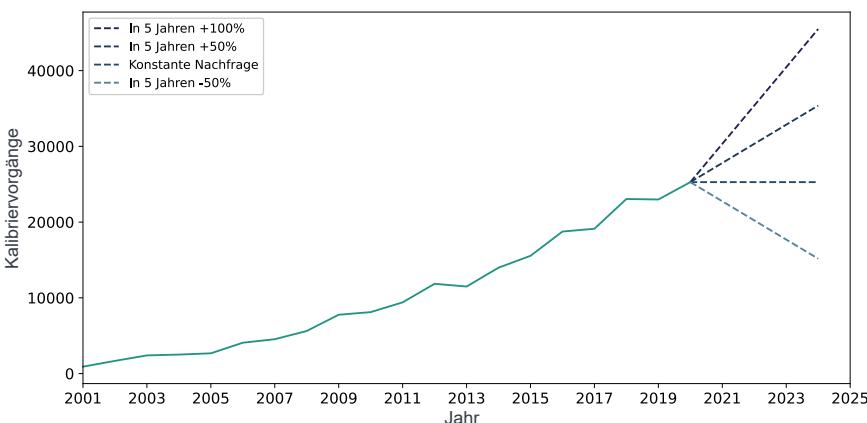


Abbildung 6-4: Darstellung verschiedener Nachfrageverläufe

Für die Kapazitätsänderung werden die aktuelle Situation, das Hinzunehmen einer Station und das Wegnehmen einer Station verglichen. Hierbei ist relevant, welcher Stationstyp in der Menge variiert wird. Die grundlegende Idee ist einerseits, einen Stationstyp zu wählen, welcher im Standort häufig als Flaschenhals auftritt. Andererseits darf eine Wegnahme des Stationstyps nicht zur Unmöglichkeit der Produktion bestimmter Gewichte führen. Als anschauliches Beispiel sei hier die Wegnahme der Akklimatisierung genannt, wodurch gar keine Produkte mehr bearbeitet werden könnten.

Es gibt verschiedene Möglichkeiten einen Engpass in der Produktion zu identifizieren. Nach Bertagnolli (2018, S. 70) besteht ein Engpass, wo sich Teile häufen. Daher wird eine Analyse der durchschnittlichen Warteschlange an den Stationen bei einer Steigerung der Nachfrage um 50 % durchgeführt. Diese ergab folgende Flaschenhälse:

*Tabelle 6-22: Übersicht über die durchschnittliche Warteschlangenlänge pro Station im Stresstest*

Stationsnummer	Durchschnittliche Warteschlangenlänge [Anzahl Aufträge]
9	71,76
17	55,70
7	42,46
22	26,91
24	23,56

Die Nummerierung folgt hierbei den Stationsnummern in Anhang A18. Der Auswertung folgend wird die Station 9 als stärkster Flaschenhals dupliziert bzw. entfernt. Die Funktion dieser Station ist durch andere Stationen substituierbar, sodass prinzipiell alle Produkte weiterhin gefertigt werden können. Des Weiteren werden auch die Preisbereitschaft und die Lieferzeitpräferenz analog zu den Szenarien im vorhergehenden Kapitel variiert. Statt einer dynamischen Änderung werden die Präferenzen jedoch statisch zu Beginn der Simulation festgelegt. Dies erleichtert die Interpretation der grafischen Darstellungen von Ergebnissen. Eine weitere Vereinfachung wird getroffen, indem jeweils nur die Preisbereitschaft oder die Lieferzeitpräferenz ausgehend vom Ist-Zustand variiert wird. Hierdurch müssen weniger Versuchskombinationen ausgeführt werden. Diese Erleichterung wird in der Annahme getroffen, dass sich die Preisbereitschaft hauptsächlich im Gewinn und die Lieferzeitpräferenz hauptsächlich im Servicegrad niederschlagen. Die Effekte werden daher isoliert voneinander betrachtet. Ein weiterer Vorteil dieses Vorgehens ist die Möglichkeit, die Zusammenhänge als Ebenen im dreidimensionalen Raum darstellen zu können. Sowohl Preisbereitschaft als auch Lieferzeitpräferenz werden von -20 % bis +20 % variiert. Es ergeben sich folgende Parametervariationen:

Tabelle 6-23: Überblick über die Parametervariationen in der erweiterten Betrachtung

Parameter	Zu prüfende Ausprägungen
Nachfrage	{+100 %, +50 %, +0 %, -50 %} ausgehend vom Status Quo
Kapazitäten	{+1, +0, -1} von Station(-styp) 9
Preisbereitschaft	{+20 %, +10 %, 0 %, -10 %, -20 %} ausgehend vom Status Quo
Lieferzeitpräferenz	{+20 %, +10 %, 0 %, -10 %, -20 %} ausgehend vom Status Quo

Um eine gewisse Sicherheit gegen Ausreißer zu erzielen, wird jeder Versuch zehnmal wiederholt. Es ergibt sich somit die folgende Anzahl an Simulationsläufen aus den 4 Nachfrageleveln, 3 Kapazitätsniveaus, 5 Preis- bzw. Lieferzeitpräferenzen und 10 Wiederholungen:

$$(4 \cdot 3 \cdot 5) \cdot 2 \cdot 10 = 1200 \text{ Simulationsläufe} \quad \text{Formel 6-1}$$

Aus den Ergebnissen lassen sich nun verschiedene Zusammenhänge aufzeigen. Eine Betrachtung der Abhängigkeit zwischen Gewinndifferenz von Lösungsmodell und Status Quo mit statischer Lieferzeit sowie Preisbereitschaft und Nachfragefaktor zeigt, dass das Lösungsmodell mit zunehmender Nachfrage besser abschneidet als der Status Quo (vgl. Abbildung 6-5).

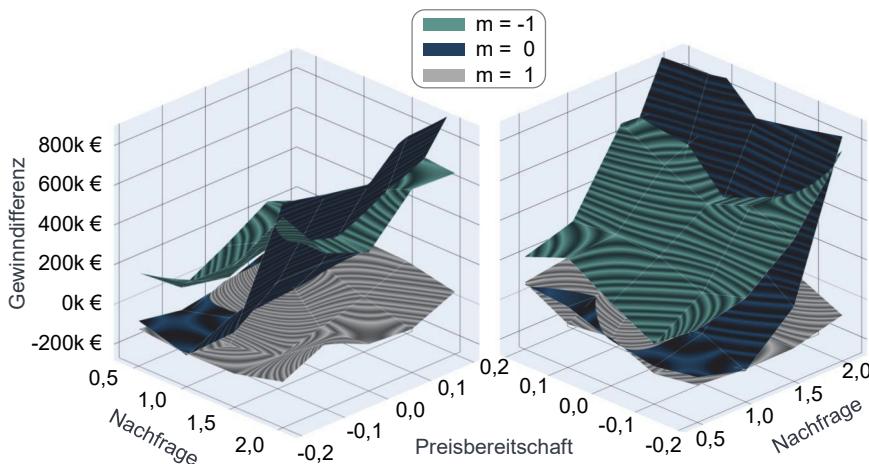


Abbildung 6-5: Darstellung der Gewinndifferenzebenen zwischen Agent und Status Quo mit statischer Lieferzeit

Dies ist besonders ausgeprägt bei einer Kapazitätseinschränkung ( $m = -1$ ) und im aktuellen Kapazitätsangebot ( $m = 0$ ), wohingegen bei einer Kapazitätserhöhung ( $m = 1$ ) sich keine ausgeprägte Differenz zeigt. Es wird deutlich, dass der Ansatz in der Lage ist, mit einem beschränkten Kapazitätsangebot bei steigender Nachfrage besser umzugehen als das Vergleichssystem. Ein positiver Zusammenhang zwischen Preisbereitschaft und Gewinndifferenz ist hingegen weniger stark ausgeprägt. Dennoch lässt sich auch hier eine leichte Steigung erkennen. Dies deutet abermals darauf hin, dass der Agent in der Lage ist, Preisbereitschaften besser auszuschöpfen als das Vergleichssystem.

Zusätzlich zu den Gewinndifferenzebenen lässt sich auch der vom Agenten erzielte, absolute Gewinn in Ebenen darstellen (Abbildung 6-6). Hier sei nochmals darauf verwiesen, dass dieser absolute Gewinn nicht den realen Unternehmensgewinn angibt, weil wichtige Positionen, wie zum Beispiel Versicherungen und Kreditzinsen, im Rahmen der Modellierung vernachlässigt wurden (vgl. Kapitel 5.1.1.1). Zudem wurde in dieser Arbeit die Gewinndifferenz als zentrale Entscheidungsgröße festgelegt, wodurch der absolute Gewinn von nachrangiger Ordnung ist. Dennoch ist die Darstellung dieser Ebenen von großem Wert, weil sie Kippunkte aufzeigen kann, bei denen sich eine Kapazitätsänderung für ein Unternehmen lohnt. Es zeigt somit der vielschichtige Mehrwert der Arbeit.

Für die Kapazitätsänderung wurde in Rücksprache mit dem Anwendungspartner angenommen, dass eine neue Kapazität des gegebenen Stationstyps 9 ungefähr 130.000 € inklusive Instandhaltung kostet (vgl. auch Kapitel 6.1.1.3). Der Verkauf einer Station wurde mit einer Einnahme von 65.000 € inklusive der Ersparnis der Instandhaltung angenommen. Die Gewinnebenen sind in Abbildung 6-6 links zu sehen. Zur besseren Sichtbarkeit der Schnittpunkte wurde die Perspektive entsprechend eingestellt.

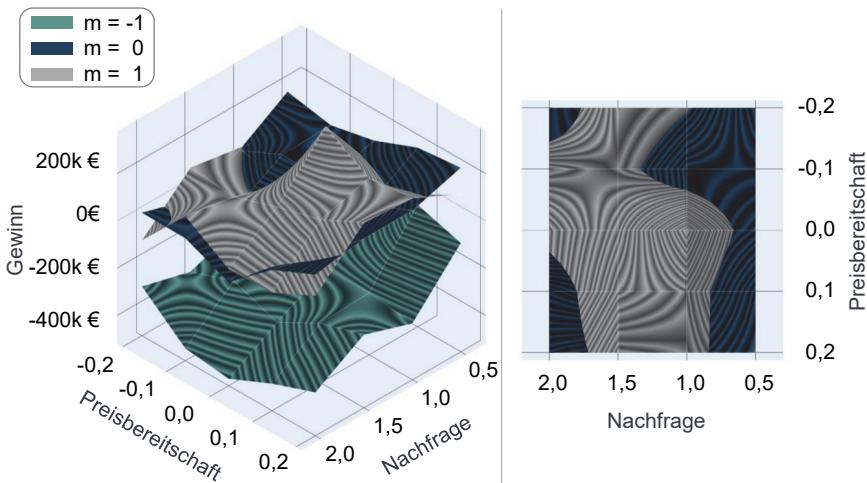


Abbildung 6-6: Darstellung der Gewinnebenen des Agenten

Wie zu sehen ist, kann sich eine Kapazitätserweiterung bereits im aktuellen Zustand ab einer Nachfrage von 100 % (1,0) und einer ursprünglichen Preisbereitschaft (0,0) lohnen. Um diesen Verlauf der beiden Gewinnebenen bei ( $m = 0$ ) und ( $m = 1$ ) besser zu visualisieren, findet sich in der Abbildung rechtsseitig noch eine Draufsicht. Dort ist zu sehen, wie es zu einer Instabilität bei 200% Nachfrage kommt, sodass dort teilweise die Ebene ( $m = 0$ ) wieder ein höheres Gewinnniveau erreicht. Für die Kapazitätssenkung lässt sich sagen, dass diese streng dominant schlechter in den betrachteten Betriebspunkten ist und daher vernachlässigt werden kann.

Zusätzlich zum Gewinn lässt sich der Servicegrad analysieren. Hierzu werden zuerst die Servicegrad-Differenzebenen betrachtet (Vgl. Abbildung 6-7). Wie bei den Gewinn-differenzebenen zeigt sich auch hier eine starke Zunahme der Servicegraddifferenz bei steigender Nachfrage. Diese Zunahme ist, wie zu erwarten, bei eingeschränkten Kapazitäten ausgeprägter. Interessant sind hier zwei Artefakte: Das erste Artefakt liegt bei  $m = -1$ , einer Nachfrage von -50 % und einer Lieferzeitpräferenz von +20 %. Dieser Betriebspunkt sollte vergleichsweise leicht zu handhaben sein. Dennoch schneidet der Agent deutlich schlechter ab als das Vergleichssystem. Das zweite Artefakt liegt bei  $m = 0$ , einer Nachfrage von 200 % und einer Lieferzeitpräferenz von +20 %. Auch hier schneidet der Agent deutlich schlechter als das Vergleichssystem ab.

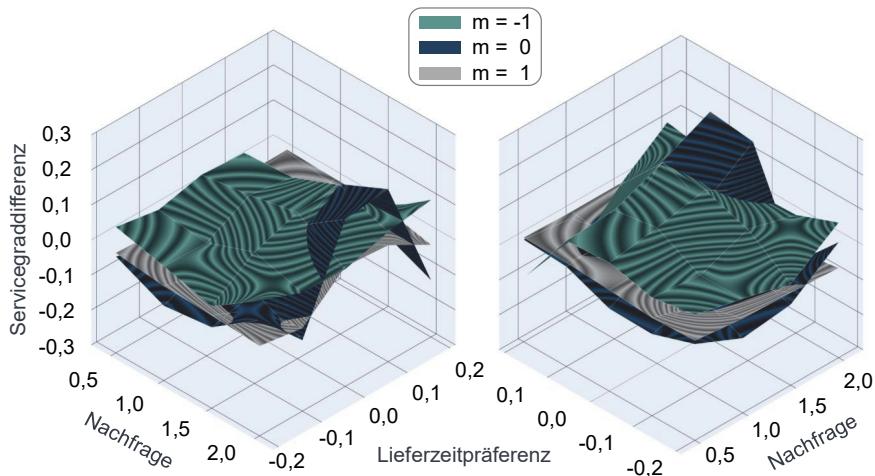


Abbildung 6-7: Darstellung der Servicegrad-Differenzebenen zwischen Agent und Status Quo mit statischer Lieferzeit

Eine Betrachtung der vom Agenten erzielten, absoluten Servicegradebenen (vgl. Abbildung 6-8) zeigt die zuvor beschriebenen Ausreißer nicht. Daher kann festgestellt werden, dass in dem Betriebspunkt tatsächlich das Vergleichssystem besser abschneidet.

Im Übrigen zeigen die Servicegradebenen den zu erwartenden Verlauf: Mit mehr Kapazitäten steigt der Servicegrad, mit zunehmender Nachfrage nimmt er ab. Interessant zu sehen ist, dass die Ebenen über die Lieferzeitpräferenz verhältnismäßig stabil bleiben, was für die Fähigkeit zur Auftragsnivellierung spricht. Diese Fähigkeit nimmt mit zunehmender Nachfrage und sinkenden Kapazitäten ab, weil das Produktionssystem immer weniger Lücken in der Ablaufplanung hat, in die neue Aufträge eingeschoben werden können.

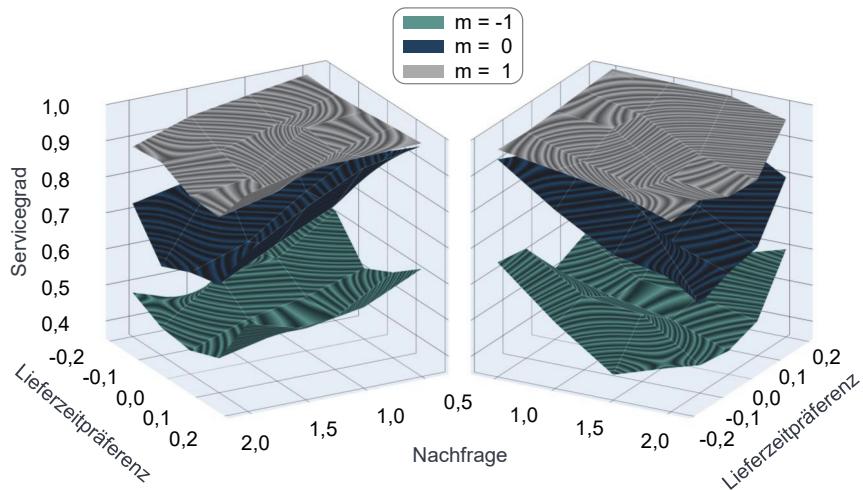


Abbildung 6-8: Darstellung der Servicegrad-Ebenen des Agenten

## 6.2 Anwendung im Bereich der Pneumatikzylinderproduktion

Der zweite Anwendungsfall dient vor allem dem Nachweis der allgemeinen Anwendbarkeit des Ansatzes. Um diesen Zweck zu erfüllen, wurde als starker Kontrast zum ersten Anwendungsfall ein repräsentatives, weltweit agierendes Unternehmen aus der Handhabungstechnik mit mehreren Produktionsstandorten gewählt. Demgemäß werden auch mehrere Standorte berücksichtigt. Einige Unternehmen aus der Handhabungstechnik besitzen Online-Plattformen für den Direktvertrieb ihrer Produkte, die hypothetisch als infrastrukturelle Grundlage für eine interaktive dynamische Lieferzeit-Preisgestaltung dienen könnten (Kollmann 2020, S. 173). Der Anwendungsfall bietet daher auch aus Sicht der Praxisnähe die ideale Ausgangslage zur Erprobung des Ansatzes.

Der zweite Anwendungsfall basiert auf den vom Verfasser der Arbeit betreuten Abschlussarbeiten A\_Ochs (2021) und A\_Henzi (2021).

## 6.2.1 Vorstellung des Anwendungsfalls

Im Fokus dieses Anwendungsfalls steht ein weltweit agierendes Unternehmen, mit über 20.000 Mitarbeitenden, über 200 Niederlassungen und über 10 Produktionsstandorten. Letztere befinden sich unter anderem in Deutschland, China und den USA, welche die Kernabsatzmärkte darstellen. Diese drei Standorte werden daher im Rahmen der Erprobung fokussiert. Zu den Kernprodukten des betrachteten Unternehmens zählen verschiedene pneumatische und elektrische Antriebslösungen, aus denen sich ein Produktkatalog von über 30.000 Produkten ergibt, wovon 2200 jeden Tag gefertigt werden. Alle Produkte unterteilen sich wiederum in eine Vielzahl von Varianten. Mit den Produkten werden Umsätze pro Jahr in Höhe von über 3 Mrd. € erzielt. Bei der Produktion von Handhabungstechnik ist die Zerspanung der zentral kritische Produktionsschritt, der sich wesentlich auf die Kosten und den Gewinn niederschlägt. Daher wird in diesem Anwendungsfall die Zerspanung als einstufiger Prozess modelliert.

Zur weiteren Eingrenzung wird ein Pneumatikzylinder als Beispielprodukt festgelegt. Es handelt sich hierbei um ein repräsentatives Produkt aus dem Handhabungsbereich. Anhand dieses Beispielprodukts wird nun der Anwendungsfall weiter konkretisiert.

### 6.2.1.1 Vorstellung der Produktvarianten

Der Pneumatikzylinder weist ein breites Einsatzspektrum vor. Sein heterogenes Anwendungsgebiet zeigt sich exemplarisch an seinem Einsatz im Bereich der Batterieherstellung, der Textilindustrie und sogar im Bereich der Produktion von Notfall-Beatmungsgeräten<sup>5</sup>. Das Einsatzspektrum wird unter anderem durch eine große Variantenvielfalt ermöglicht. Die Variantentreiber sind unter anderem der Kolbenhub und der Kolbendurchmesser. Hieraus allein lassen sich bereits rund neunzig relevante Varianten bilden, die im Rahmen der Erprobung betrachtet werden<sup>6</sup>. Darüber hinaus kann es weitere variable Produktmerkmale geben, wie zum Beispiel eine besondere Wärmebeständigkeit. Diese werden jedoch nicht betrachtet, da sie nicht direkt zerspanungsrelevant sind. Ein exemplarischer Pneumatikzylinder ist in Anhang A25 zu sehen.

---

<sup>5</sup> Vgl. <https://www.textilegence.com/en/festo-automation-digitalization-itma/> [24.04.2022]

<sup>6</sup> Vgl. <https://www.landefeld.de/gruppe/de/kompaktzylinder-adn/FTO-536374> [24.04.2022]

Da in diesem Anwendungsfall keine verkaufsrelevanten Daten vorlagen, mussten diese aus öffentlichen Daten abgeleitet werden. In Anhang A26 findet sich eine mathematische Herleitung der verkaufsrelevanten Daten. Die Ergebnisse sind in der nachfolgenden Tabelle 6-24 zusammengefasst.

*Tabelle 6-24: Verkaufspreis und -volumen des Pneumatikzylinders per Durchmesser*

Kolbendurchmesser in mm	Ø12	Ø16	Ø20	Ø25	Ø32	Ø40	Ø50	Ø63	Ø80	Ø100
Ø Verkaufspreis in €	40,08	41,97	44,36	51,24	59,91	77,51	89,78	111,21	140,97	176,84
Verkaufsvolumen pro Jahr in Stück	3.778	7.568	11.346	15.124	18.914	18.914	15.124	11.346	7.568	3.778
Anzahl Produktvarianten	7	8	9	9	10	10	10	9	9	9
Verkaufsvolumen pro Monat pro Produktvariante in Stück	45	79	105	140	158	158	126	105	70	35

### 6.2.1.2 Herleitung des Kundenstamms

Zur Anwendung des Ansatzes wird angenommen, dass die Kunden sich in einen der acht Kundentypen gemäß Kapitel 5.2.1.1 einteilen. Die konkreten Werte basieren auf Expertenbefragungen. Die Lieferzeitpräferenzen ließen sich hierbei konkret aus historischen Bestellungen ableiten. Die Preisbereitschaft und Lieferzeitkorridore hingegen musste aus dem Direktvertrieb geschätzt werden. Die Häufigkeit der acht Kundentypen wurde gleichverteilt angenommen. Die konkreten Werte für die acht Kundentypen finden sich in Tabelle 6-25.

*Tabelle 6-25: Kundentypen des Anwendungsfalls*

Kunde	$a_{i,v}$	$b_{i,v}$	$d_{i,v}^u$	Zeitpräferenz	preissensitiv	Lieferzeit-sensitiv
1	1,0	5,0	1,5	kurzfristig	X	
2	0,8	3,0	0,5	kurzfristig	X	X
3	0,7	15,0	1,4	langfristig	X	
4	1,0	17,0	0,6	langfristig	X	X
5	0,2	8,0	1,0	kurzfristig		
6	0,1	3,0	0,5	kurzfristig		X
7	0,05	15,0	1,4	langfristig		
8	0,1	14,0	0,6	langfristig		X

Die Kunden bestellen im Mittelwert alle 1,5 Monate, wobei die zugrundeliegende Verteilung unbekannt ist und somit als Exponentialverteilung angenommen wird (vgl. Kapitel 5.2.1.2). Bezüglich der Bestellmenge liegt nur das Gesamtvolumen vor, sodass die Menge pro Bestellung geschätzt werden muss. Die Bestellmenge wird entsprechend als Poissonverteilung modelliert und ergibt sich aus der Kombination von Verkaufsvolumina und Nachfragefrequenz. Die Herleitung findet sich in Anhang A27. Für normale Kunden läuft die Bestellung über eine Plattform automatisch und die Lieferzeit wird statisch vorgegeben. Es wird ein Lieferversprechen von sieben Tagen angestrebt, wobei eine Lieferung in unter drei Tagen im Standardgeschäft nicht möglich ist. Für Verspätungskosten werden die gleichen Annahmen getroffen wie im ersten Anwendungsfall.

### 6.2.1.3 Herleitung des Produktionssystems an drei Standorten

In diesem Abschnitt werden alle notwendigen Daten zur Modellierung des Produktionsystems hergeleitet. Der Fokus bei der Modellierung des Produktionssystems liegt auf der Zerspanung. Um diese als einstufigen Prozess im Umweltmodell abzubilden, wird eine repräsentative Referenzmaschine herangezogen. Mithilfe von Ausgangswerten der Referenzmaschine und weiteren Annahmen zu Standortgegebenheiten lässt sich die Fertigungszeit für die einzelnen Varianten und sodann die Gesamtfertigungszeit bestimmen. Die ausführliche Herleitung zu den Fertigungszeiten findet sich in Anhang A28. Die Werte zu den Fertigungszeiten lassen sich Tabelle 6-26 entnehmen.

*Tabelle 6-26: Überblick zur Gesamtfertigungszeit und den Fertigungszeiten pro Variante*

Kolbendurchmesser in mm	Ø12	Ø16	Ø20	Ø25	Ø32	Ø40	Ø50	Ø63	Ø80	Ø100
Verkaufsvolumen stk/a (1)	3.778	7.568	11.346	15.124	18.914	18.914	15.124	11.346	7.568	3.778
Durchschnittliche Fertigungszeit in min. (2)	2,19	2,40	2,66	2,87	3,41	3,80	4,28	5,21	6,12	7,18
Fertigungszeit pro Jahr in min. (3)=(1)*(2)	8.261	18.153	30.211	43.386	64.434	71.803	64.780	59.136	46.302	27.142
Gesamtfertigungszeit Summe über (3)	<b>433.608</b>									

Für die Bestimmung der Stationsanzahl bzw. Maschinenanzahl wird nun angenommen, dass allgemein im Zwei-Schicht-Betrieb produziert wird. Außerdem wird angenommen, dass die gefertigten Komponenten und damit die betrachteten Maschinen nicht allein

für den einen Pneumatikzylinder produzieren. Stattdessen wird angenommen, dass noch vier weitere Produkte gefertigt werden. Mit den bekannten sechs Stationstypen bzw. Maschinentypen aus Anhang A28 und den Ausführungen in Anhang A29 ergibt sich eine Anzahl von sieben Maschinen.

Weiter wird angenommen, dass sich die sieben Maschinen auf die drei Standorte Deutschland, China und USA gleichermaßen verteilen. Die siebte Maschine ist dem deutschen Standort mit dem höchsten Automatisierungsgrad zugewiesen. Es ergibt sich die Maschinenaufteilung wie in Tabelle 6-27 dargestellt.

*Tabelle 6-27: Überblick über die Maschinenaufteilung über die Standorte*

Land	Stationstyp	Stationsnummer
Deutschland	Höher automatisiert	1
Deutschland	Höher automatisiert	2
Deutschland	Weniger automatisiert	3
USA	Höher automatisiert	4
USA	Weniger automatisiert	5
China	Höher automatisiert	6
China	Weniger automatisiert	7

Die Logistik von verschiedenen Standorten zu verschiedenen Kunden wird nicht berücksichtigt. Die PPS fungiert in diesem Anwendungsfall zentral für alle Standorte.

Zuletzt werden Kostenfaktoren betrachtet. Hier werden als Erstes die variablen Kostenfaktoren fokussiert. Da diese nicht explizit vorliegen, werden sie im Folgenden über eine Deckungsbeitragsrechnung berücksichtigt. Es wird zunächst angenommen, dass eine konstante relative Marge über alle Produkte angestrebt wird. Mit dieser Annahme kann die Deckungsbeitragsquote als verwandte Kennzahl herangezogen und alle variablen Kosten in den Preis integriert werden. Die Deckungsbeitragsquote ist nach Thommen & Achleitner et al. (2020, S. 194, 388) definiert. Sie wird im Rahmen dieser Erprobung mit 34 % der Umsatzerlöse gemäß einer Expertenbefragung angenommen. Für etwaige Lagerhaltungskosten werden wiederum basierend auf den Angaben von Biedermann (2008, S. 43) 15 % pro Jahr bzw. 0,041 % pro Tag allgemein festgelegt.

Zuletzt werden noch die Maschineninvestitionskosten bestimmt. Hierzu wurde eine Anfrage bei einem Hersteller durchgeführt und daraus die folgende Investitionstabelle abgeleitet (vgl. auch Anhang A30).

*Tabelle 6-28: Investitionstabelle für Maschinen aufgeschlüsselt nach Standorten*

Land	Stationstyp	Investitionskosten
Deutschland	Höher automatisiert	230.000 €
Deutschland	Weniger automatisiert	218.500 €
USA	Höher automatisiert	207.000 €
USA	Weniger automatisiert	195.500 €
China	Höher automatisiert	184.000 €
China	Weniger automatisiert	172.500 €

## 6.2.2 Exemplarische Anwendung und Ergebnisse

Wie im ersten Anwendungsfall erfolgt nun die Anwendung des Ansatzes. Es wird wiederum zuerst eine Hyperparameteroptimierung durchgeführt und als Zweites das Potenzial des Ansatzes auf Signifikanz geprüft. Im Gegensatz zum ersten Anwendungsfall wird auf eine erweiterte Betrachtung des Ansatzes hinsichtlich der optimalen Kapazitäten und ihrer Abhängigkeit von Nachfrage, Preisbereitschaft und Lieferzeitpräferenz verzichtet. Stattdessen wird direkt in der Hyperparameteroptimierung und im Vergleich mit dem Status Quo die Möglichkeit zur Einsparung von Kapazitäten berücksichtigt.

### 6.2.2.1 Hyperparameteroptimierung

Analog zur Hyperparameteroptimierung in Anwendungsfall 1 werden die durchschnittlichen Belohnungsfunktionswerte in der automatischen Suche als erstes Auswahlkriterium herangezogen. Sodann werden die drei besten Konfigurationen aus der automatischen Suche gegenübergestellt und eine Konfiguration für die Prüfung der Hypothesen ausgewählt. Die Hyperparameteroptimierung basiert hierbei auf der vom Verfasser der Arbeit betreuten Abschlussarbeit A\_Henzi (2021).

Bevor die Suche beginnen kann, müssen zunächst die spezifischen Steuer- und Rauschfaktoren bestimmt werden. Die spezifischen Steuerfaktoren wurden basierend auf einer Expertenbefragung festgelegt und sind in Tabelle 6-29 zu sehen.

Tabelle 6-29: Spezifische Steuerfaktoren im zweiten Anwendungsfall

Name des konstant gehaltenen Parameters	Wert
Strafe für einen abgelehnten Auftrag $W_1$	15
Strafe für Verfrühung $W_2$	5
Strafe für geplante Verspätung $W_3$	175
Belohnung für Preisaufschlag $W_4$	$100 \cdot p_{rel,i,v}$
Strafe für realisierte Verspätung $W_5$	175
Belohnung für mögliche Kapazitätseinsparung $W_6$	5
Belohnung für realisierte Kapazitätseinsparung $W_7$	10

Wie aus der Vorstellung des Anwendungsfalls zu entnehmen ist, werden 6 Stationstypen bzw. hier Maschinentypen mit je einer Station untersucht. Der Standort Deutschland besitzt als einziges noch eine zusätzliche Station des höher automatisierten Typs, sodass sich eine Stationsanzahl von sieben ergibt. Bei den Kundentypen wird angenommen, dass diese durch die 8 vorgestellten Typen aus Kapitel 5.2.1.1 repräsentiert werden können. Die Stations- bzw. Maschinentypen und Kundentypen stellen die konstant gehaltenen Rauschfaktoren dar. Sie sind in Tabelle 6-30 zusammengefasst.

Tabelle 6-30: Konstant gehaltene Rauschfaktoren im zweiten Anwendungsfall

Name des konstant gehaltenen Rauschfaktors	Wert
Anzahl Kunden(-typen) $I$	8
Anzahl der Stationstypen $MS_{typ}$	6
Anzahl Stationen je Stationstyp $O_{Styp}$	1 (+1 für Standort DE)

Bei den zufallsbasierten Rauschfaktoren gilt es, die Bestellhäufigkeit und -menge abzubilden. Gemäß Kapitel 6.2.1.2 wird die Bestellhäufigkeit einer Exponentialverteilung folgend und die Bestellmenge als Poissonverteilung angenommen. Die zufallsbasierten Rauschfaktoren sind in Tabelle 6-31 zusammengefasst.

Tabelle 6-31: Zufallsbasierte Rauschfaktoren im zweiten Anwendungsfall

Name des zufallsbasierten Rauschfaktors	Wert
Bestellhäufigkeit $BH_{i,j}$	$F_{exp} \text{ mit } \frac{1}{\lambda_{exp}} = 1,5 \text{ Monate}$ $= 64.800 \text{ Minuten}$
Bestellmenge $BM_{i,j}$	Produkt- und kundenspezifisch poissonverteilt (vgl. Anhang A27)
Fertigungszeiten $fz_{j,q,typ}$	vgl. Anhang A28 $\pm 5\%$

Auch im zweiten Anwendungsfall ist es notwendig, eine Einschwingphase und eine Trainingsphase für das System zu definieren. Da im zweiten Anwendungsfall ein deutlich größeres Auftragsvolumen besteht, ist es möglich, die Simulationszeit der Trainingsphase für den Agenten zu verkleinern. Sie sollte jedoch trotzdem größer als ein Jahr sein, um die Belohnung einer Kapazitätseinsparung auslösen zu können. Die Phasen ergeben sich gemäß Abbildung 6-9. Die Bewertungsphase wurde hierbei auf 525.600 min festgelegt, weil das genau einem Simulationsjahr entspricht.

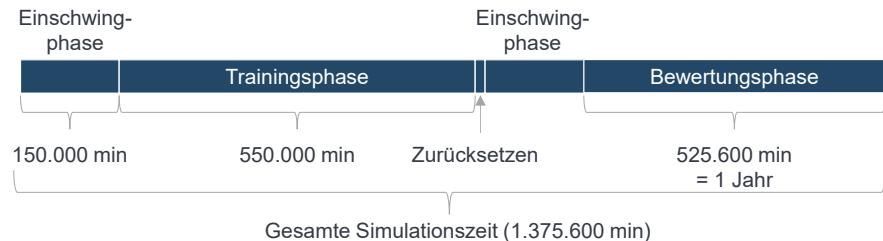


Abbildung 6-9: Übersicht über die Simulationsphasen im Anwendungsfall 2

Es wurden 40 Konfigurationen innerhalb des vorgegebenen Suchraums getestet. Durch die Parallelisierung auf acht Kernen benötigte die Suche ungefähr 15 Stunden. Die Konfiguration zur Durchführung der Zufallssuche ist Teil der frei verfügbaren Implementierung (Stamer 2021b). Eine Übersicht über die Belohnungsverteilung der gefundenen Konfigurationen findet sich in der Abbildung 6-10.

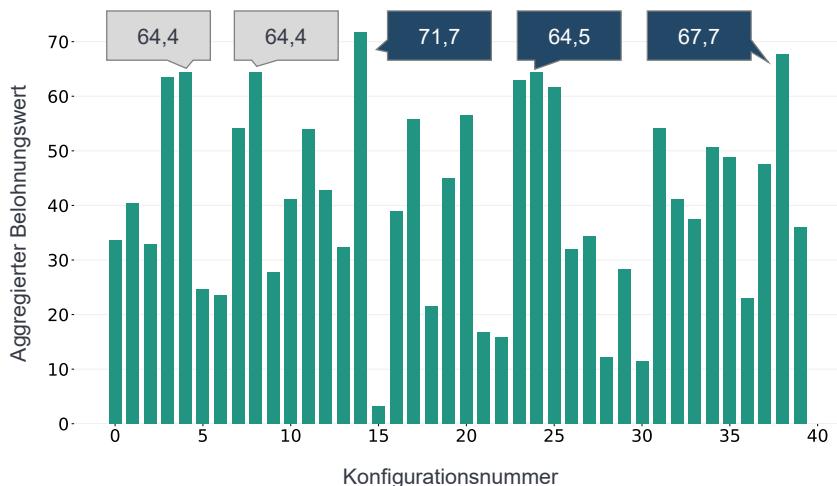


Abbildung 6-10: Verteilung der Belohnungswerte in der Hyperparameteroptimierung

Es werden nun nachfolgend die besten drei Konfigurationen aus der Suche entnommen und in ausführlichen Simulationen zehnmal wiederholt. Die gemittelten Ergebnisse werden miteinander verglichen mit dem Ziel, die beste Konfiguration für den Anwendungsfall zu finden.

### Szenario 1: Statische, heterogene Kundenpräferenzen

Für Szenario 1 ergeben sich die Werte wie folgt:

*Tabelle 6-32: Vergleich der Konfigurationen in Szenario 1 (Statische Präferenzen)*

	Konfiguration 1	Konfiguration 2	Konfiguration 3
Gewinn [€]	8.435.432	7.596.636	8.066.312
Servicegrad [%]	91,5 %	90,4 %	96,2 %
Verspätungen [-]	2.441	2.758	1.113
Entgangene Aufträge [-]	12.883	18.269	13.841
Summe Liefetermin-abweichungen [min]	22.759.650	135.353.900	35.613.510
Eingesparte Stationen [-]	0	0,14	0

Es zeigt sich, dass die erste Konfiguration in drei von sechs Ausprägungen dominiert. Die zweitbeste Konfiguration bei statischen Kundenpräferenzen ist die Nummer drei. Diese dritte Konfiguration ist jedoch beim Gewinn nur 4 % und bei entgangenen Aufträgen nur 7 % hinter der ersten Konfiguration. Wie eingangs beschrieben, wird hier auch die Einsparung von Kapazitäten integriert betrachtet. Es zeigt sich, dass nur Konfiguration 2 teilweise zu einer Einsparung von Stationen geführt hat.

### Szenario 2: Heterogene Kundenpräferenzen mit Drift zur Kurzfristigkeit

Im zweiten Szenario ändern sich während der Bewertungsphase die Kundenpräferenzen hin zu kurzfristigeren Terminen, wie es zum Beispiel bei kürzer werdenden Produktlebenszyklen passieren kann. Die Ergebnisse finden sich in der nachfolgenden Tabelle:

Tabelle 6-33: Vergleich der Konfigurationen in Szenario 2 (Drift zu Kurzfristigkeit)

	Konfiguration 1	Konfiguration 2	Konfiguration 3
Gewinn [€]	5.197.001	6.988.671	7.481.713
Servicegrad [%]	98,8 %	95,3 %	99,8 %
Verspätungen [-]	311	1.169	49
Entgangene Aufträge [-]	24.720	20.531	16.332
Summe Lieferterminabweichungen [min]	73.639.160	119.525.600	55.417.350
Eingesparte Stationen [-]	0,29	0,14	0

Es zeigt sich, dass unter dynamischem Verhalten Konfiguration 3 deutlich anpassungsfähiger ist als Konfiguration 1. Konfiguration 3 ist in diesem Szenario die beste Lösung mit Ausnahme der eingesparten Maschinen. Zur Erinnerung sei hier erwähnt, dass die Lieferterminabweichungen die Summe aus Verspätung und Verfrühung ist. Daher erklärt sich die niedrige Ausprägung bei der Anzahl der Verspätungen im Verhältnis zu der Summe an Lieferterminabweichungsminuten über alle drei Konfigurationen hinweg. Insgesamt zeigt sich hier jedoch ein positiver Zusammenhang zwischen der Lieferterminabweichung und der Anzahl an Verspätungen.

### Szenario 3: Heterogene Kunden mit Drift zu geringerer Preisbereitschaft

Das dritte Szenario simuliert eine sinkende Preisbereitschaft. Das bedeutet, dass es zum Preiskampf kommt.

*Tabelle 6-34: Vergleich der Konfigurationen in Szenario 3  
(Sinkende Preisbereitschaft)*

	Konfiguration 1	Konfiguration 2	Konfiguration 3
Gewinn [€]	4.285.115	6.955.750	7.523.246
Servicegrad [%]	99,4 %	99,1 %	99,5 %
Verspätungen [-]	126	219	140
Entgangene Aufträge [-]	24.515	20.039	14.558
Summe Liefetermin-abweichungen [min]	114.389.600	141.626.100	34.918.610
Eingesparte Stationen [-]	0,14	0,14	0

Im dritten Szenario ist erneut die Konfiguration 3 dominierend. Aus den Experimenten zeigt sich, dass es unterschiedliche Strategien gibt, die zu einer hohen Belohnung des Agenten führen können. Es scheint somit eine Pareto-Front zu geben. Die Konfiguration 2 beispielsweise führt in allen drei Szenarien zu Maschineneinsparungen. Jedoch sind die Gewinnwerte dafür nicht dominierend. Konfiguration 1 könnte besonders für heterogene, statische Märkte geeignet sein, da sie in diesem Szenario dominiert.

Basierend auf einer Dominanzanalyse über alle drei Szenarien hinweg wird für den nachfolgenden Hypothesentest Konfiguration 3 gewählt. Konfiguration 3 konnte sich in den meisten Bewertungsgrößen behaupten und erreichte in zwei von drei Fällen den höchsten Gewinn. Konfiguration 3 ist Teil der frei verfügbaren Implementierung (Stamer 2021b).

### 6.2.2.2 Potenzial im Vergleich zum Status Quo

Ebenfalls in gleicher Weise wie im Kapitel 6.1.2.2 wird nun die Signifikanzuntersuchung hinsichtlich des Potenzials durchgeführt. Es werden wiederum die drei bekannten Szenarien unterschieden. Die Details der Durchführung finden sich in Anhang A31.

Insgesamt zeigt sich, dass bei dem zweiten Anwendungsfall der vorgestellte Ansatz in allen drei Szenarien signifikant besser als der Status Quo in Bezug auf den Gewinn ist. Jedoch haben sich nicht alle Überlegungen aus der Einleitung bestätigt. Eine Übersicht darüber, in welchen Szenarien der Agent in welcher Zielgröße besser als der Status Quo abgeschnitten hat, lässt sich der nachfolgenden Tabelle entnehmen.

*Tabelle 6-35: Darstellung der Verbesserung zum Status Quo durch Mittelwertvergleich*

Legende: Agent ist...	Szenario 1		Szenario 2		Szenario 3	
	Statische Lieferzeit	Flexible Lieferzeit	Statische Lieferzeit	Flexible Lieferzeit	Statische Lieferzeit	Flexible Lieferzeit
Gewinn	✓	✓	✓	✓	✓	✓
Servicegrad	✓	✓	✓	✓	✓	✓
Verspätungen	✓	✓	✓	✓	✓	✓
Entgangene Aufträge	✓	✗	✗	✗	✓	✗
Liefertermin- abweichungen	✗	✗	✗	✗	✗	✗
Eingesparte Stationen	—	—	—	—	—	—

In der Einleitung wird die Vermutung aufgestellt, dass mit einer dynamischen Lieferzeit-Preisgestaltung höhere Gewinne erzielt, Kundenbedürfnisse besser bedient und gleichzeitig Kapazitäten nivelliert werden können. Die Produktionsnivellierung sollte hierbei zu Einsparungen von Stationen führen. Der Agent erzielt zwar einen höheren Gewinn und schafft es gleichzeitig, weniger Verspätungen zu erzeugen. Jedoch konnten in keiner Konfiguration Kapazitäten gespart werden. Es lässt sich jedoch vermuten, dass es in dem gegebenen Anwendungsfall von Vorteil war, mehr Aufträge zu bedienen, anstatt Kapazitäten einzusparen. Ein Indiz hierauf ist das Ergebnis der Hyperparameteroptimierung, welche mit Konfiguration 2 zeigte, dass Kapazitätseinsparungen im Vergleich

zur Verarbeitung eines höheren Auftragsvolumens unterlegen sind. Unter der Annahme eines konstanten Auftragsvolumens gegenüber dem Vergleichssystem würden hingegen weniger Kapazitäten benötigt. Dies zeigte sich bereits im ersten Anwendungsfall. Die Anzahl der entgangenen Aufträge deutet gleichzeitig darauf hin, dass der Agent versucht, möglichst vorteilhafte Kunden, also mit hoher Zahlungsbereitschaft und zur Auslastung passender Lieferzeitpräferenz, zu bedienen.

## 7 Diskussion, kritische Würdigung und Ausblick

In diesem Kapitel werden die in der Erprobung erzielten Ergebnisse diskutiert (Kapitel 7.1), die Erfüllung der Zielsetzung und Anforderungen beurteilt (Kapitel 7.2) und darauf aufbauend ein Ausblick auf mögliche zukünftige Forschung gegeben (Kapitel 7.3).

### 7.1 Diskussion der Ergebnisse

Die Ergebnisse dieser Arbeit entstammen der Erprobung in zwei Anwendungsfällen, der Kalibrierung von Hochpräzisionsgewichten und der Produktion von Pneumatikzylinder. Das Vorgehen zur Erlangung der Ergebnisse besteht aus zwei Schritten: Der Hyperparameteroptimierung zur Identifikation einer geeigneten Konfiguration und dem Vergleich zwischen dem in dieser Arbeit entwickelten Lösungsmodell und dem Vergleichssystem in zuvor definierten Szenarien.

In der Hyperparameteroptimierung zeigt sich sowohl im ersten Anwendungsfall als auch im zweiten Anwendungsfall, dass die Leistung des Lösungsmodells von der Wahl der richtigen Konfiguration stark abhängt. Gleichzeitig können die besten Konfigurationen der Hyperparameteroptimierung akzeptable Werte in den aufgestellten Zielgrößen erreichen. Hiermit ist gemeint, dass keine Zielgröße zu Gunsten einer anderen vollständig vernachlässigt wird. Somit erfüllt die Hyperparameteroptimierung ihren Zweck.

Im Vergleich des Lösungsmodells mit dem Vergleichssystem wird als erstes die zentrale Zielgröße, der Unternehmensgewinn, fokussiert. In beiden Anwendungsfällen und jeweils allen drei Szenarien konnte der Agent des Lösungsmodells einen statistisch signifikant höheren Unternehmensgewinn als das Vergleichssystem erzielen. Dieses Ergebnis beweist, dass der Agent in der Lage ist, individuelle Angebote zu erstellen und so einen höheren Preis sowie einen höheren Gesamtnutzen zu erzielen, ohne mit übermäßigen Verspätungskosten den zusätzlichen Umsatz wieder zunichte zu machen. Letztere Aussage bezieht sich auf den Zusammenhang, dass die Verspätungskosten mit 200% des Preises nach acht Tagen sehr strikt angesetzt werden und somit eine Strategie einheitlich kurzer Lieferzeitangebote mit vielen Verspätungen unmöglich ist. Der Agent ist nur in der Lage, einen höheren Gewinn zu erzielen, wenn er die Verspätungen gleichzeitig begrenzt.

In einer erweiterten Betrachtung ergibt sich, dass der Agent besonders bei hoher Nachfrage und begrenzten Kapazitäten einen höheren Unternehmensgewinn erreicht als das Vergleichssystem. In dieser erweiterten Betrachtung werden Kapazitäten manuell

hinzugefügt bzw. entfernt. Hierbei zeigt sich, dass sogar eine Erhöhung der Maschinenkapazität im ersten Anwendungsfall bei einer Nachfrage zum jetzigen Zeitpunkt sinnvoll wäre. Entsprechend spart der Agent auch von sich aus keine Stationen ein, sondern nutzt diese, um die Kundenpräferenzen besser bedienen zu können.

Neben dem Unternehmensgewinn als zentrale Zielgröße werden als weitere Zielgrößen der Servicegrad, die Anzahl Verspätungen, die Anzahl entgangener Aufträge und die Lieferterminabweichung bestimmt. Bei diesen Zielgrößen zeigt sich in beiden Anwendungsfällen ein gemischtes Bild. Setzt sich der Agent bei einer Zielgröße in einem Szenario durch, unterliegt er unter Umständen in einem anderen Szenario dem Vergleichssystem. Im zweiten Anwendungsfall hingegen zeigen sich im ersten Szenario Ergebnisse, bei denen der Agent in allen Zielgrößen bis auf die Lieferterminabweichung besser abschneidet als das Vergleichssystem. Diese Ergebnisse sind von besonderem Interesse, denn sie deuten darauf hin, dass es prinzipiell möglich ist, das scheinbare Paradoxon der gleichzeitigen Steigerung gegenläufiger Zielgrößen zu realisieren. Gegenläufige Zielgrößen sind zum Beispiel der Unternehmensgewinn, die Anzahl an Aufträgen und der Servicegrad. Dieses scheinbare Paradoxon wird möglich, indem der Agent in der Lage ist, nicht allein die Kundenpräferenzen optimal zu bedienen, sondern gleichzeitig eine Verschiebung von Aufträgen dergestalt umzusetzen, dass damit Verspätungen vermieden werden.

Bei einem Blick auf die entgangenen Aufträge fällt auf, dass der Agent über beide Anwendungsfälle hinweg in neun von zwölf Vergleichen mehr Aufträge ablehnt. Gleichzeitig ist der Unternehmensgewinn stets höher. Dies deutet darauf hin, dass der Agent sich auf ausgewählte Kunden fokussiert und diese optimal zu bedienen versucht. Fraglich ist, ob sich diese Strategie der Fokussierung in der Praxis verallgemeinern lässt. Hierzu lohnt es sich, die mathematischen Zusammenhänge, auf denen der Ansatz fußt, nochmals zu rekapitulieren. Der Agent bildet seine optimale Strategie auf Basis der Belohnungsfunktion aus. Die Belohnungsfunktion muss hierbei an dem wahren Optimierungsziel des Anwenders ausgerichtet sein. Je besser beide übereinstimmen, desto besser kann der Agent abschneiden. In dieser Arbeit ist das übergeordnete Ziel die Gewinnoptimierung auf Basis der Zielfunktion, welche in Kapitel 5.1.1 aufgestellt wird. Im Rahmen der Arbeit zeigt sich, dass die Belohnungsfunktion ihren Zweck erfüllt und bis zu einem gewissen Grad an der Gewinnoptimierung ausgerichtet ist. Das Lösungsmodell funktioniert also so, wie es vom Zielsystem vorgegeben ist. Folglich kann eine Verallgemeinerung der Strategie einer Kundenfokussierung nur insofern zutreffend

sein, wie das in dieser Arbeit aufgestellte Zielsystem mit den Zielen weiterer Unternehmen übereinstimmt. Es verbleibt für den Anwender die individuelle Aufgabe, das Zielsystem korrekt aufzustellen, um die Unternehmensziele zu repräsentieren. Hierbei zeigt sich eine Stärke der Arbeit: Aufgrund der hohen Adaptivität des Verfahrens lässt sich das Gleichgewicht durch Justierung der Belohnungsfunktionsparameter leicht einstellen. Dies zeigt sich anhand der Erprobung in zwei heterogenen Anwendungsfällen, welche unterschiedliche Konfigurationen hinsichtlich Produktionssystem und Kunden in der Umwelt nutzen. Dennoch erzielt der Agent stets einen höheren Gewinn als das Vergleichssystem.

## 7.2 Kritische Würdigung

Anschließend an die Diskussion der Ergebnisse erfolgt nun eine kritische Würdigung der Arbeit. Die zentrale Forschungsfrage dieser Arbeit lautet:

*Wie kann eine dynamische Lieferzeit-Preisgestaltung im Bereich der variantenreichen Produktion unter Berücksichtigung von Produktionsstandort und Netzwerk vorteilhaft gestaltet werden?*

Zu dieser Forschungsfrage wird die Hypothese aufgestellt, dass ein Lösungsmodell zur dynamischen Lieferzeit-Preisgestaltung in der Lage sein kann, sowohl der Herausforderung steigender Individualisierung als auch der Herausforderung steigender Volatilität im Produktionssystem zu begegnen.

Um der Hypothese und der Forschungsfrage nachzugehen, ist ein Teilziel dieser Arbeit, zunächst ein geeignetes Zielsystem aufzustellen. Das hierfür erarbeitete Zielsystem stellt den Unternehmensgewinn in den Fokus, wie es auch im Stand der Forschung verbreitet ist. Darüber hinaus werden jedoch auch weitere Kennzahlen, wie zum Beispiel die Anzahl entgangener Aufträge oder der Servicegrad, in das Zielsystem aufgenommen, um weitere relevante Aspekte im Auftragsmanagement abzubilden. Das Design des Zielsystems ist besonders kritisch für die Qualität des Ansatzes, da der Agent seine Aktionen über die Belohnungsfunktion an dem Zielsystem ausrichtet. Daher müssen die Wahl der Kennzahlen bzw. Faktoren und deren Gewichtung im Zielsystem hinterfragt werden. Für die vorgestellten Anwendungsfälle stellt sich das Zielsystem als geeignet heraus. Jedoch könnten neue Zusammenhänge in weiteren Anwendungsfällen die Aufnahmen weiterer Kennzahlen notwendig machen, um weiterhin zufriedenstellende Ergebnisse zu erzielen.

Neben dem Aufstellen eines Zielsystems ist ein Teilziel dieser Arbeit, ein geeignetes Lösungsverfahren auszuwählen. Mögliche Lösungsverfahren werden in dieser Arbeit vorgestellt und hierbei ein Schwerpunkt auf das Reinforcement Learning gelegt. Es zeigt sich in der Analyse des Standes der Forschung, dass das Reinforcement Learning in der dynamischen Lieferzeit-Preisgestaltung die Anforderungen, zum Beispiel bei der Adaptivität und Skalierbarkeit, besser erfüllt als andere Ansätze. In der weiteren Ausarbeitung wird dann aus dem Reinforcement Learning das Actor Critic Verfahren gewählt. Ein Grund für das Actor Critic Verfahren ist die Möglichkeit der Nutzung einer kontinuierlichen Funktion, welche sich wiederum aus der optimalen Segregation der Kundentypen ableitet. Obgleich dieser Ansatz funktioniert, könnte die Frage auftreten, ob eine solche Segregation notwendig ist, da sie das Verfahren verkompliziert. Stattdessen wäre es möglich, einen diskreten Ansatz mit einer limitierten Anzahl von Kundentypen einzusetzen. Eine Herausforderung wäre hierbei jedoch, die Kundentypen geeignet zu bestimmen. Demgegenüber ist das vorgestellte Verfahren in der Lage, beliebige Typen mit unterschiedlichen Präferenzprofilen zu adaptieren.

Nach der Auswahl eines geeigneten Lösungsverfahrens wird die Umsetzung als Teilziel fokussiert. Hierfür wird das allgemeine Reinforcement Learning Modell bestehend aus Agent, Umwelt und Interaktion herangezogen. Die Umwelt des Agenten wird explizit modelliert. Die Modellierung erscheint zunächst im Widerspruch zu der Wahl eines modellfreien Verfahrens. In der Praxis wäre die explizite Modellierung der Umwelt jedoch nicht notwendig und der Agent könnte direkt eingesetzt werden. Die Modellierung der Umwelt erfolgte hier mit dem primären Ziel, den Einsatz ohne Eingriff in ein reales System über eine Simulation erproben zu können. Obgleich die Modellierung der Umwelt für den Einsatz des Agenten also auf den ersten Blick nicht relevant ist, hängt das Design des Agenten jedoch mit der Modellierung der Umwelt zusammen. Die zugrundeliegenden Annahmen in der Umweltmodellierung, wie zum Beispiel die Existenz nur eines Hochpunktes in der Nutzenfunktion des Kunden, sind daher besonders kritisch zu hinterfragen und müssten empirisch geprüft werden. Eine weitere Annahme betrifft die Ablaufplanung. Diese wird mit einer fristbasierten Steuerung abgebildet. Auch wenn der Einsatz einer fristbasierten Steuerung bei Aufträgen mit variablen Lieferzeiten naheliegend erscheint, können dennoch in anderen Anwendungsfällen andere Steuerungskonzepte auftreten. Für diese Steuerungskonzepte bleibt offen, ob das Lösungsmodell bzw. der Agent zu guten Ergebnissen kommt. Aufgrund der ausgeprägten Fähigkeit zur Adaption, kann jedoch vermutet, dass der Agent dennoch gute Ergebnisse

---

erzielt. Es zeigt sich an dieser Stelle außerdem eine allgemeine Stärke der Arbeit aus Sicht der Forschung: Die Umsetzung des Lösungsverfahrens erfolgte in Form einer quelloffenen, der Öffentlichkeit zugänglichen Software, die zudem modular aufgebaut ist. Daher können neue Modellaspekte leicht integriert und getestet werden.

Zuletzt besteht ein Teilziel darin, ein Vorgehen zu entwickeln, mit welchem der Lösungsansatz im Rahmen der Erprobung zuverlässig getestet werden kann. Hierfür werden eine Hyperparameteroptimierung und ein Hypothesentest entwickelt. Im Rahmen der Erprobung wird die Überlegenheit des Lösungsmodells gegenüber einem definierten Vergleichssystem mithilfe der Hypothesentests nachgewiesen. Das Vergleichssystem wird aus dem Status Quo im Auftragsmanagement abgeleitet. Der Agent kann sich jedoch nicht bei allen Kennzahlen behaupten, sodass keine strenge Dominanz vorherrscht. In einer erweiterten Betrachtung wird die Stabilität des Verfahrens in Bezug auf variierende Kapazitäten und Nachfrageriveaus ausführlicher untersucht. Es zeigt sich hier die Vorteilhaftigkeit besonders bei Kapazitätsengpässen und hoher Nachfrage. Durch die Erprobung in zwei Anwendungsfällen und je drei Szenarien wird die generelle Anwendbarkeit des Ansatzes gezeigt.

Unabhängig von den Teilzielen dieser Arbeit stellt sich für die Anwendung des Ansatzes die Frage der Akzeptanz. Solange das gesamte Auftragsmanagement nicht vollautomatisch abläuft, sondern sich ein Mensch als Entscheider im System befindet, ist die Akzeptanz der Ergebnisse eine relevante Problemstellung. Reinforcement Learning an sich ist ein komplexes Verfahren, welches nur sehr schwer nachzuvollziehen ist. Ein Vorteil dieser Arbeit ist der Einsatz einer Simulation, sodass verschiedene Alternativen in verschiedenen Szenarien gezeigt werden können. Das Aufzeigen konsistenter Ergebnisse könnte die Akzeptanz erhöhen. Dennoch bietet die Arbeit keine durchgängige, deterministische Erklärbarkeit der Ergebnisse, sodass hier ein offener Punkt für den Transfer in die Praxis besteht.

Bei der Verfolgung der aufgestellten Teilziele wurden stets die allgemeinen Anforderungen an den Lösungsansatz berücksichtigt, um das Gesamtziel zu erreichen. Insgesamt ließe sich die zentrale Forschungsfrage mit der vorliegenden Arbeit also als beantwortet ansehen. Eine detaillierte Betrachtung zeigt, dass die zugrundeliegenden Hypothesen größtenteils zutreffen. Mithilfe des Lösungsansatzes ist es möglich, heterogene Kundenpräferenzen zu antizipieren und entsprechende Lieferzeit-Preisfunktionen aufzustellen. Die Prämisse, dass hierbei auch die Auslastung der Produktion in einer

Weise berücksichtigt werden würde, dass Kapazitäten nivelliert und sogar eingespart werden können, trifft teilweise zu. In der Erprobung wird keine Einsparung von Kapazitäten erzielt. In der erweiterten Betrachtung wird klar, dass sogar eine Kapazitätserhöhung sinnvoll wäre. Jedoch zeigen sich im zweiten Anwendungsfall auch Ergebnisse, bei denen der Agent in allen Kennzahlen bis auf die Lieferterminabweichung besser abschneidet als das Vergleichssystem. Somit ist es prinzipiell möglich, das scheinbare Paradoxon eines höheren Servicegrads, höheren Gewinns und mehr bearbeiteten Aufträgen zu erreichen. An dieser Stelle wären weitergehende Untersuchungen notwendig, um die Bereiche besser zu definieren, in denen der Agent in allen Zielgrößen besser abschneidet als das Vergleichssystem. Mit den jetzigen Anwendungen ist der Agent einzig im erzielten Gewinn konstant besser als das Vergleichssystem, was jedoch auch das zentrale Bewertungskriterium dieser Arbeit ist.

### 7.3 Ausblick

Ausgehend von dem dargelegten Lösungsansatz bieten sich zahlreiche Anknüpfungspunkte für neue und aufbauende Forschungsarbeiten. Es lassen sich drei Kategorien unterscheiden: Die erste Kategorie sind inkrementelle Verbesserungen mit nahezu gleichem Forschungsziel. Die zweite Kategorie behandelt Weiterentwicklungen mit verwandtem Forschungsziel. Die dritte Kategorie umfasst Arbeiten zur Untersuchung gänzlich neuer Forschungsfragen.

Bei den inkrementellen Verbesserungen des Ansatzes zählen vor allem solche, die das Modell in bestehenden Modellelementen detaillieren oder Modellelemente mit geringfügigem Ergebniseinfluss ergänzen. So könnten weitere Arbeiten Änderungen der Parametrisierung von Kunden, Produkten oder Stationen und deren Wirkung auf die Ergebnisse untersuchen. Das grundlegende Lösungsmodell als solches würde in seiner Logik folglich unverändert bleiben. Darüber hinaus könnte auch das Systemverhalten angepasst oder erweitert werden. Der vorgestellte Ansatz modelliert einige Modellelemente, wie zum Beispiel die Kundennutzenfunktion oder die Auslieferung von Produkten, deterministisch. Es wäre möglich, diese stochastisch abzubilden und den Einfluss auf die Stabilität des Systems zu untersuchen. In eine ähnliche Richtung ginge eine Erweiterung des bestehenden Produktionssystemmodells um neue Elemente, wie eine Intralogistik, eine taktische Produktionsplanung oder ein komplexeres Lagermanagementsystem. Denkbar wäre auch eine Optimierung des Lernprozesses an sich. Der einfachste Ansatz wäre die Hyperparameteroptimierung länger und mit größerem

Suchraum durchzuführen. Mit mehr Aufwand wäre ein intelligentes Entscheidungssystem für Hyperparameteroptimierung verbunden. Hier könnte beispielsweise der Verlauf der erzielten Belohnungswerte während des Simulationslaufs beobachtet und dieser bei instabilem Verhalten oder vergleichsweise niedrigen Belohnungswerten frühzeitig terminiert werden.

Zu den Weiterentwicklungen mit verwandtem Forschungsziel sollen alle Ideen zählen, die auf dieselbe Motivation einer dynamischen Lieferzeit-Preisgestaltung aufbauen, jedoch einen signifikanten Unterschied in der Zielstellung oder im Lösungsansatz aufweisen. Hierbei lohnt sich ein Blick in den Stand der Forschung dieser Arbeit. Verwandte Arbeiten behandeln dediziert die Annahme oder Ablehnung von Aufträgen, bevor es überhaupt zu einer Lieferzeit-Preisgestaltung käme. Das Annahmeproblem ließe sich also unter Umständen externalisieren. Weitere Arbeiten legen einen Schwerpunkt auf die Optimierung der Maschinenbelegung im Rahmen der Ablaufplanung. In dieser Arbeit basiert die Ablaufplanung auf einer eigens entwickelten Heuristik, die sich an den Grundlagen orientiert. Die Ablaufplanung ist dahingehend ausgelegt, dass sie im Gesamtkontext mit dem Agenten funktioniert. Diese Heuristik der Ablaufplanung erlaubt nicht das Verschieben eines für eine Maschine bereits eingeplanten Auftrags auf eine andere Maschine. Wenn das Problem als Bin-Packing-Problem interpretiert wird, zeigt sich das Verbesserungspotenzial. Prinzipiell wären dann alle Ansätze im Bereich der Bin-Packing-Problematik als mögliche Verbesserung denkbar. Jedoch sollten Forschungsarbeiten mit diesem Ziel dem Lernprozess des Agenten weiterhin Rechnung tragen, da andernfalls die Funktionsfähigkeit der dynamischen Lieferzeit-Preisgestaltung gefährdet wäre. Ein dritter Ansatzpunkt bietet sich bei der Identifikation der Kunden. Im bisherigen Ansatz muss der Agent den Typ eines einzelnen Kunden oder einer Kundengruppe erlernen. Dies birgt Nachteile in der Skalierung und der resultierenden Komplexität. Für sehr große Kundenanzahlen ist hier die Einführung eines vorgesetzten, kundenindividuellen Klassifikators denkbar. Der Kundenklassifikator lässt sich in Form eines neuronalen Netzes in Kombination mit einem Suchalgorithmus umsetzen, der als Eingang die Lieferzeit-Preisgestaltungsvariablen des Agenten erhält. Als Output soll der vereinbarte Preis und die Lieferzeit des Kunden ausgegeben werden. Sobald der Klassifikator ausreichend trainiert ist, kann durch einfache Rastersuche möglicher Lieferzeit-Preis-Funktionen der Typ des Kunden geschätzt werden. Die Klassifikation kann dann beispielsweise anhand der acht vorgestellten Kundentypen erfolgen. Eine

Diskretisierung in mehr als acht Kundentypen wäre ebenfalls möglich. Auch die Implementierung einer Nachfrageprognose zur Verbesserung der Nivellierung durch den Agenten wäre ein Ansatz. Hierbei würde der Agent die prognostizierte Nachfrage als zusätzlichen Zustandsparameter erhalten. Ein solches Vorgehen bietet sich besonders für regelmäßige Nachfrageschwankungen an, die gut prognostiziert werden können. Ein letzter Anknüpfungspunkt in dieser Kategorie könnte den Fokus auf die Auswirkung des Ansatzes auf das ganze Wertschöpfungsnetzwerk legen. Durch Verschaltung von mehreren Systemen zur dynamischen Lieferzeit-Preisgestaltung, wie es hier erarbeitet ist, kann ein Wertschöpfungsnetzwerk simuliert werden. Hierbei wäre ersichtlich, wie eine durchgängige dynamische Lieferzeit-Preisgestaltung real die Agilität des Netzwerks beeinflussen würde.

Die dritte Kategorie mit weiterführenden Arbeiten könnte den bestehenden Ansatz auch zur Beantwortung gänzlich verschiedener Fragestellungen nutzen. Abstrakt gesehen ließen sich im bestehenden Ansatz die Modellkomponenten des Umweltmoduls durch andere Modellkomponenten ersetzen. Es ergäbe sich ein komplett neues System. Grundlage eines Ansatzes müsste stets eine ereignisdiskrete Simulation sein, die jedoch im Produktionsumfeld üblich ist. Der Agent ließe sich dann für Fragestellungen im Sinne einer Regelung einsetzen. Denkbar wäre der Einsatz beispielsweise für eine Produktionssteuerung oder die Rückführung von Qualitätsdaten zur Anpassung von Produktionsprozessen. Jedoch ist das Actor-Critic-Verfahren nicht für alle Fälle optimal geeignet. Der Einsatz müsste folglich im Einzelnen geprüft werden. Eine gute Daumenregel zur Beurteilung bietet hier die Betrachtung von Zustands- und Aktionsraum: Sind beide kontinuierlich, so ist das ein Argument für das Actor-Critic-Verfahren.

## 8 Zusammenfassung

Globale Wertschöpfungsnetzwerke, welche zur dominierenden Wertschöpfungsform geworden sind, produzieren zum einen eine wachsende Anzahl an Produktvarianten und müssen zum anderen auf zunehmende Volatilität reagieren. Dies lässt sich unter anderem zurückführen auf die stärkere Ausrichtung an den Kundenanforderungen, die zunehmende Vernetzung und die resultierende Störungsempfindlichkeit. In einem solchen Umfeld gilt es für Unternehmen, die eigenen Fähigkeiten zur Produktion ideal mit der individuellen Nachfrage im Netzwerk zu harmonisieren, um profitabel zu sein. Eine zentrale Funktion hat hierbei das Auftragsmanagement, welches die Lieferzeit-Preise-Paare pro Produktvariante gestaltet. Bei der Gestaltung muss berücksichtigt werden, dass sowohl Kundennachfrage als auch ein produzierender Standort dynamische Faktoren sind, die sich mit der Zeit ständig ändern. Um die nötige Adaptivität zu erreichen, bietet sich deshalb der Einsatz eines dynamisch-adaptiven Verfahrens aus dem Reinforcement Learning an, um die Lieferzeit-Preisgestaltung durchzuführen.

Ausgehend von dieser Motivation wird in der Arbeit ein Reinforcement Learning Verfahren umgesetzt, welches eine dynamische Lieferzeit-Preisgestaltung in variantenreicher Produktion ermöglicht. Das Verfahren erzielt eine Reduktion der auf die Produktion durchschlagenden Nachfragevolatilität durch eine preisgesteuerte Nivellierung von Nachfragespitzen bei gleichzeitiger Umsatzsteigerung durch eine individuell zugeschnittene Lieferzeit-Preis-Funktion.

Im Rahmen der Arbeit wird hierzu als Erstes ein Zielsystem aufgestellt. Als zweites wird ausgehend vom allgemeinen Reinforcement Learning Modell ein konkretes Lösungsmodell, bestehend aus Umwelt, Agent und deren Interaktion, entwickelt. Die Umwelt wird unterteilt in eine Standortebene und eine Netzwerkebene. Auf der Netzwerkebene ist ein markanter Bestandteil der Arbeit die Modellierung der Kunden mit individuellen Lieferzeit- und Preispräferenzen. Auf der Standortebene ist die Produktionssteuerung herauszuheben, die individuell für das Gesamtsystem ausgehend von einer ver-spätungsminimierenden Heuristik entwickelt ist. Das Agentenmodul behandelt die Schnittstellenfunktionen zwischen Agent und Umwelt, die Architektur und den Trainingsprozess. Hierbei werden Freiheitsgrade definiert, die anwendungsfallspezifisch fixiert werden. Die Freiheitsgrade bestehen aus einstellbaren Parametern wie zum Beispiel der Anzahl der Neuronen im neuronalen Netzwerk des Agenten oder der Lernrate.

Umweltmodul und Agentenmodul werden im Rahmen des Versuchsmoduls in ein experimentierbares Lösungsmodell überführt. Das Versuchsmodul dient des Weiteren der Versuchsplanung, der Orchestrierung der Experimente und der Analyse der Ergebnisse. Für den Potenzialnachweis werden drei Szenarien und ein Vergleichssystem mit zwei Ausprägungen aufgestellt, mit denen die zentrale Hypothese einer Gewinnsteigerung zu prüfen war. Der Lösungsansatz ist ohne proprietäre Software implementiert. Die resultierende Software ist frei verfügbar und verwendbar (Stamer 2021b).

Die Erprobung des Ansatzes basiert auf zwei Anwendungsfällen aus den vom BMBF geförderten Forschungsprojekten BaSys4SupplyQ und ReKoNeT. Der erste Anwendungsfall entstammt der Kalibrierung von Hochpräzisionsgewichten. Der zweite Anwendungsfall ist in der Produktion von Pneumatikzylindern angesiedelt. Die heterogenen Anwendungsfälle haben zum Ziel, die breite Anwendbarkeit und Stabilität des Ansatzes zu demonstrieren. Im Ergebnis zeigt sich, dass der Ansatz zur dynamischen Lieferzeit-Preisgestaltung funktioniert. Das System ist in der Lage, mit dynamischen, heterogenen Kundenpräferenzen umzugehen und gleichzeitig den Servicegrad zu verbessern, trotz variierender Kapazitätsauslastungen. Insgesamt kann der entwickelte Ansatz folglich die Forschungsfrage beantworten und somit einen Beitrag zur Steigerung der Wirtschaftlichkeit durch Beherrschung der Volatilität im Kontext variantenreicher Produktion erzielen.

Für zukünftige Forschungsarbeiten gibt es gleichzeitig eine Reihe möglicher Anknüpfungspunkte. Konkret zu nennen sind die Ausgestaltung des Umweltmodells, wie zum Beispiel die Erweiterung um weitere Steuerungskonzepte, und die Untersuchung der Auswirkung einer dynamischen Lieferzeit-Preisgestaltung im ganzen Netzwerk mit mehreren Wertschöpfungsstufen. Bei letzterem ergäbe sich die Frage, inwiefern die dynamische Lieferzeit-Preisgestaltung in der Lage wäre, die Agilität des gesamten Wertschöpfungsnetzwerks positiv zu beeinflussen.

## 9 Liste der eigenen Veröffentlichungen

Stamer, Peukert & Lanza 2021

Stamer, F., Peukert, S., Lanza, G. (2021). „Order Planning“ in *Global Manufacturing Management. Management for Professionals*, Hrsg. Friedli, T., Lanza, G., Remling, D., Springer, Cham. ISBN: 9783030727390.

Stamer & Maier et al. 2021

Stamer, F., Maier, S., Peukert, S., Lanza, G. (2021). „Adaptive and Dynamic Feedback Loops between Production System and Production Network based on the Asset Administration Shell“, *15th CIRP Conference on Intelligent Computation in Manufacturing Engineering*. Gulf of Naples, Italy

Stamer & Wegmann et al. 2021

Stamer, F., Wegmann, M., Peukert, S., Lanza, G. (2021). „Competence Development Program to Promote Collaborative and Agile Production Networks Based on a Learning Factory“, *Proceedings of the Conference on Learning Factories (CLF) 2021*, S. 1-6

Stamer & Bizenberger et al. 2020

Stamer, F., Bizenberger, P., Häfner, B., Lanza, G. (2020). „Deriving Collaboration Cases in Production Networks Considering Smart Services“. *Procedia CIRP*, 96, S. 33-38.

Stamer & Steinke et al. 2020

Stamer, F., Steinke, M., Silbernagel, R., Häfner, B., Lanza, G. (2020). „Using Smart Services as a Key Enabler for Collaboration in Global Production Networks“, *Procedia CIRP*, 93, S. 730-735.

Stamer & Leschinski et al. 2020

Stamer, F., Leschinski, G., Häfner, B., Lanza, G., Korczok, A. (2020): „Smart Services in Produktionsnetzwerken modellieren - Smart Services im Kontext der Kollaboration in Produktionsnetzwerken“, *wt Werkstattstechnik*, 110 (7/8), S. 541-544.

Stamer & Labbouz et al. 2019

Stamer, F.; Labbouz, A.; Häfner, B. & Lanza, G. (2019), „Framework für Smart Services als Grundlage für Kollaboration im Kontext aufkommender industrieller Dienstleistungen“, *Proceedings of the 9th WGP Jahreskongress 2019*, Hrsg. Wulfsberg, J. P., Hintze, W. & Behrens, B., Apprimus Verlag, Aachen, S. 603-612.

Silbernagel & Stamer et al. 2019

Silbernagel, R.; Stamer, F.; Häfner, B.; Linzbach, J. & Lanza, G. (2019). „Kollaboration in globalen Wertschöpfungsnetzwerken“. *ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb*, 5, S. 314-317

Friedli & Schuh et al. 2019

Friedli, T.; Schuh, G.; Lanza, G.; Remling, D.; Gützlaff, A. & Stamer, F. (2019). „Next Level Production Networks“, *ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb*, 114 (3), S. 101-104.

Storms & Roggendorf et al. 2017

Storms, S., Roggendorf, S., Stamer, F., Obdenbusch, M. et al. (2017). „PLM -supported automated process planning and partitioning for collaborative assembly processes based on a capability analysis”, *Proceedings of the 7th WGP Jahrestkongress 2017*, Hrsg. Schmitt, R., Schuh, G., Apprimus Verlag, Aachen, S. 241.

## 10 Literaturverzeichnis

Verweise gemäß dem Schema *A\_Nachname Jahr* beziehen sich auf studentische Arbeiten am wbk des Karlsruher Instituts für Technologie (KIT), die von dem Autor der Dissertation angeleitet wurden.

### A\_Henzi 2021

A\_Henzi, M. (2021), *Reinforcement Learning zur dynamischen Preisgestaltung in Produktionsnetzwerken*. Masterarbeit, Karlsruher Institut für Technologie (KIT), wbk Institut für Produktionstechnik, Karlsruhe.

### A\_Lorenz 2021

A\_Lorenz, P. (2021), *Entwicklung eines Nachfrage-Prognosesystems am Beispiel der Kalibrierung von Hochpräzisionsgewichten*. Masterarbeit, Karlsruher Institut für Technologie (KIT), wbk Institut für Produktionstechnik, Karlsruhe.

### A\_Ochs 2021

A\_Ochs, F. (2021), *Simulation eines globalen Produktionsnetzwerkes mit dynamischer Preisgestaltung zur Ermittlung des Potenzials von Lieferzeitflexibilität*. Masterarbeit, Karlsruher Institut für Technologie (KIT), wbk Institut für Produktionstechnik, Karlsruhe.

### A\_Werle 2021

A\_Werle, A. (2021), *Ereignisdiskrete Simulation eines Produktionsnetzwerkes zur Ermittlung des Potentials einer dynamischen Preis-Lieferzeit-Bestimmung*. Masterarbeit, Karlsruher Institut für Technologie (KIT), wbk Institut für Produktionstechnik, Karlsruhe.

### acatech 2017

acatech (Hrsg.) (2017), *Kollaboration als Schlüssel zum erfolgreichen Transfer von Innovationen. Handlungsempfehlungen für Forschung und Entwicklung*, Utzverlag, München. ISBN: 9783831646173.

### Afèche 2013

Afèche, P. (2013), „Incentive-Compatible Revenue Management in Queueing Systems: Optimal Strategic Delay“, *Manufacturing & Service Operations Management*, 15 (3), S. 423–443.

### Afèche & Pavlin 2016

Afèche, P. & Pavlin, J. M. (2016), „Optimal Price/Lead-Time Menus for Queues

- with Customer Choice: Segmentation, Pooling, and Strategic Delay“, *Management Science*, 62 (8), S. 2412–2436.
- Aggarwal 2015  
Aggarwal, C. C. (2015), *Data Mining*, Springer, Cham. ISBN: 9783319141411.
- Aggarwal 2018  
Aggarwal, C. C. (2018), *Neural Networks and Deep Learning*, Springer, Cham. ISBN: 9783319944623.
- Ahn, Gümüş & Kaminsky 2007  
Ahn, H.-s.; Gümüş, M. & Kaminsky, P. (2007), „Pricing and Manufacturing Decisions When Demand Is a Function of Prices in Multiple Periods“, *Operations Research*, 55 (6), S. 1039–1057.
- Akan, Ata & Olsen 2012  
Akan, M.; Ata, B. & Olsen, T. (2012), „Congestion-Based Lead-Time Quotation for Heterogenous Customers with Convex-Concave Delay Costs. Optimality of a Cost-Balancing Policy Based on Convex Hull Functions“, *Operations Research*, 2012 (60), S. 1505–1519.
- Albana, Frein & Hammami 2018  
Albana, A. S.; Frein, Y. & Hammami, R. (2018), „Effect of a lead time-dependent cost on lead time quotation, pricing, and capacity decisions in a stochastic make-to-order system with endogenous demand“, *International Journal of Production Economics*, 203, S. 83–95.
- Arnold & Isermann et al. 2008  
Arnold, D.; Isermann, H.; Kuhn, A.; Tempelmeier, H. & Furmans, K. (2008), *Handbuch Logistik*, Springer, Berlin, Heidelberg. ISBN: 9783540729280.
- Arredondo & Martinez 2010  
Arredondo, F. & Martinez, E. (2010), „Learning and adaptation of a policy for dynamic order acceptance in make-to-order manufacturing“, *Computers & Industrial Engineering*, 58 (1), S. 70–83.
- Arulkumaran & Deisenroth et al. 2017  
Arulkumaran, K.; Deisenroth, M. P.; Brundage, M. & Bharath, A. A. (2017), „A Brief Survey of Deep Reinforcement Learning“, *IEEE Signal Processing Magazine*, 34 (6), S. 1–16.

## Ata &amp; Olsen 2013

Ata, B. & Olsen, T. L. (2013), „Congestion-based leadtime quotation and pricing for revenue maximization with heterogeneous customers“, *Queueing Systems*, 73, S. 35–78.

## Backhaus &amp; Erichson et al. 2018

Backhaus, K.; Erichson, B.; Plinke, W. & Weiber, R. (2018), *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*, Springer, Berlin, Heidelberg. ISBN: 9783662566541.

## Bahrami, Chen &amp; Wong 2021

Bahrami, S.; Chen, Y. C. & Wong, V. W. S. (2021), „Deep Reinforcement Learning for Demand Response in Distribution Networks“, *IEEE Transactions on Smart Grid*, 12 (2), S. 1496–1506.

## Banks 2010

Banks, J. (2010), *Discrete-event system simulation*, Pearson; Pearson Education, Upper Saddle River, NJ. ISBN: 9780138150372.

## Bartholomae &amp; Wiens 2016

Bartholomae, F. W. & Wiens, M. (2016), *Spieltheorie. Ein anwendungsorientiertes Lehrbuch*, Springer Gabler, Wiesbaden. ISBN: 9783834944191.

## Barut &amp; Sridharan 2005

Barut, M. & Sridharan, V. (2005), „Revenue Management in Order-Driven Production Systems“, *Decision Sciences*, 36 (2), S. 287–316.

## Baykasoğlu &amp; Subulan et al. 2020

Baykasoğlu, A.; Subulan, K.; Güçdemir, H.; Dudaklı, N. & Eren Akyol, D. (2020), „Revenue management for make-to-order manufacturing systems with a real-life application“, *The Engineering Economist*, 65 (1), S. 27–65.

## Benioudakis, Burnetas &amp; Ioannou 2021

Benioudakis, M.; Burnetas, A. & Ioannou, G. (2021), „Lead-time quotations in unobservable make-to-order systems with strategic customers: Risk aversion, load control and profit maximization“, *European Journal of Operational Research*, 289 (1), S. 165–176.

## Bergstra &amp; Bengio 2012

Bergstra, J. & Bengio, Y. (2012), „Random Search for Hyper-Parameter Optimization“, *Journal of Machine Learning Research*, 13 (10), S. 281–305.

## Bertagnolli 2018

Bertagnolli, F. (2018), *Lean Management*, Springer Fachmedien Wiesbaden, Wiesbaden. ISBN: 9783658131234.

## Bichler &amp; Trommsdorff 2021

Bichler, A. & Trommsdorff, V. (2021), „Präferenzmodelle bei der Conjointanalyse“ in *Conjointanalyse*, Hrsg. D. Baier & M. Brusch, Springer, Berlin, Heidelberg, S. 81–93.

## Biedermann 2008

Biedermann, H. (2008), *Ersatzteilmanagement. Effiziente Ersatzteilloistik für Industrieunternehmen*, Springer, Berlin, Heidelberg. ISBN: 9783540008507.

## Blazewicz &amp; Ecker et al. 2019

Blazewicz, J.; Ecker, K. H.; Pesch, E.; Schmidt, G.; Sterna, M. & Weglarz, J. (2019), *Handbook on Scheduling*, Springer International Publishing, Cham. ISBN: 9783319998480.

## Bloech &amp; Bogaschewsky et al. 2014

Bloech, J.; Bogaschewsky, R.; Buscher, U.; Daub, A.; Götze, U. & Roland, F. (2014), *Einführung in die Produktion*, Springer, Berlin, Heidelberg. ISBN: 9783642318924.

## Bohnstedt 2014

Bohnstedt, J. (2014), *Vertragsrecht im Einkauf. Erfolgsfaktor im Supply Change Risk Management (SCRM)*, Springer Gabler, Wiesbaden. ISBN: 9783658032487.

## Bondoux &amp; Nguyen et al. 2020

Bondoux, N.; Nguyen, A. Q.; Fiig, T. & Acuna-Agost, R. (2020), „Reinforcement learning applied to airline revenue management“, *Journal of Revenue and Pricing Management*, 19 (5), S. 332–348.

## Bungartz &amp; Zimmer et al. 2013

Bungartz, H.-J.; Zimmer, S.; Buchholz, M. & Pflüger, D. (2013), *Modellbildung und Simulation*, Springer, Berlin, Heidelberg. ISBN: 9783642376559.

Calin 2020

Calin, O. L. (2020), *Deep learning architectures. A mathematical approach*, Springer, Cham, Switzerland. ISBN: 9783030367206.

Capocchi 2019

Capocchi, A. (2019), *Economic Value and Revenue Management Systems*, Springer International Publishing, Cham. ISBN: 9783030024161.

Çelik & Maglaras 2008

Çelik, S. & Maglaras, C. (2008), „Dynamic Pricing and Lead-Time Quotation for a Multiclass Make-to-Order Queue“, *Management Science*, 54 (6), S. 1132–1146.

Chaharsooghi, Honarvar & Modarres 2011

Chaharsooghi, S. K.; Honarvar, M. & Modarres, M. (2011), „A multi-stage stochastic programming model for dynamic pricing and lead time decisions in multi-class make-to-order firm“, *Scientia Iranica*, 18 (3), S. 711–721.

Charnsirisakskul, Griffin & Keskinocak 2006

Charnsirisakskul, K.; Griffin, P. M. & Keskinocak, P. (2006), „Pricing and scheduling decisions with leadtime flexibility“, *European Journal of Operational Research*, 171 (1), S. 153–169.

Crawley 2015

Crawley, M. J. (2015), *Statistics. An introduction using R*, Wiley, Chichester. ISBN: 9781118941096.

Dean, Voss & Draguljić 2017

Dean, A.; Voss, D. & Draguljić, D. (2017), *Design and Analysis of Experiments*, Springer International Publishing, Cham. ISBN: 9783319522487.

DIN 60050-351:2014-09

DIN 60050-351:2014-09 (2014), *Internationales Elektrotechnisches Wörterbuch*, Deutsches Institut für Normung e.V., Beuth Verlag, Berlin.

Ding & Huang et al. 2020

Ding, Z.; Huang, Y.; Yuan, H. & Dong, H. (2020), „Introduction to Reinforcement Learning“ in *Deep Reinforcement Learning*, Hrsg. H. Dong, Z. Ding & S. Zhang, Springer Singapore, Singapore, S. 47–123.

DMG MORI Academy GmbH 2021

DMG MORI Academy GmbH (2021), *Kurzangebot. CTX alpha 500 - perfekt für die Ausbildung Industrie 4.0* [19.08.2021].

Domschke & Drexel et al. 2015

Domschke, W.; Drexel, A.; Klein, R. & Scholl, A. (2015), *Einführung in Operations Research*, Springer, Berlin, Heidelberg. ISBN: 9783662482155.

Dyckhoff 2006

Dyckhoff, H. (2006), *Produktionstheorie. Grundzüge industrieller Produktionswirtschaft : mit 20 Tabellen*, Springer, Berlin, Heidelberg. ISBN: 9783540326007.

Eley 2012

Eley, M. (2012), *Simulation in der Logistik*, Springer, Berlin, Heidelberg. ISBN: 9783642273728.

ElHafsi 2000

ElHafsi, M. (2000), „An operational decision model for lead-time and price quotation in congested manufacturing systems“, *European Journal of Operational Research*, 126 (2), S. 355–370.

Erlach 2010

Erlach, K. (2010), *Wertstromdesign*, Springer, Berlin, Heidelberg. ISBN: 9783540898665.

Esfandiari, Mashreghi & Emami 2019

Esfandiari, S.; Mashreghi, H. & Emami, S. (2019), „Coordination of Order Acceptance, Scheduling and Pricing Decisions in Unrelated Parallel Machine Scheduling“, *Journal of Industrial Engineering & Production Research*, 30, S. 195–205.

Eversheim 1996

Eversheim, W. (1996), *Organisation in der Produktionstechnik*, Springer, Berlin, Heidelberg. ISBN: 9783642877384.

Feng, Liu & Liu 2011

Feng, J.; Liu, L. & Liu, X. (2011), „An Optimal Policy for Joint Dynamic Price and Lead-Time Quotation“, *Operations Research*, 59 (6), S. 1523–1527.

Feng & Zhang 2017

Feng, J. & Zhang, M. (2017), „Dynamic quotation of leadtime and price for a Make-

To-Order system with multiple customer classes and perfect information on customer preferences“, *European Journal of Operational Research*, 258 (1), S. 334–342.

García, Luengo & Herrera 2015

García, S.; Luengo, J. & Herrera, F. (2015), *Data Preprocessing in Data Mining*, Springer International Publishing, Cham. ISBN: 9783319102467.

Garmdare, Lotfi & Honarvar 2018

Garmdare, H. S.; Lotfi, M. M. & Honarvar, M. (2018), „Integrated model for pricing, delivery time setting, and scheduling in make-to-order environments“, *Journal of Industrial Engineering International*, 14, S. 55–64.

Gass & Fu 2013

Gass, S. I. & Fu, M. C. (2013), *Encyclopedia of Operations Research and Management Science*, Springer US, Boston, MA. ISBN: 9781441911377.

Geier 2014

Geier, S. (2014), *Demand Fulfillment bei Assemble-to-Order-Fertigung. Analyse, Optimierung und Anwendung in der Computerindustrie*, Springer Fachmedien Wiesbaden, Wiesbaden. ISBN: 9783658048785.

Giesecking 2009

Giesecking, T. (2009), *Gewinnoptimale Preisbestimmung in werbefinanzierten Märkten*, Springer Gabler, Wiesbaden. ISBN: 9783834918970.

Gönsch 2017

Gönsch, J. (2017), „A survey on risk-averse and robust revenue management“, *European Journal of Operational Research*, 263 (2), S. 337–348.

Gosavii, Bandla & Das 2002

Gosavii, A.; Bandla, N. & Das, T. K. (2002), „A reinforcement learning approach to a single leg airline revenue management problem with multiple fare classes and overbooking“, *IIE Transactions*, 34 (9), S. 729–742.

Gottmann 2016

Gottmann, J. (2016), *Produktionscontrolling*, Springer Fachmedien Wiesbaden, Wiesbaden. ISBN: 9783658019501.

Grondman & Busoniu et al. 2012

Grondman, I.; Busoniu, L.; Lopes, G. A. D. & Babuska, R. (2012), „A Survey of Actor-Critic Reinforcement Learning: Standard and Natural Policy Gradients“, *IEEE Transactions on Systems, Man, and Cybernetics*, 42 (6), S. 1291–1307.

Guadix & Cortés et al. 2010

Guadix, J.; Cortés, P.; Onieva, L. & Muñozuri, J. (2010), „Technology revenue management system for customer groups in hotels“, *Journal of Business Research*, 63 (5), S. 519–527.

Guhlich, Fleischmann & Stolletz 2015

Guhlich, H.; Fleischmann, M. & Stolletz, R. (2015), „Revenue management approach to due date quoting and scheduling in an assemble-to-order production system“, *OR Spectrum : Quantitative Approaches in Management*, 37, S. 951–982.

Gutenschwager & Rabe et al. 2017

Gutenschwager, K.; Rabe, M.; Spieckermann, S. & Wenzel, S. (2017), *Simulation in Produktion und Logistik*, Springer, Berlin, Heidelberg. ISBN: 9783662557440.

Hafizoğlu, Gel & Keskinocak 2016

Hafizoğlu, A. B.; Gel, E. S. & Keskinocak, P. (2016), „Price and Lead Time Quotation for Contract and Spot Customers“, *Operations Research*, 64 (2), S. 406–415.

Hancock & Khoshgoftaar 2020

Hancock, J. T. & Khoshgoftaar, T. M. (2020), „Survey on categorical data for neural networks“, *Journal of Big Data*, 7 (1), S. 1–41.

Hao & Yu 2012

Hao, J. & Yu, J. (2012), „Agent-Based Dynamic Order Acceptance Policy in Make-to-Order Manufacturing“. *Fourth International Conference on Computational and Information Sciences*. Chongqing, China, IEEE. ISBN: 9781467324069.

Hao, Yu & Wu 2013

Hao, J.; Yu, J. J. & Wu, M. C. (2013), „Dynamic Joint Decision on Price and Delivery Date in MTO Manufacturer Based on Agent“, *Energy Development*, 860-863, S. 2812–2816.

Helming, Schmidt & Kompa 2012

Helming, J.; Schmidt, C. & Kompa, S. (2012), „Grundlagen der Produktionsplanung

und -steuerung 2. Zeitdynamische Simulation in der Produktion“ in *Produktionsplanung und -steuerung 2. Grundlagen der PPS*, Hrsg. G. Schuh & V. Stich, Springer, Berlin, Heidelberg, 195-231.

Helmold 2020

Helmold, M. (2020), *Total Revenue Management (TRM)*, Springer International Publishing, Cham. ISBN: 9783030469849.

Herzog 2017

Herzog, A. (2017), *Callcenter – Analyse und Management*, Springer Fachmedien Wiesbaden, Wiesbaden. ISBN: 9783658183080.

Hillig 2006

Hillig, T. (2006), *Verfahrensvarianten der Conjoint-Analyse zur Prognose von Kaufentscheidungen*, DUV, Wiesbaden. ISBN: 9783824483297.

Hing, van Harten & Schuur 2007

Hing, M. M.; van Harten, A. & Schuur, P. C. (2007), „Reinforcement learning versus heuristics for order acceptance on a single resource“, *Journal of Heuristics*, 13 (2), S. 167–187.

Hintsches 2012

Hintsches, A. (2012), *Dynamische Kapazitätssteuerung bei kundenindividueller Auftragsproduktion in der stahlverarbeitenden Industrie*, Springer Gabler, Wiesbaden. ISBN: 9783834942074.

Hotz 2007

Hotz, I. (2007), *Simulationsbasierte Frühwarnsysteme zur Unterstützung der operativen Produktionssteuerung und -planung in der Automobilindustrie*. Dissertation, Magdeburg.

Huber 2014

Huber, S. (2014), *Informationsintegration in dynamischen Unternehmensnetzwerken. Architektur, Methode und Anwendung*, Springer Gabler, Wiesbaden. ISBN: 9783658077471.

Ivanov & Pavlov et al. 2016

Ivanov, D.; Pavlov, A.; Dolgui, A.; Pavlov, D. & Sokolov, B. (2016), „Disruption-

- driven supply chain (re)-planning and performance impact assessment with consideration of pro-active and recovery policies“, *Transportation Research*, 90, S. 7–24.
- Ivanov, Tsipoulanidis & Schönberger 2019  
Ivanov, D.; Tsipoulanidis, A. & Schönberger, J. (2019), *Global Supply Chain and Operations Management*, Springer International Publishing, Cham. ISBN: 9783319943121.
- Jaehn & Pesch 2019  
Jaehn, F. & Pesch, E. (2019), *Ablaufplanung*, Springer, Berlin, Heidelberg. ISBN: 9783662587799.
- Jammernegg, Reiner & Wakolbinger 2018  
Jammernegg, W.; Reiner, G. & Wakolbinger, T. (2018), „Circular Supply Chain: Combining Supply Chain Strategy and Circular Economy“ in *Handbuch Produktions- und Logistikmanagement in Wertschöpfungsnetzwerken*, Hrsg. H. Corsten, R. Gössinger & T. S. Spengler, De Gruyter, Oldenbourg, 67–85.
- Jin 2012  
Jin, X. (2012), *Optimal price and delivery time quotation with production scheduling for make-to-order manufacturing*. Dissertation, Nanyang Technological University, Singapur.
- Jin, Li & Sivakumar 2013  
Jin, X.; Li, K. & Sivakumar, A. I. (2013), „Scheduling and optimal delivery time quotation for customers with time sensitive demand“, *International Journal of Production Economics*, 145 (1), S. 349–358.
- Juan & Keenan et al. 2021  
Juan, A. A.; Keenan, P.; Martí, R.; McGarragh, S.; Panadero, J.; Carroll, P. & Oliva, D. (2021), „A review of the role of heuristics in stochastic optimisation: from metaheuristics to learnheuristics“, *Annals of Operations Research*.
- Kádár & Egri et al. 2018  
Kádár, B.; Egri, P.; Pedone, G. & Chida, T. (2018), „Smart, simulation-based resource sharing in federated production networks“, *CIRP Annals*, 67 (1), S. 503–506.

## Kahneman &amp; Tversky 1979

Kahneman, D. & Tversky, A. (1979), „Prospect theory. An analysis of decision under risk“, *Econometrica : journal of the Econometric Society*, 47 (2), S. 263–291.

## Kellner, Lienland &amp; Lukesch 2020

Kellner, F.; Lienland, B. & Lukesch, M. (2020), *Produktionswirtschaft*, Springer, Berlin, Heidelberg. ISBN: 9783662614457.

## Kimes &amp; Wirtz 2015

Kimes, S. E. & Wirtz, J. (2015), *Revenue management. Advanced strategies and tools to enhance firm profitability*, Now Publishers, Boston. ISBN: 9781680830286.

## Klein &amp; Steinhardt 2008

Klein, R. & Steinhardt, C. (2008), *Revenue Management*, Springer, Berlin, Heidelberg. ISBN: 9783540688433.

## Kleppmann 2020

Kleppmann, W. (2020), *Versuchsplanung. Produkte und Prozesse optimieren*, Hanser, München. ISBN: 9783446463974.

## Klocke &amp; König 2008

Klocke, F. & König, W. (2008), *Fertigungsverfahren 1. Drehen, Fräsen, Bohren*, Springer, Berlin, Heidelberg. ISBN: 9783540234586.

## Kollmann 2020

Kollmann, T. (2020), *Handbuch Digitale Wirtschaft*, Springer Fachmedien Wiesbaden, Wiesbaden. ISBN: 9783658172909.

## Kolmar 2017

Kolmar, M. (2017), *Grundlagen der Mikroökonomik*, Springer, Berlin, Heidelberg. ISBN: 9783662550434.

## Kubickova 2021

Kubickova, M. (2021), „Revenue management in manufacturing: systematic review of literature“, *Journal of Revenue and Pricing Management*, 21, S. 1–6.

## Kühn 2006

Kühn, W. (2006), *Digitale Fabrik. Fabriksimulation für Produktionsplaner*, Hanser, München. ISBN: 3446406190.

## Kummer 2013

Kummer, S. (2013), *Grundzüge der Beschaffung, Produktion und Logistik*, Pearson, München. ISBN: 9783868940930.

## Kunath &amp; Winkler 2019

Kunath, M. & Winkler, H. (2019), „Adaptive Assistenzsysteme zur Entscheidungsunterstützung für die dynamische Auftragsabwicklung: Konzeptionelle Überlegungen und Anwendungsszenarien unter Berücksichtigung des Digitalen Zwillings des Produktionssystems“ in *Handbuch Industrie 4.0 und Digitale Transformation*, Hrsg. R. Obermaier, Springer Fachmedien Wiesbaden, Wiesbaden, S. 269–294.

## Lanza &amp; Ferdows et al. 2019

Lanza, G.; Ferdows, K.; Kara, S.; Mourtzis, D.; Schuh, G.; Váncza, J.; Wang, L. & Wiendahl, H.-P. (2019), „Global production networks: Design and operation“, *CIRP Annals*, 68 (2), S. 823–841.

## Lanza &amp; Nyhuis et al. 2018

Lanza, G.; Nyhuis, P.; Fisel, J.; Jacob, A.; Nielsen, L.; Schmidt, M. & Stricker, N. (2018), *Wandlungsfähige, menschzentrierte Strukturen in Fabriken und Netzwerken der Industrie 4.0*, Utzverlage, München.

## Lanza &amp; Treber 2019

Lanza, G. & Treber, S. (2019), „Transparency increase in global production networks based on multi-method simulation and metamodeling techniques“, *CIRP Annals*, 68 (1), S. 439–442.

## Leimeister 2020

Leimeister, J. M. (2020), *Dienstleistungsengineering und -management*, Springer, Berlin, Heidelberg. ISBN: 9783662598573.

## Li, Lin &amp; Ye 2014

Li, Y.; Lin, Q. & Ye, F. (2014), „Pricing and promised delivery lead time decisions with a risk-averse agent“, *International Journal of Production Research*, 52 (12), S. 3518–3537.

## Li, Wang &amp; Sawhney 2012

Li, X.; Wang, J. & Sawhney, R. (2012), „Reinforcement learning for joint pricing, lead-time and scheduling decisions in make-to-order systems“, *European Journal of Operational Research*, 221 (1), S. 99–109.

Liu, Lu & Qi 2020

Liu, Z.; Lu, L. & Qi, X. (2020), „Price quotation for orders with different due dates“, *International Journal of Production Economics*, 220 (3), S. 1–23.

Lohnert & Fischer 2019

Lohnert, N. & Fischer, K. (2019), „Einsatz des Revenue Managements in der Make-to-Order-Produktion“ in *Logistik im Wandel der Zeit – Von der Produktionssteuerung zu vernetzten Supply Chains*, Hrsg. M. Schröder & K. Wegner, Springer Fachmedien Wiesbaden, Wiesbaden, S. 235–262.

Lu, Hong & Zhang 2018

Lu, R.; Hong, S. H. & Zhang, X. (2018), „A Dynamic pricing demand response algorithm for smart grid: Reinforcement learning approach“, *Applied Energy*, 220, S. 220–230.

März & Krug et al. 2011

März, L.; Krug, W.; Rose, O. & Weigert, G. (2011), *Simulation und Optimierung in Produktion und Logistik*, Springer, Berlin, Heidelberg. ISBN: 9783642145353.

Öner-Közen & Minner 2018

Öner-Közen, M. & Minner, S. (2018), „Dynamic pricing, leadtime quotation and due date based priority dispatching“, *International Journal of Production Research*, 56 (15), S. 5118–5130.

OIML R 111-1

OIML R 111-1 (2004), *Metrological and technical requirements*, Organisation Internationale de Métrologie Légale.

Ouazene & Yalaoui et al. 2017

Ouazene, Y.; Yalaoui, F.; Kelly, R. & Idjeraoui, T. (2017), „Coordination and optimization of dynamic pricing and production decisions“. *IEEE Symposium Series on Computational Intelligence (SSCI)*. Honolulu, Hawaii, IEEE, S. 1–6. ISBN: 9781538627266.

Pekgün 2010

Pekgün, P. (2010), „Pricing and Lead-Time Decisions“ in *Wiley Encyclopedia of Operations Research and Management Science*, Hrsg. J. J. Cochran, L. A. Cox, P. Keskinocak, J. P. Kharoufeh & J. C. Smith, John Wiley & Sons, Hoboken, NJ, USA.

## Pieper 2017

Pieper, M. (2017), *Mathematische Optimierung*, Springer Fachmedien Wiesbaden, Wiesbaden. ISBN: 9783658169749.

## Piller 2008

Piller, F. T. (2008), *Mass Customization. Ein wettbewerbsstrategisches Konzept im Informationszeitalter*, Deutscher Universitätsverlag, Wiesbaden. ISBN: 3835003550.

## Pinedo 2016

Pinedo, M. L. (2016), *Scheduling. Theory, Algorithms, and Systems*, Springer International Publishing, Cham. ISBN: 9783319265780.

## Plaat 2020

Plaat, A. (2020), *Learning to Play. Reinforcement Learning and Games*, Springer International Publishing, Cham. ISBN: 9783030592370.

## Plambeck 2004

Plambeck, E. L. (2004), „Optimal Leadtime Differentiation via Diffusion Approximations“, *Operations Research*, 52 (2), S. 213–228.

## Queiroz &amp; Ivanov et al. 2020

Queiroz, M. M.; Ivanov, D.; Dolgui, A. & Fosso Wamba, S. (2020), „Impacts of epidemic outbreaks on supply chains: mapping a research agenda amid the COVID-19 pandemic through a structured literature review“, *Annals of Operations Research*, S. 1–38.

## Rabe, Spiekermann &amp; Wenzel 2008

Rabe, M.; Spiekermann, S. & Wenzel, S. (2008), *Verifikation und Validierung für die Simulation in Produktion und Logistik*, Springer, Berlin, Heidelberg. ISBN: 9783540352815.

## Rao 2014

Rao, V. R. (2014), *Applied Conjoint Analysis*, Springer, Berlin, Heidelberg. ISBN: 9783540877523.

## Rebala, Ravi &amp; Churiwala 2019

Rebala, G.; Ravi, A. & Churiwala, S. (2019), *An Introduction to Machine Learning*, Springer International Publishing, Cham. ISBN: 9783030157289.

## Salewski 2018

Salewski, H. (2018), „Auftragsorientierte Produktionsprogrammplanung“ in *Handbuch Produktions- und Logistikmanagement in Wertschöpfungsnetzwerken*, Hrsg. H. Corsten, R. Gössinger & T. S. Spengler, De Gruyter, Oldenbourg, S. 230–245.

## Scharf, Helfenstein &amp; Jäger 2021

Scharf, F.; Helfenstein, F. & Jäger, J. (2021), „Actor vs Critic: Learning the Policy or Learning the Value“ in *Reinforcement Learning Algorithms: Analysis and Applications*, Hrsg. B. Belousov, H. Abdulsamad, P. Klink, S. Parisi & J. Peters, Springer International Publishing, Cham, S. 123–133.

## Schenk, Wirth &amp; Müller 2014

Schenk, M.; Wirth, S. & Müller, E. (2014), *Fabrikplanung und Fabrikbetrieb*, Springer, Berlin, Heidelberg. ISBN: 9783642054587.

## Schuh, Brandenburg &amp; Cuber 2012

Schuh, G.; Brandenburg, U. & Cuber, S. (2012), „Grundlagen der Produktionsplanung und -steuerung. Aufgaben“ in *Produktionsplanung und -steuerung 1. Grundlagen der PPS*, Hrsg. G. Schuh & V. Stich, Springer, Berlin, Heidelberg, 29-81.

## Schuh, Brosze &amp; Brandenburg 2012

Schuh, G.; Brosze, T. & Brandenburg, U. (2012), „Grundlagen der Produktionsplanung und -steuerung. Aachener PPS-Modell“ in *Produktionsplanung und -steuerung 1. Grundlagen der PPS*, Hrsg. G. Schuh & V. Stich, Springer, Berlin, Heidelberg, S. 11–28.

## Schuh, Schenk &amp; Servos 2015

Schuh, G.; Schenk, M. & Servos, N. (2015), „Design of a Simulation Model for the Assessment of a Real-time Capable Disturbance Management in Manufacturing Supply Chains“, *Procedia Manufacturing*, 3, S. 425–432.

## Schuh &amp; Schmidt 2014

Schuh, G. & Schmidt, C. (Hrsg.) (2014), *Produktionsmanagement. Handbuch Produktion und Management 5*, Springer, Berlin, Heidelberg. ISBN: 978-3-642-54287-9.

Schuh, Schmidt & Adema 2014

Schuh, G.; Schmidt, C. & Adema, J. (2014), „Produktionsmanagement. Auftragsmanagement“ in *Produktionsmanagement. Handbuch Produktion und Management* 5, Hrsg. G. Schuh & C. Schmidt, Springer, Berlin, Heidelberg, S. 109–150.

Schuh, Schmidt & Helmig 2012

Schuh, G.; Schmidt, C. & Helmig, J. (2012), „Grundlagen der Produktionsplanung und -steuerung. Prozesse“ in *Produktionsplanung und -steuerung 1. Grundlagen der PPS*, Hrsg. G. Schuh & V. Stich, Springer, Berlin, Heidelberg, 109-194.

Seo & Chang et al. 2021

Seo, D.-W.; Chang, K.; Cheong, T. & Baek, J.-G. (2021), „A reinforcement learning approach to distribution-free capacity allocation for sea cargo revenue management“, *Information Sciences*, 571 (3), S. 623–648.

Sibbertsen & Lehne 2021

Sibbertsen, P. & Lehne, H. (2021), *Statistik. Einführung für Wirtschafts- und Sozialwissenschaftler*, Springer, Berlin, Heidelberg. ISBN: 9783662626955.

Siebertz, van Bebber & Hochkirchen 2017

Siebertz, K.; van Bebber, D. & Hochkirchen, T. (2017), *Statistische Versuchsplanung*, Springer, Berlin, Heidelberg. ISBN: 9783662557426.

Simon & Fassnacht 2019

Simon, H. & Fassnacht, M. (2019), *Price Management*, Springer International Publishing, Cham. ISBN: 9783319994550.

Slamanig 2011

Slamanig, M. (2011), *Produktwechsel als Problem im Konzept der Mass Customization. Theoretische Überlegungen und empirische Befunde*. Zugl.: Klagenfurt, Univ., Diss., 2010, Springer Gabler, Wiesbaden. ISBN: 9783834926463.

So & Song 1998

So, K. C. & Song, J.-S. (1998), „Price, delivery time guarantees and capacity selection“, *European Journal of Operational Research*, 111, S. 28–49.

Spengler, Rehkopf & Volling 2007

Spengler, T.; Rehkopf, S. & Volling, T. (2007), „Revenue management in make-to-order manufacturing—an application to the iron and steel industry“, *OR Spectrum : Quantitative Approaches in Management*, 29 (1), S. 157–172.

## Stachowiak 1973

Stachowiak, H. (1973), *Allgemeine Modelltheorie*, Springer, Wien. ISBN: 0387811060.

## Stamer 2021a

Stamer, F. (2021), *An Electron app using Vizgraph and JSON text to visualize state graphs*. DOI: 10.5281/zenodo.5633115.

## Stamer 2021b

Stamer, F. (2021), *Dynamic Pricing in Production Networks. An Implementation with Reinforcement Learning in Typescript*. DOI: 10.5281/zenodo.6468683.

## Stasinski 2020

Stasinski, K. (2020), „A Literature Review on Dynamic Pricing - State of Current Research and New Directions“ in *Advances in Computational Collective Intelligence*, Hrsg. M. Hernes, K. Wojtkiewicz & E. Szczerbicki, Springer International Publishing, Cham, S. 465–477.

## Stein 2021

Stein, O. (2021), „Unrestringierte Optimierung“ in *Grundzüge der Nichtlinearen Optimierung*, Hrsg. O. Stein, Springer, Berlin, Heidelberg, S. 23–113.

## Steven 2007

Steven, M. (2007), *Handbuch Produktion. Theorie - Management - Logistik - Controlling*, Kohlhammer, Stuttgart. ISBN: 9783170183124.

## Sutton &amp; Barto 2018

Sutton, R. S. & Barto, A. G. (2018), *Reinforcement learning. An introduction*, MIT Press, Cambridge. ISBN: 9780262193986.

## Thommen &amp; Achleitner et al. 2020

Thommen, J.-P.; Achleitner, A.-K.; Gilbert, D. U.; Hachmeister, D.; Jarchow, S. & Kaiser, G. (2020), *Allgemeine Betriebswirtschaftslehre*, Springer Fachmedien Wiesbaden, Wiesbaden. ISBN: 9783658272456.

## Tomczak &amp; Heidig 2014

Tomczak, T. & Heidig, W. (2014), *Revenue Management aus der Kundenperspektive*, Springer Fachmedien Wiesbaden, Wiesbaden. ISBN: 9783658007348.

## Váncza &amp; Monostori et al. 2011

Váncza, J.; Monostori, L.; Lutters, D.; Kumara, S. R.; Tseng, M.; Valckenaers, P. &

- van Brussel, H. (2011), „Cooperative and responsive manufacturing enterprises“, *CIRP Annals*, 60 (2), S. 797–820.
- Vengerov 2008
- Vengerov, D. (2008), „A gradient-based reinforcement learning approach to dynamic pricing in partially-observable environments“, *Future Generation Computer Systems*, 24 (7), S. 687–693.
- VDI 3633 Blatt 1
- VDI 3633 Blatt 1 (2014), *Simulation von Logistik-, Materialfluß- und Produktionsystemen*, Verein Deutscher Ingenieure, Beuth Verlag, Berlin.
- Volling 2009
- Volling, T. (2009), *Auftragsbezogene Planung bei variantenreicher Serienproduktion. Eine Untersuchung mit Fallstudien aus der Automobilindustrie*. Zugl.: Braunschweig, Techn. Univ., Diss., 2008, Springer Gabler, Wiesbaden. ISBN: 9783834914774.
- Vosooghidizaji, Taghipour & Canel-Depitre 2020
- Vosooghidizaji, M.; Taghipour, A. & Canel-Depitre, B. (2020), „Supply chain coordination under information asymmetry: a review“, *International Journal of Production Research*, 58 (6), S. 1805–1834.
- Wadhwa & Mishra et al. 2010
- Wadhwa, S.; Mishra, M.; Chan, F. T.S. & Ducq, Y. (2010), „Effects of information transparency and cooperation on supply chain performance: a simulation study“, *International Journal of Production Research*, 48 (1), S. 145–166.
- Wang & Wang 2019
- Wang, S. & Wang, X. (2019), „Parallel Machine Scheduling with Pricing and Rejection“. *16th International Conference on Service Systems and Service Management (ICSSSM)*. Shenzhen, China, IEEE, S. 1–5. ISBN: 9781728119410.
- Weber 2020
- Weber, F. (2020), *Künstliche Intelligenz für Business Analytics*, Springer Fachmedien Wiesbaden, Wiesbaden. ISBN: 9783658297725.
- Weichselbaum & Rothböck 2007
- Weichselbaum, E. & Rothböck, M. (2007), „Das Konzept Fixe Lieferzeiten (LiFix)“

in *Steuerung von Supply Chains. Strategien - Methoden - Beispiele*, Hrsg. P. Klaus, Springer Gabler, Wiesbaden, S. 165–204.

Werner 2015

Werner, H. (2015), „Supply Chain Performance messen“, *Controlling & Management Review*, 59 (1), S. 18–25.

Werner 2020

Werner, H. (2020), *Supply Chain Management*, Springer Fachmedien Wiesbaden, Wiesbaden. ISBN: 9783658324285.

Werners 2013

Werners, B. (2013), *Grundlagen des Operations Research*, Springer, Berlin, Heidelberg. ISBN: 9783642401015.

Wiendahl 2011

Wiendahl, H.-H. (2011), *Auftragsmanagement der industriellen Produktion*, Springer, Berlin, Heidelberg. ISBN: 9783642191480.

Wiendahl & ElMaraghy et al. 2007

Wiendahl, H.-P.; ElMaraghy, H. A.; Nyhuis, P.; Zäh, M. F.; Wiendahl, H.-H.; Duffie, N. & Brieke, M. (2007), „Changeable Manufacturing - Classification, Design and Operation“, *CIRP Annals*, 56 (2), S. 783–809.

Winter 2019

Winter, S. (2019), *Grundzüge der Spieltheorie*, Springer, Berlin, Heidelberg. ISBN: 9783662582145.

Woratschek & Roth 2017

Woratschek, H. & Roth, S. (2017), „Die Ermittlung von Preisbereitschaften und Preisreaktionen als Basis zur Bestimmung des Kundenwerts“ in *Kundenwert*, Hrsg. S. Helm, B. Günter & A. Eggert, Springer Fachmedien Wiesbaden, Wiesbaden, S. 321–346.

Wu, Chuang & Hsu 2014

Wu, I.-L.; Chuang, C.-H. & Hsu, C.-H. (2014), „Information sharing and collaborative behaviors in enabling supply chain performance: A social exchange perspective“, *International Journal of Production Economics*, 148, S. 122–132.

Wu & Kazaz et al. 2012

Wu, Z.; Kazaz, B.; Webster, S. & Yang, K.-K. (2012), „Ordering, Pricing, and Lead-

- Time Quotation Under Lead-Time and Demand Uncertainty“, *Production and Operations Management*, 21 (3), S. 576–589.
- Xiao, Yang & Shen 2011  
Xiao, T.; Yang, D. & Shen, H. (2011), „Coordinating a supply chain with a quality assurance policy via a revenue-sharing contract“, *International Journal of Production Research*, 49 (1), S. 99–120.
- Ye, Sun & Li 2016  
Ye, T.; Sun, H. & Li, Z. (2016), „Coordination of pricing and leadtime quotation under leadtime uncertainty“, *Computers & Industrial Engineering*, 102, S. 147–159.
- Yu, Zhao & Sun 2013  
Yu, D. Z.; Zhao, X. & Sun, D. (2013), „Optimal Pricing and Capacity Investment for Delay-Sensitive Demand“, *IEEE Transactions on Engineering Management*, 60 (1), S. 124–136.
- Zatta 2016  
Zatta, D. (2016), *Revenue Management in Manufacturing*, Springer International Publishing, Cham. ISBN: 9783319302393.
- Zhang & Yu 2020  
Zhang, H. & Yu, T. (2020), „Taxonomy of Reinforcement Learning Algorithms“ in *Deep Reinforcement Learning*, Hrsg. H. Dong, Z. Ding & S. Zhang, Springer Singapore, Singapore, S. 125–133.
- Zhao, Stecke & Prasad 2012  
Zhao, X.; Stecke, K. E. & Prasad, A. (2012), „Lead Time and Price Quotation Mode Selection: Uniform or Differentiated?“, *Production and Operations Management*, 21 (1), S. 177–193.
- Zimmermann 2008  
Zimmermann, H.-J. (2008), *Operations Research*, Vieweg+Teubner, Wiesbaden. ISBN: 9783834804556.

## Abbildungsverzeichnis

Abbildung 1-1: Überblick über die Herausforderungen im Auftragsmanagement	2
Abbildung 1-2: Überblick über Ausgangslage, Hypothesen und Zielsetzung der Arbeit	4
Abbildung 2-1: Ebenen der Produktion in Anlehnung an Wiendahl & ElMaraghy et al. (2007, S. 785f)	7
Abbildung 2-2: Einordnung der variantenreichen Produktion über den Kundenentkopplungspunkt und die zugehörigen Fertigungstypen	9
Abbildung 2-3: Übersicht über die Aufgaben der PPS in Anlehnung an Schuh, Brosze & Brandenburg (2012, S. 21) und markierte Schwerpunkte von Kapitel 2.2	10
Abbildung 2-4: Soll-Prozess des Auftragsmanagements in der variantenreichen Produktion in Anlehnung an Schuh, Schmidt & Helmig (2012, S. 143, 172)	12
Abbildung 2-5: Überblick über die Einteilung von Präferenzintegrationsmodellen und Hervorhebung der im Kontext einer Lieferzeit-Preisgestaltung relevanten Elemente	18
Abbildung 2-6: Konzeptionelles Modell des Reinforcement Learning nach (Sutton & Barto 2018, S. 48)	25
Abbildung 2-7: Darstellung von relevanten Aktivierungsfunktionen	29
Abbildung 2-8: Grundlegende Architektur eines vorwärtsgerichteten neuronalen Netzes	30
Abbildung 2-9: Klassifikation von Reinforcement Learning Verfahren	31
Abbildung 2-10: Actor Critic Modell in Anlehnung an Scharf, Helfenstein & Jäger (2021, S. 125)	33
Abbildung 5-1: Überblick zum konkretisierten Lösungsmodell	61
Abbildung 5-2: Überblick über Modellkomponenten und Interaktion auf Netzwerkebene mit Einordnung in die Umwelt des Lösungsmodells	64

Abbildung 5-3: Zugrundeliegende Nutzenfunktion $u_i$ des Kunden in Anlehnung an A_Ochs (2021)	67
Abbildung 5-4: Effekt einer Änderung von $d_{i,v}^u$	71
Abbildung 5-5: Darstellung der acht verschiedenen Kundentypen	71
Abbildung 5-6: Übersicht über Modellkomponenten der Standortebene mit Einordnung in die Umwelt des Lösungsmodells	76
Abbildung 5-7: Beispieldarstellung eines azyklischen Graphen als Repräsentation eines Fertigungsplans	80
Abbildung 5-8: Darstellung des Kaskadeneffekts bei der Einplanung von Aufträgen	83
Abbildung 5-9: Darstellung der beiden Sichtweisen auf den Agenten	85
Abbildung 5-10: Belohnungswerte bei Beendigung einer Kundeninteraktion	86
Abbildung 5-11: Belohnungswerte nach Beendigung eines Simulationstages	87
Abbildung 5-12: Belohnungswerte nach Beendigung eines Simulationsjahres	87
Abbildung 5-13: Berechnung der Belohnungswerte in der zeitlichen Abfolge	88
Abbildung 5-14: Nutzenfunktion in Kombination mit Lieferzeit-Preis-Funktion	95
Abbildung 5-15: Einordnung des Kapitels 5.3.4 in die Sichtweisen	98
Abbildung 5-16: Architektur des Agenten	100
Abbildung 5-17: Überblick zum Aufbau des Versuchsmoduls	104
Abbildung 5-18: Vorgehen bei der Hyperparameterbestimmung	106
Abbildung 5-19: Schematische Darstellung der Architektur	113
Abbildung 6-1: Nachfrageverlauf im ersten Anwendungsfall mit Extrapolationen	117
Abbildung 6-2: Übersicht über die Simulationsphasen im Anwendungsfall 1	124

Abbildung 6-3: Verteilung der Belohnungswerte in der Hyperparameteroptimierung	125
Abbildung 6-4: Darstellung verschiedener Nachfrageverläufe	133
Abbildung 6-5: Darstellung der Gewinndifferenzebenen zwischen Agent und Status Quo mit statischer Lieferzeit	135
Abbildung 6-6: Darstellung der Gewinnebenen des Agenten	137
Abbildung 6-7: Darstellung der Servicegrad-Differenzebenen zwischen Agent und Status Quo mit statischer Lieferzeit	138
Abbildung 6-8: Darstellung der Servicegrad-Ebenen des Agenten	139
Abbildung 6-9: Übersicht über die Simulationsphasen im Anwendungsfall 2	146
Abbildung 6-10: Verteilung der Belohnungswerte in der Hyperparameteroptimierung	147
Abbildung A1-1: Vorgehensmodell zur Durchführung von Simulationsstudien nach VDI 3633 Blatt 1	XII
Abbildung A7-1: Schematische Gegenüberstellung von Gitter- und Zufallssuche	XVII
Abbildung A26-1: Dreiecksverteilung der Stückzahl verkaufter ADN-Kompaktzylinder nach Kolbendurchmesser in mm	XLII

## Tabellenverzeichnis

Tabelle 3-1: Übersicht über die Anforderungen an den Lösungsansatz	35
Tabelle 3-2: Überblick zum Forschungsdefizit	51
Tabelle 5-1: Erläuterung zu den Variablen der Gewinnrechnung	57
Tabelle 5-2: Grenzen der Nutzenfunktionsparameter	70
Tabelle 5-3: Übersicht zur Kundenkomponente	72
Tabelle 5-4: Übersicht zur Nachfragekomponente	73
Tabelle 5-5: Übersicht zur Produktkomponente	73
Tabelle 5-6: Übersicht zur Kundenauftragskomponente	74
Tabelle 5-7: Übersicht zur Netzwerkkomponente	75
Tabelle 5-8: Übersicht zur Stationskomponente	78
Tabelle 5-9: Übersicht zur Stationstypkomponente	79
Tabelle 5-10: Übersicht zur Produktionsauftragskomponente	80
Tabelle 5-11: Übersicht zur Fertigungsplankomponente	81
Tabelle 5-12: Übersicht zur Standortkomponente	84
Tabelle 5-13: Überblick über die Zustandsparameter	93
Tabelle 5-14: Grenzen der Funktionsparameter	96
Tabelle 5-15: Übersicht über die variierbaren Steuerfaktoren der Hyperparameteroptimierung inkl. Eingrenzung aus manueller Suche	107
Tabelle 5-16: Übersicht über konstant bestimmte Steuerfaktoren	107
Tabelle 5-17: Übersicht über anwendungsfallspezifische Steuerfaktoren	108
Tabelle 5-18: Übersicht über Rauschfaktoren	108
Tabelle 5-19: Umgang mit verschiedenen Auftragslasten in Anwendungsfällen	110
Tabelle 6-1: Überblick über durchschnittliche Verkaufspreise	118
Tabelle 6-2: Überblick zu den Kundentypen	119
Tabelle 6-3: Spezifische Steuerfaktoren im ersten Anwendungsfall	123

Tabelle 6-4: Konstant gehaltene Rauschfaktoren im ersten Anwendungsfall	123
Tabelle 6-5: Zufallsbasierte Rauschfaktoren im ersten Anwendungsfall	123
Tabelle 6-6: Vergleich der Konfigurationen in Szenario 1 (Statische Präferenzen)	126
Tabelle 6-7: Vergleich der Konfigurationen in Szenario 2 (Drift zu Kurzfristigkeit)	126
Tabelle 6-8: Vergleich der Konfigurationen in Szenario 3 (Sinkende Preisbereitschaft)	127
Tabelle 6-9: Ergebnisse – statische Lieferzeit – Szenario 1 – Anwendungsfall 1	128
Tabelle 6-10: Hypothesentest geg. statische Lieferzeit – Szenario 1 – Anwendungsfall 1	128
Tabelle 6-11: Ergebnisse – flexible Lieferzeit – Szenario 1 – Anwendungsfall 1	128
Tabelle 6-12: Hypothesentest geg. flexible Lieferzeit – Szenario 1 – Anwendungsfall 1	129
Tabelle 6-13: Ergebnisse – statische Lieferzeit – Szenario 2 – Anwendungsfall 1	129
Tabelle 6-14: Hypothesentest geg. statische Lieferzeit – Szenario 2 – Anwendungsfall 1	129
Tabelle 6-15: Ergebnisse – flexible Lieferzeit – Szenario 2 – Anwendungsfall 1	130
Tabelle 6-16: Hypothesentest geg. flexible Lieferzeit – Szenario 2 – Anwendungsfall 1	130
Tabelle 6-17: Ergebnisse – statische Lieferzeit – Szenario 3 – Anwendungsfall 1	130
Tabelle 6-18: Hypothesentest geg. statische Lieferzeit – Szenario 3 – Anwendungsfall 1	130
Tabelle 6-19: Ergebnisse – flexible Lieferzeit – Szenario 3 – Anwendungsfall 1	131

Tabelle 6-20: Hypothesentest geg. flexible Lieferzeit – Szenario 3 – Anwendungsfall 1	131
Tabelle 6-21: Darstellung der Verbesserung zum Status Quo durch Mittelwertvergleich	132
Tabelle 6-22: Übersicht über die durchschnittliche Warteschlangenlänge pro Station im Stresstest	134
Tabelle 6-23: Überblick über die Parametervariationen in der erweiterten Betrachtung	135
Tabelle 6-24: Verkaufspreis und -volumen des Pneumatikzylinders per Durchmesser	141
Tabelle 6-25: Kundentypen des Anwendungsfalls	141
Tabelle 6-26: Überblick zur Gesamtfertigungszeit und den Fertigungszeiten pro Variante	142
Tabelle 6-27: Überblick über die Maschinenaufteilung über die Standorte	143
Tabelle 6-28: Investitionstabelle für Maschinen aufgeschlüsselt nach Standorten	144
Tabelle 6-29: Spezifische Steuerfaktoren im zweiten Anwendungsfall	145
Tabelle 6-30: Konstant gehaltene Rauschfaktoren im zweiten Anwendungsfall	145
Tabelle 6-31: Zufallsbasierte Rauschfaktoren im zweiten Anwendungsfall	146
Tabelle 6-32: Vergleich der Konfigurationen in Szenario 1 (Statische Präferenzen)	148
Tabelle 6-33: Vergleich der Konfigurationen in Szenario 2 (Drift zu Kurzfristigkeit)	149
Tabelle 6-34: Vergleich der Konfigurationen in Szenario 3 (Sinkende Preisbereitschaft)	150
Tabelle 6-35: Darstellung der Verbesserung zum Status Quo durch Mittelwertvergleich	151
Tabelle A1-1: Unterscheidung von Simulationsmodelleigenschaften in Anlehnung an Gutenschwager & Rabe et al. (2017, S. 16)	XI

Tabelle A8-1: Alternative Softwarepakete für die ereignisdiskrete Simulation	XIX
Tabelle A8-2: Alternative Softwarepakete für das Reinforcement Learning	XIX
Tabelle A9-1: Überblick über alle möglichen Parameter einer Konfigurationsdatei	XXI
Tabelle A18-1: Maschinenspezifische Prozesszeiten für Toleranzklasse E2 (Teil 1)	XXXIII
Tabelle A18-2: Maschinenspezifische Prozesszeiten für Toleranzklasse E2 (Teil 2)	XXXIV
Tabelle A28-1: Überblick über die Bearbeitungszeitfaktoren	XLIV
Tabelle A28-2: Übersicht über die Fertigungszeiten im zweiten Anwendungsfall	XLV
Tabelle A31-1: Ergebnisse – statische Lieferzeit – Szenario 1 – Anwendungsfall 2	XLVII
Tabelle A31-2: Hypothesentest geg. statische Lieferzeit – Szenario 1 – Anwendungsfall 2	XLVIII
Tabelle A31-3: Ergebnisse – flexible Lieferzeit – Szenario 1 – Anwendungsfall 2	XLVIII
Tabelle A31-4: Hypothesentest geg. flexible Lieferzeit – Szenario 1 – Anwendungsfall 2	XLVIII
Tabelle A31-5: Ergebnisse – statische Lieferzeit – Szenario 2 – Anwendungsfall 2	XLIX
Tabelle A31-6: Hypothesentest geg. statische Lieferzeit – Szenario 2 – Anwendungsfall 2	XLIX
Tabelle A31-7: Ergebnisse – flexible Lieferzeit – Szenario 2 – Anwendungsfall 2	XLIX
Tabelle A31-8: Hypothesentest geg. flexible Lieferzeit – Szenario 2 – Anwendungsfall 2	L
Tabelle A31-9: Ergebnisse – statische Lieferzeit – Szenario 3 – Anwendungsfall 2	L

Tabelle A31-10: Hypothesentest geg. statische Lieferzeit – Szenario 3 –  
Anwendungsfall 2

L

Tabelle A31-11: Ergebnisse – flexible Lieferzeit – Szenario 3 –  
Anwendungsfall 2

LI

Tabelle A31-12: Hypothesentest geg. flexible Lieferzeit – Szenario 3 –  
Anwendungsfall 2

LI

## Anhang

### A1 Ausführungen zu den Grundlagen der Simulation im Produktionsumfeld

Nachfolgend werden ergänzende Grundlagen zur Simulation im Produktionsumfeld dargelegt, die zum Verständnis der Arbeit beitragen. Hierzu wird zunächst auf Simulation im Produktionsumfeld allgemein eingegangen. Im zweiten Teil wird dann ein normtes Vorgehensmodell zur Durchführung einer Simulationsstudie erläutert.

#### A1.1 Weiterführende Grundlagen zu Simulation im Produktionsumfeld

Zu Beginn von Kapitel 2.4 wurde bereits erläutert, dass die Simulation bei der Analyse, Kontrolle, Entwicklung und Planung von Systemen sowie zur Entscheidungsunterstützung eingesetzt werden kann. Darüber hinaus kann die Simulation auch zur Optimierung eines Systems verwendet werden (Herzog 2017, S. 152). Jedoch ist die Simulation selbst kein Optimierungsverfahren (Steven 2007, S. 483). Sie kann lediglich zur Findung besserer Parameterkonfigurationen genutzt werden (VDI 3633 Blatt 1, S. 3).

Zusammenhängend mit dem Simulationsbegriff werden in der Literatur weitere zentrale Begriffe verwendet, die nachfolgend erläutert werden: Um ein *System* zu analysieren und somit Erkenntnisse über dieses zu erlangen, werden *Experimente* im Rahmen einer *Simulationsstudie* durchgeführt (VDI 3633 Blatt 1, S. 3f). Ein System bezeichnet eine Menge von in Beziehung stehenden Elementen, die sich von ihrer Umwelt abgrenzen und als Ganzes gesehen werden (Gutenschwager & Rabe et al. 2017, S. 10f; DIN 60050-351:2014-09). Ein Experiment, auch als Simulationslauf bezeichnet, ist das Nachbilden eines Systems mit einem Simulationsmodell über einen bestimmten Zeitraum (Eley 2012, S. 4). Das hierfür benötigte Simulationsmodell bildet das System auf abstrakte Art und Weise ab (Eley 2012, S. 3). Typischerweise werden in einem Modell nur die zur Erfüllung des Modellzwecks relevanten Systemelemente, -eigenschaften und -beziehungen abgebildet (Gutenschwager & Rabe et al. 2017, S. 19; Stachowiak 1973, S. 131–133). Das gesamte Vorgehen von der Modellbildung über die Durchführung von Experimenten und der Interpretation der Ergebnisse bis hin zur Übertragung auf den Anwendungsfall wird als Simulationsstudie bezeichnet. Ein Modell zur Durchführung einer Simulationsstudie wird detaillierter in Anhang A1.2 erläutert.

Wie zu Beginn des Kapitels 2.4 beschrieben, sind Simulationen bei manchen Problemstellungen die einzige Möglichkeit zur Analyse und Entscheidungsunterstützung. In jedem Fall ist es jedoch hilfreich, zur Bewertung der Eignung einer Simulation die Vorteile und Nachteile des Verfahrens zu kennen. Die Vorteile wurden bereits zu Beginn von Kapitel 2.4 beschrieben. Sie bestehen aus der Möglichkeit verschiedene Hypothesen und Szenarien zu prüfen, ohne das betrachtete System zu beeinflussen. Es ist dadurch leichter möglich Wirkzusammenhänge zu erkennen.

Diesen Vorteilen einer Simulation stehen jedoch auch Nachteile gegenüber. Zunächst entsteht ein zeitlicher und monetärer Aufwand bei der Durchführung einer Simulationsstudie (Kühn 2006, S. 26). Die Datenbeschaffung und -aufbereitung spielt zudem eine kritische Rolle bei der Simulation und kann als ein Schwachpunkt identifiziert werden (Hotz 2007, S. 13). Die Qualität der Datengrundlage hat unmittelbaren Einfluss auf die Ergebnisse und damit auf die Schlussfolgerungen der Simulationsstudie. Weiterhin ist ein spezifischer Nachteil stochastischer Simulationsmodelle, dass sie selbst nur stochastische Ergebnisse liefern, die es zu interpretieren gilt (Hotz 2007, S. 14). Hier ist die Überprüfung der statistischen Signifikanz der Ergebnisse wichtig.

Eine Klassifizierung von Simulationsmodellen, die ein System abbilden, wird anhand ihres Verhaltens vorgenommen. Zunächst können dynamische und statische Modelle unterschieden werden (Gutenschwager & Rabe et al. 2017, S. 15f). Diese Unterscheidung charakterisiert ein Modell hinsichtlich seiner Veränderungen über die Zeit. Weiter kann ein Modell deterministischer oder stochastischer Natur sein (Gutenschwager & Rabe et al. 2017, S. 18). In den stochastischen Modellen spielt der Zufall eine wichtige Rolle, wohingegen in deterministischen Simulationen keine Zufallsereignisse vorkommen. Aus dieser Eigenschaft folgt die unmittelbare Schlussfolgerung, dass stochastische Simulationsmodelle mit einer gegebenen Parametrisierung vielfach ausgeführt werden müssen, um zuverlässige Erkenntnisse über ein System zu erlangen (Eley 2012, S. 4). Deterministische Simulationsmodelle hingegen müssen hierfür nur exakt einmal ausgeführt werden (Eley 2012, S. 4). Weiter kann ein Modell terminierend oder nicht terminierend sein (Gutenschwager & Rabe et al. 2017, S. 17f). Hiermit wird unterschieden, ob ein Modell eindeutige Startbedingungen und ein natürliches Ende besitzt. Die Zustandsänderungen in einem Modell können kontinuierlich oder ereignisdiskret erfolgen (Gutenschwager & Rabe et al. 2017, S. 16). Eine Übersicht über diese Klassifikation, die die eben dargestellte Einteilung wiedergibt, findet sich in der folgenden Tabelle A1-1.

*Tabelle A1-1: Unterscheidung von Simulationsmodelleigenschaften in Anlehnung an Gutenschwager & Rabe et al. (2017, S. 16)*

Eigenschaft	Mögliche Ausprägungen	
Zeitverhalten	Statisch	Dynamisch
Zufallsverhalten	Stochastisch	Deterministisch
Terminierung	Terminierend	Nicht-terminierend
Art der Zustandsänderung	Ereignisdiskret	Kontinuierlich

Um nun den Bogen von den allgemeingültigen Simulationsmodelleigenschaften zurück zur spezifischen Simulation im Produktionsumfeld zu schlagen, lässt sich untersuchen, welche Simulationsmodelle von ihrer Natur her im Produktionsumfeld dominierend sind. Hier zeigt sich, dass vor allem die ereignisdiskrete Simulation anerkannt und verbreitet ist (Steven 2007, S. 483; Gutenschwager & Rabe et al. 2017, S. 24f; März & Krug et al. 2011, S. 14). Obgleich die Modellierung besonders von Netzwerken in einem Simulationsmodell eine Herausforderung ist, hat sich die ereignisdiskrete Simulation vielfach als alternativlos gezeigt (Wadhwa & Mishra et al. 2010, S. 149). Dies liegt an der Möglichkeit, Netzwerke mit vielen Produktionsstandorten im Modell zu modularisieren, um die Komplexität zu beherrschen (Wadhwa & Mishra et al. 2010, S. 149). Nichtsdestotrotz fokussiert sich die Forschung auf die Betrachtung eines fokalen Standorts mit mehreren Verbindungen ins Netzwerk (Wadhwa & Mishra et al. 2010, S. 149). Eine der Ursachen kann in der vorhandenen Informationsasymmetrie gesehen werden, die die vollständige und detaillierte Abbildung eines Netzwerks erschwert und auch nicht die Praxis aus Sicht eines Unternehmens widergäbe (Váncza & Monostori et al. 2011, S. 801; Vosooghidizaji, Taghipour & Canel-Depitre 2020, S. 1805).

## **A1.2 Vorgehensmodell zur Durchführung einer Simulationsstudie**

Mithilfe eines Vorgehensmodells für Simulationsstudien soll eine fachgerechte Durchführung gewährleistet werden, um somit belastbare Ergebnisse zu erhalten (Rabe, Spiekermann & Wenzel 2008, S. 27). In der Literatur existiert eine ganze Reihe an Vorgehensmodellen (vgl. Rabe, Spiekermann & Wenzel 2008, S. 27, 29-32), jedoch hat sich im deutschsprachigen Raum das Vorgehensmodell nach der Richtlinie VDI 3633 etabliert (Rabe, Spiekermann & Wenzel 2008, S. 27).

Das Vorgehensmodell nach der Richtlinie VDI 3633 besteht aus sechs Hauptschritten, welche in Abbildung A1-1 dargestellt sind und im Folgenden erläutert werden.

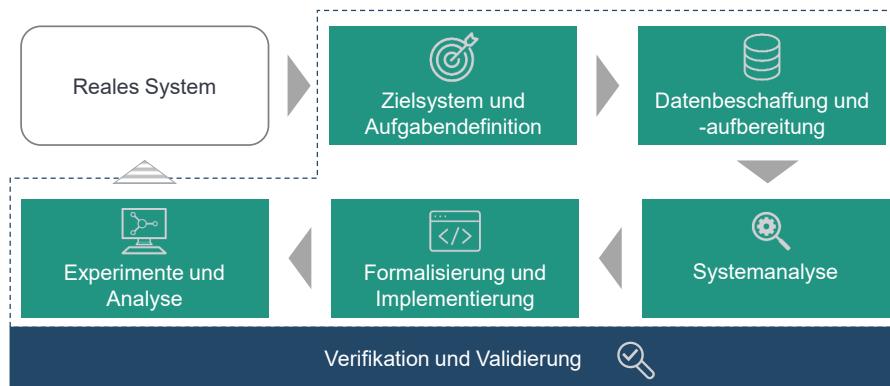
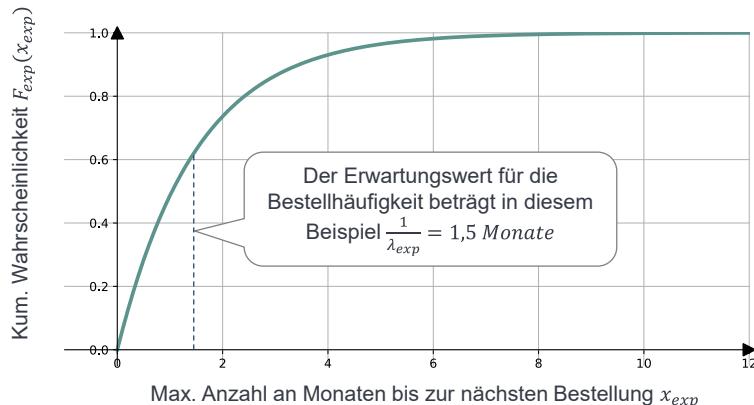


Abbildung A1-1: Vorgehensmodell zur Durchfhrung von Simulationsstudien nach VDI 3633 Blatt 1

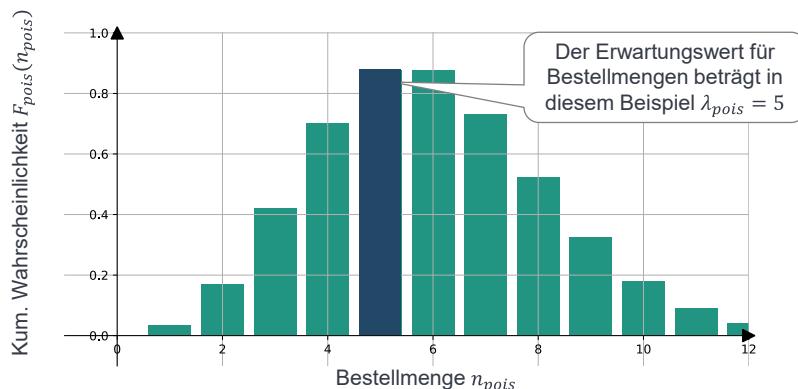
Im ersten Schritt des Vorgehensmodells geht es um die Aufgabendefinition und das Zielsystem. Das Zielsystem beinhaltet üblicherweise ein Hauptziel und daraus abgeleitete Teilziele. Nach der Entwicklung des Zielsystems wird parallel mit der Datensammlung und -aufbereitung begonnen. In diesem Schritt werden die für die Simulation notwendigen Daten zusammengestellt. Generell wird hierbei mit einer groben, analytischen Schätzung auf Basis der Daten begonnen, um ein erstes Bild des Systems zu erhalten, bevor die Daten im Detail erhoben und verarbeitet werden. Dieses Vorgehen steht im Einklang mit dem Schritt Systemanalyse, welche parallel zu Beginn der Datensammlung durchgeführt wird. Die Systemanalyse mündet in ein Konzeptmodell. Im Rahmen der Systemanalyse entsteht folglich eine erste Abbildung des Systems, wodurch Detailgrad und Systemgrenzen festgelegt werden. Das Konzeptmodell ist jedoch noch nicht experimentierbar. Auf Basis der Erkenntnisse der Systemanalyse wird das Konzeptmodell deshalb im Rahmen der formalen Modellierung und Implementierung in ein experimentierbares Simulationsmodell überführt. Die Modellierung und Implementierung sind hierbei eng verzahnt. Ein wichtiger Schritt im Vorgehensmodell ist die Verifikation und Validierung. Sie verläuft über den gesamten Zeitraum der Simulationsstudie parallel zu allen Schritten und soll die Korrektheit und Verwendbarkeit aller (Teil-)Ergebnisse sicherstellen. Mit dem fertigen Simulationsmodell können nun Experimente und Analysen durchgeführt werden. Die erste Aufgabe bei diesem Schritt ist, die Parametervariationen so zu planen, dass das Simulationsziel mit möglichst wenigen Durchläufen erreicht werden kann. Gleichzeitig ist der Weg zum Simulationsziel meist

ein systematisches Probieren und kann daher nicht deterministisch geplant werden. Nach der Durchführung der Experimente müssen die gewonnenen Experimentergebnisse ausgewertet und interpretiert werden. Die Auswertung der Experimentergebnisse ist von besonderer Bedeutung, da sie wesentlich die Qualität der Rückschlüsse beeinflusst. (VDI 3633 Blatt 1)

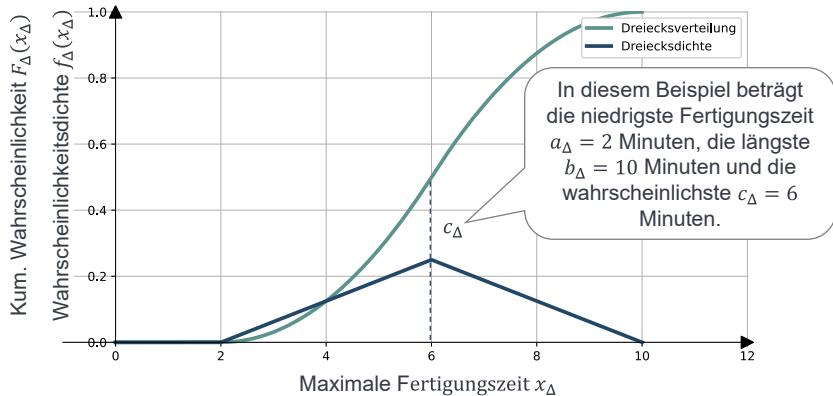
## A2 Beispiel einer Exponentialverteilung als Bestellhäufigkeit



## A3 Beispiel einer Poissonverteilung als Bestellmenge



## A4 Beispiel einer Dreiecksverteilung als Fertigungszeit



## A5 Algorithmus zur Ablaufplanung

Mit den aufgestellten Regeln aus Kapitel 5.2.2.5 stellt sich die Heuristik wie folgt dar: Zunächst werden die gemäß Fertigungsplan in Frage kommenden Stationen nach ihrer Auslastung aufsteigend sortiert. Die Intralogistik wird hierbei vernachlässigt. Dann werden die Schonstationen übersprungen. Als nächstes wird über die verbleibenden Stationen iteriert und jeweils geprüft, ob der Produktionsauftrag noch in die Warteschlange passt, ohne sich oder andere Produktionsaufträge zu verspäten. Kann der Produktionsauftrag in keine aktive Station eingeplant werden, wird eine Schonstation verwendet. Die Entscheidung, ob ein Produktionsauftrag noch in die Warteschlange passt, wird durch den Vergleich der Schlupfzeit des einzusortierenden Produktionsauftrags mit den Schlupfzeiten der wartenden Produktionsaufträge erreicht. Die Produktionsaufträge sind nach aufsteigender Schlupfzeit sortiert. Der zusätzliche Produktionsauftrag wird in die Warteschlange einsortiert, wenn gilt:

$$t_{schlupf,pa,o}^- < t_{schlupf,pa,o} < t_{schlupf,pa,o}^+ \quad \text{Formel A5-1}$$

$t_{schlupf,pa,o}$  gibt hierbei die Schlupfzeit des aktuell betrachteten Produktionsauftrags  $pa_{aktuell,o}$  an.  $t_{schlupf,pa,o}^-$  bzw.  $t_{schlupf,pa,o}^+$  ist die Schlupfzeit des Produktionsauftrags, der unmittelbar vor bzw. unmittelbar nach dem aktuell betrachteten Produktionsauftrag in der Warteschlange  $W_{Schlange,o}$  der aktuell betrachteten Station käme.

Die Schlupfzeit wird durch folgende Formel bestimmt:

$$t_{schlupf,pa,o} = x_{i,v} - t_{i,pa} \text{ für } pa \in W_{Schlange,o}, pa \in PA_v \quad \text{Formel A5-2}$$

Obgleich die Formel zunächst trivial erscheint, verbirgt sich dahinter ein Problem: Erstens kann die verbleibende Durchlaufzeit  $t_{i,pa}$  zum Zeitpunkt der Planung stochastisch sein. Zweitens ist die verbleibende Durchlaufzeit bei einem mehrstufigen Prozess abhängig von allen Folgeschritten. Dies stellt ebenfalls eine Herausforderung dar, sofern die Folgeschritte nicht deterministisch sind, es also verschiedene Schrittfolgen geben kann.

Für die Planung mit nicht deterministischen Produktionsaufträgen kann mit einer Konfidenz gearbeitet werden, um mit einer definierten Wahrscheinlichkeit Verspätungen zu vermeiden. Voraussetzung hierfür ist die Berechnung der Fertigungszeitverteilung. Diese lässt sich aus der Abfolgen aller Möglichkeiten im Fertigungsplan des aktuellen Produktionsauftrags ableiten. Da der Fertigungsplan diskrete Ereignisse abbildet, ergibt sich auch die resultierende Gesamtverteilung möglicher verbleibender Fertigungszeiten als diskrete Menge. Diese diskrete Gesamtverteilung lässt sich nun per Rückwärtsinduktion bestimmen. Mithilfe eines Konfidenzniveaus kann dann eingestellt werden, welche Fertigungszeit mindestens anzunehmen ist. Um dieses Vorgehen zu veranschaulichen, sei hier folgendes, vereinfachtes Beispiel gegeben: Ein Produktionsauftrag hat einen bestimmten Zustand. Gemäß des zugehörigen Fertigungsplans folgt ein Transformationsschritt mit einer Dauer von 2 min. Mit 50 % Wahrscheinlichkeit muss danach noch eine Nacharbeit erfolgen, welche 1 min in Anspruch nimmt. Andernfalls ist der Produktionsauftrag abgeschlossen. Die Fertigungszeit ist nun mit 50 % 2 min und mit 50 % 3 min lang. Wenn die Konfidenz auf  $\geq 50\%$  liegt, so muss von einer verbleibenden Fertigungszeit von 3 min ausgegangen werden, andernfalls von 2 min.

Sobald ein neuer Produktionsauftrag nun auf diese Art und Weise in eine Warteschlange eingeschoben wird, ist es wichtig zu beachten, dass die verbleibende Durchlaufzeit aller nachfolgenden Produktionsaufträge entsprechend um die Verschiebung aktualisiert wird.

## A6 Ergänzung zur Bestimmung des Nutzenmaximums

$$\frac{du_i}{dx_{i,v}} = u'_i = -a_{i,v}e^{(x_{i,v}-b_{i,v})} + a_{i,v} + K_{i,v}e^{(L_{i,v}-x_{i,v})} - K_{i,v} = 0 \quad \text{Formel 5-37}$$

$$\begin{aligned} 0 &= -a_{i,v}e^{(x_{i,v}-b_{i,v})} + K_{i,v}e^{(L_{i,v}-x_{i,v})} + a_{i,v} - K_{i,v} \\ &= -a_{i,v}e^{x_{i,v}}e^{-b_{i,v}} + K_{i,v}e^{L_{i,v}}e^{-x_{i,v}} + a_{i,v} - K_{i,v} \end{aligned} \quad \text{Formel A6-1}$$

$$0 = -a_{i,v} \cdot e^{2x_{i,v}} \cdot e^{-b_{i,v}} + K_{i,v} \cdot e^{L_{i,v}} + (a_{i,v} - K_{i,v}) \cdot e^{x_{i,v}} \quad \text{Formel A6-2}$$

$$\begin{aligned} 0 &= -a_{i,v} \cdot v^2 \cdot e^{-b_{i,v}} + K_{i,v} \cdot e^{L_{i,v}} + (a_{i,v} - K_{i,v}) \cdot v \\ &\text{mit } v = e^{x_{i,v}} \text{ als Substitut folgt} \end{aligned} \quad \text{Formel A6-3}$$

$$v_{1/2} = \frac{-(a_{i,v} - K) \pm \sqrt{(a_{i,v} - K_{i,v})^2 + 4 \cdot a_{i,v} \cdot e^{-b_{i,v}} \cdot K_{i,v} \cdot e^{L_{i,v}}}}{-2 \cdot a_{i,v} \cdot e^{-b_{i,v}}} \quad \text{Formel A6-4}$$

$$\begin{aligned} x_{i,v,1/2}^{opt} \\ = \ln \left( \frac{(a_{i,v} - K_{i,v}) \mp \sqrt{(a_{i,v} - K_{i,v})^2 + 4 \cdot a_{i,v} \cdot e^{-b_{i,v}} \cdot K_{i,v} \cdot e^{L_{i,v}}}}{2a_{i,v}e^{-b_{i,v}}} \right) \end{aligned} \quad \text{Formel 5-38}$$

## A7 Erläuterungen zur Wahl der zufallsbasierten Suche

Als Suchstrategie einer Hyperparameteroptimierung kann in Analogie zu einer teilstufiellen Versuchsplanung eine Gittersuche verwendet werden. Dabei hat sich jedoch gezeigt, dass eine zufallsbasierte Suche im Raum möglicher Steuerfaktorausprägungen in den meisten Fällen zu schnelleren und besseren Ergebnissen führt als eine vorstrukturierte Gittersuche (Bergstra & Bengio 2012, S. 302). Die Hyperparameteroptimierung mithilfe der Gittersuche und mithilfe der zufallsbasierten Suche für einen zweidimensionalen Raum ist in der nachfolgenden Abbildung schematisch dargestellt.

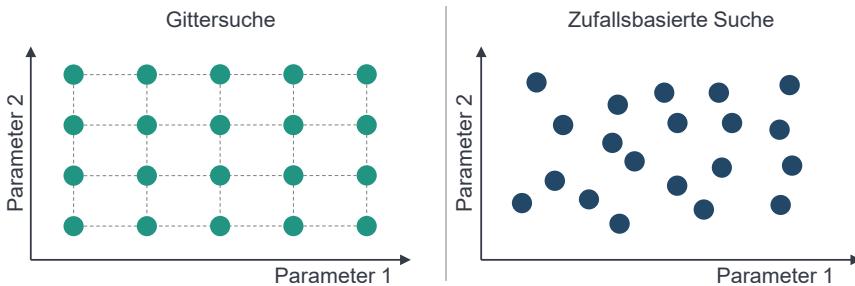


Abbildung A7-1: Schematische Gegenüberstellung von Gitter- und Zufallssuche

Neben dem Erzielen besserer Ergebnisse hat die zufallsbasierte Suche gegenüber einer Gittersuche weitere Vorteile nach Bergstra & Bengio (2012, S. 302):

- Experimente können zu jedem Zeitpunkt abgebrochen werden und es ergibt sich eine vollständige, gültige Versuchsreihe. Dies ist bei klassischen Versuchsplänen nicht möglich.
- Falls neue Rechenressourcen verfügbar werden, können zusätzliche Experimente als Ergänzung gestartet werden, ohne dass eine Abstimmung mit den laufenden Experimenten nötig werden würde.
- Alle Versuche laufen asynchron, wodurch Komplexität in Form von Abhängigkeiten reduziert wird.
- Falls ein Experiment auf einer Rechenressource scheitern sollte, zum Beispiel aufgrund von Speicherlimitationen, scheitert nicht die gesamte Versuchsreihe, sondern die bestehenden Ergebnisse können verwendet werden und das Experiment kann unmittelbar fortgesetzt werden.

Aus diesen Gründen wird bei der Durchführung auf eine zufallsbasierte Suche gesetzt.

## A8 Überlegungen zur Softwarearchitektur

Ein zentraler Bestandteil, auf welchen diese Arbeit aufsetzt, ist die Simulation und das Reinforcement Learning. Um den Lösungsansatz prüfen zu können, war es notwendig das Lösungsmodell auch softwareseitig umzusetzen. Die Entwicklung der Software birgt dabei viele eigene Herausforderungen und Fragestellungen, die mit dem eigentli-

chen Ansatz zunächst keinen Bezug aufweisen. Nachfolgend werden die grundsätzlichen Überlegungen und Entscheidungen dargelegt, welche den Softwaredemonstrator beeinflusst haben.

Zu Beginn der Softwareentwicklung war aufgrund der Überlegungen in Kapitel 5.1 klar, dass eine ereignisdiskrete Simulation und ein Actor Critic Reinforcement Learning zum Einsatz kommen sollen. Der Schwerpunkt der Arbeit lag nicht auf der grundlegend neuen Entwicklung von Simulation und Reinforcement Learning. Deshalb war es ein Ziel der Überlegungen, möglichst viele generische Aspekte über existierende Softwarelösungen abzubilden. Für die ereignisdiskrete Simulation bieten sich zahlreiche, bestehende Lösungen an. Es gibt sowohl quelloffene Softwarepakete als auch proprietäre Softwarelösungen, wie beispielsweise *Anylogic* und *PlantSim*, welche die Kriterien einer ereignisdiskreten Simulation im Produktionsumfeld prinzipiell erfüllen. Auf der anderen Seite gibt es ebenso eine Mehrzahl an Möglichkeiten zur Umsetzung eines Actor Critic Reinforcement Learning Agenten. Hier kommt die Rechengeschwindigkeit zum Tragen: Zur Umsetzung einer praktikablen Software ist es notwendig, die Rechenoperationen so effizient wie möglich zu gestalten. Deshalb ist das Zusammenspiel der beiden Komponenten Simulation und Agent kritisch. Es ist klar, dass beide Komponenten hochfrequent miteinander interagieren müssen und daher diese Interaktion einen großen Einfluss auf die Rechenzeit haben wird. Proprietäre Softwarelösungen bieten zwar gewisse Schnittstellen, um eine Integration mit anderen Softwarepaketen zu ermöglichen. Diese Schnittstellen sind jedoch stark eingeschränkt. Dieser Aspekt steht einer hochfrequenten Interaktion entgegen. Darüber hinaus stellen proprietäre Lösungen ein Risiko dar, weil sich jeder Zeit eine zunächst unbekannte Einschränkung ergeben kann, welche die Umsetzung zum Scheitern verurteilt. Ein drittes Argument ist die Verfügbarkeit für die Forschungswelt. Proprietäre Softwarelizenzen verursachen Kosten und erschweren anderen Forschern die Ergebnisse zu prüfen bzw. mit Ihnen weiter arbeiten zu können. Daher wurden proprietäre Softwarelösungen frühzeitig ausgeschlossen.

Das Argument der hochfrequenten Interaktion legt nahe, dass sowohl Simulation als auch Agent in derselben Programmiersprache entwickelt werden, um beide unmittelbar im selben Code miteinander zu verbinden. Daher war es notwendig eine Lösung zu finden, bei der die verwendeten Softwarepakete in derselben Programmiersprache zur Verfügung stehen. Auf der Seite der quelloffenen Simulationsbibliotheken ergaben sich folgende, in Frage kommenden Optionen:

*Tabelle A8-1: Alternative Softwarepakete für die ereignisdiskrete Simulation*

Softwarepaket	Programmiersprache
SimJS	Javascript
Gekko	Python
SimPy	P
Tortuga	Java
SimSharp	C#
SimJava	Java
SharpSim	C#

Weitere existierende Alternativen wurden frühzeitig verworfen. Hierbei wurden folgende Kriterien untersucht:

- Wird das Softwarepaket noch gepflegt?
- Ist es als Softwarebibliothek (und nicht als grafische Software) gedacht?
- Bildet es essenzielle Funktionen ab (bspw. Zufallszahlengenerator)?

Für die Entwicklung des Agenten stellten sich folgende Alternativen dar:

*Tabelle A8-2: Alternative Softwarepakete für das Reinforcement Learning*

Softwarepaket	Programmiersprache
Tensorflow.js	Javascript
Tensorflow	Python
KerasRL	Python
PyQLearning	Python
MAgent	Python
RLGraph	Python
ML.NET	C#
Weka	Java
Mahout	Java

Es handelt sich hierbei nur um eine Auswahl einiger Softwarepakete ohne Anspruch auf Vollständigkeit. Es zeigt sich jedoch, dass die Mehrzahl an Softwarepaketen in *Python* geschrieben ist. Durch die starke Repräsentanz von *Python* im Bereich des maschinellen Lernens ist davon auszugehen, dass die Softwarepakete eine gute Dokumentation haben. *Python* hat jedoch einen wesentlichen Nachteil: Es besitzt keine Typensicherheit. Hierdurch ist erst zur Laufzeit bekannt, welcher Datentyp eine Variable

annimmt. Dieser Aspekt ist nachteilhaft, weil hierdurch automatische Codeanalysen zur Fehlerprävention eingeschränkt werden. In der Folge muss deutlich mehr Zeit für die manuelle Fehlersuche eingeplant werden. Hinzu kommt, dass sich Software besonders in der Forschung häufig wandelt. Durch die Änderung der Software kann es zu unvorhersehbaren Folgefehlern kommen, deren Behebung zusätzliche Zeit kostet. Typensichere Sprachen wie *C#* oder *Java* haben dieses Problem nicht. Jedoch bieten die Softwarepakte dieser Sprachen unzureichend Möglichkeiten zur Umsetzung von Reinforcement Learning. An dieser Stelle soll nochmals das Argument der Rechenzeit aufgegriffen werden. Die Rechenzeit ist abhängig von der Effizienz der entwickelten Software auf der einen Seite, und vom Hardwareeinsatz auf der anderen Seite. Hierbei gilt zu beachten, dass eine Skalierung der Rechenhardware nur horizontal und nicht vertikal kostengünstig möglich ist. Damit ist gemeint, dass die Kosten für mehr Hardware derselben Art linear steigen. Eine Steigerung der einzelnen Hardware ist nur sehr begrenzt möglich und die Kosten steigen nicht linear. Daher sollte ein Augenmerk auf die Möglichkeit zur horizontalen Skalierung liegen. Hierbei bietet es sich an, die Software als Serverlösung zu gestalten und in einen Container zu kapseln. In der Folge ist die Verteilung auf derselben Hardware auf mehrere Rechenkerne und auf mehreren Hardwareinstanzen einfach möglich. Im Umfeld der Serverapplikationen tut sich *NodeJs* her vor. Es handelt sich dabei um eine *javascript*-basierte Laufzeitumgebung. Auch *Python* bietet Lösungen zur Entwicklung von Serverapplikationen. Für *Javascript* existiert jedoch die erweiterte Programmiersprache *Typescript*. Diese erlaubt die Umsetzung von typensicherer Programmierung in *Javascript*. Aufgrund dieser Vorteile und der existierenden Lösungen wurde *Typescript* in Verbindung mit *NodeJs* ausgewählt. Die ereignisdiskrete Simulation basiert hierbei auf *SimJS*. Der Agent basiert auf *Tensorflow.js*. Das Softwarepaket wurde jedoch maßgeblich weiterentwickelt.

Neben der ereignisdiskreten Simulation und dem Agenten ist eine wichtige Komponente die Datenbank zur Speicherung der Experimentergebnisse. Hier wurden einfache Datendateien, lokale Datenbanken und serverbasierte Datenbanken verglichen. Nachteilig bei den einfachen Datendateien ist die Effizienz der Speicher- und Leseoperationen. Für ein einmaliges Laden von Parametern zu Beginn einer Ausführung sind Datendateien aufgrund ihrer einfachen Handhabbarkeit vorteilhaft. Während der Ausführung sind die Effizienz von Speicher- und Leseoperationen für die Lösung kritisch, da im Rahmen des Versuches mehrere Gigabyte an Daten anfallen können. Auf der an-

deren Seite bringt eine serverbasierte Datenbank die beste Leistung, aber auch zusätzliche Komplexität in die Software: Die REST-basierte Schnittstelle muss entsprechend implementiert und abgesichert werden. Dagegen bietet ein lokales Datenbanksystem einen Kompromiss beider Lösungen: Speicher- und Leseoperationen sind effizienter als bei einfachen Datendateien, jedoch entfällt die komplexe Schnittstelle zu einer vollwertigen, serverbasierten Datenbank. Im Umfeld der lokalen Datenbanken hat sich *SQLite* hervorgetan, welches auf dem weit verbreiteten *SQL*-Standard aufsetzt. Daher fiel die Wahl auf *SQLite*.

Die hier dargelegten Überlegungen führen nun zu der im Kapitel 5.5.1 vorgestellten Softwarearchitektur.

## A9 Erläuterungen zur Konfigurationsdatei

Die Konfigurationsdatei hat für die Umgebung teilweise spezifische Parameter und teilweise allgemeingültige Parameter. Eine Umgebung ist abhängig vom Anwendungsfall und dem Experimentablauf. Die spezifischen Parameter einer Umgebung können auch bei Ausführung anderer Umgebungen gesetzt werden, haben jedoch dann keine Auswirkung. Das Versuchsmodul projiziert die übergebene Konfigurationsdatei auf ein Standardobjekt, welches Standardwerte aller Parameter enthält. Wird ein relevanter Wert in der Konfigurationsdatei nicht gesetzt, so wird der Standardwert verwendet. Alle möglichen Parameter finden sich inklusive Erklärung in der nachfolgenden Tabelle:

*Tabelle A9-1: Überblick über alle möglichen Parameter einer Konfigurationsdatei*

Kategorie	Parameter	Datenformat	Beschreibung
Allgemeine Einstellungen	createLogReport	Boolean	Steuert die Erstellung von Log-Einträgen
	maximumReportLength	Boolean	Steuert die maximale Anzahl zu visualisierender Datenpunkte zur Laufzeit
	parallelProcessing	Boolean	Steuert, ob die Simulation auf mehreren Kernen ausgeführt werden soll
	inMemoryCache	Boolean	Steuert, ob die Datenbank zur Laufzeit im Arbeitsspeicher zwischengespeichert wird
	seed	Zahl	Der Startwert für den Zufallszahlengenerator. Bei einem Wert von -1 wird ein Wert zufällig gezogen.

Umwelteinstellungen	agentmodel	Text	Speicherplatz für das Agentenmodell im JSON-Format
	name	Text	Name des Experiments
	environment	Text	Umgebung, welche ausgeführt werden soll.
	startDate	Text	Startdatum der Simulation
	totalSimTime	Zahl	Simulationszeit, welche simuliert werden soll. -1 bedeutet unendlich lange.
	maxEvents	Zahl	Maximale Anzahl an diskreten Ereignissen, die ausgeführt werden soll. -1 bedeutet unendlich viele.
	customerAmount	Zahl	Anzahl der Kunden
	useKI	Boolean	Entscheidet, ob die künstliche Intelligenz zur Setzung der Preise verwendet wird
	replayAfterAmount	Zahl	Entscheidet, nach wie vielen Verhandlungen der Agent trainiert werden soll
	staticLeadTime	Zahl	Statische Lieferzeit im Fall statischer Preissetzung
	staticPriceRelativeSurcharge	Zahl	Preisaufschlag im Fall statischer Preissetzung
	dynamicLeadTime	Boolean	Steuert die statische Preissetzung. Bei true wird immer die angeforderte Lieferzeit des Kunden akzeptiert
	inventoryHoldingCost	Zahl	Lagerhaltungskosten (relativ zum Basispreis)
	delayPenaltyCost	Zahl	Verspätungsstrafe pro Verzugstag (relativ zum Basispreis)
	maximumDelayPenalty	Zahl	Maximale Verspätungsstrafe (relativ zum Basispreis)
	minimumParameterA	Zahl	Im Fall eines Kundenpräferenzdrifts, gibt den minimalen Wert für den Parameter a an.
	maximumParameterA	Zahl	Im Fall eines Kundenpräferenzdrifts, gibt den maximalen Parameter für a an.
	minimumParameterB	Zahl	vgl. minimumParameterA
	maximumParameterB	Zahl	vgl. maximumParameterA
	minimumParameterD	Zahl	vgl. minimumParameterA
	maximumParameterD	Zahl	vgl. maximumParameterA
	minimumParameterV	Zahl	vgl. minimumParameterA

maximumParameterV	Zahl	vgl. maximumParameterA
minimumParameterL	Zahl	Minimaler Wert zur Skalierung der Aktion des Agenten
maximumParameterL	Zahl	Maximaler Wert zur Skalierung der Aktion des Agenten
minimumParameterK	Zahl	vgl. minimumParameterL
maximumParameterK	Zahl	vgl. maximumParameterL
minimumParameterN_	Zahl	vgl. minimumParameterL
maximumParameterN_	Zahl	vgl. maximumParameterL
traffic	Zahl	Faktor zur Justierung sonstiger Aufträge im Produktionssystem, die nicht im Fokus stehen. Anteil der Simulationsdauer, nach der die Aufzeichnung begonnen wird
startOfBenchmarkTest	Zahl	
numberOfReplications	Zahl	Anzahl der Wiederholungen der Simulation mit einer Konfiguration
tardinessPenalty	Zahl	Verspätungsstrafe des Agenten
earlinessPenalty	Zahl	Verfrühangsstrafe des Agenten
nullOrderPenalty	Zahl	Strafe des Agenten für entgangene Aufträge
maximumPenalty	Zahl	Begrenzung der maximalen Bestrafungssumme (optional)
amountOfYears- ForUnusedMachineKPI	Zahl	Anzahl der Jahre bevor eine Maschine als ungenutzt identifiziert werden kann.
finalRewardPerUnu- sedMachine	Zahl	Belohnungswert für ungenutzte Kapazitäten
scenario	Zahl	Das zu simulierende Szenario für Experimente mit nur einem Szenario
scenarioList	Array	Liste an zu simulierenden Szenarien für Experimente mit mehreren Durchläufen
bDriftReductionInDays	Zahl	Reduktionsfaktor für den Kundenpräferenzdrift
dDriftReduction	Zahl	Vgl. bDriftReductionInDays
vDriftReduction	Zahl	Vgl. bDriftReductionInDays
reductionBase	Zahl	Reduktionsbasis als Schätzung der Aufträge pro Jahr
customerArray	Array Zahlen	aus Liste von Kundenanzahlen für Experimente mit mehreren Durchläufen

Agenteinstellungen	trafficArray	Array Zahlen	aus	Liste verschiedener Faktoren zur Skalierung sonstiger Aufträge für Experimente mit mehreren Durchläufen
	lengthEinschwingphase	Zahl		Länge der Einschwingphase bevor das Training des Agenten beginnt
	numActions	Zahl		Anzahl Aktionen des Agenten
	discountRate	Zahl		Der Abdiskontierungsfaktor
	minActorLearningRate	Zahl		Minimale Lernrate des Actors
	maxActorLearningRate	Zahl		Maximale Lernrate des Actors
	minCriticLearningRate	Zahl		Minimale Lernrate des Critics
	maxCriticLearningRate	Zahl		Maximale Lernrate des Critics
	actorLlambda	Zahl		Faktor für den Lernratenabbau des Actors
	criticLlambda	Zahl		Faktor für den Lernratenabbau des Critics
	actorHiddenLayerSize	Zahl		Anzahl Zwischenschichten des Actors
	actorNeuronsPerLayer	Zahl		Anzahl Neuronen pro Schicht
	criticHiddenLayerSize	Zahl		Anzahl Zwischenschichten des Critics
	criticNeuronsPerLayer	Zahl		Anzahl Neuronen pro Schicht
	includeSeparateTraining	Zahl		Einstellung, um Critic und Actor jeweils im Wechsel zu trainieren
	hpjsNumberOfIterationsPerConfig	Zahl		Hyperparameteroptimierung: Wiederholungen pro untersuchter Konfiguration
	hpjsNumberOfRandomSearches	Zahl		Hyperparameteroptimierung: Anzahl an zu untersuchenden Zufallskonfigurationen
	activationFunction	Zahl		Aktivierungsfunktion in den Zwischenschichten
	lowerBoundInitialization	Zahl		Untere Schranke für die Initialisierung der Neuronengewichte
	upperBoundInitialization	Zahl		Obere Schranke für die Initialisierung der Neuronengewichte
	switchAgentAfterAmountOfReplays	Zahl		Bei zwei Agenten: Anzahl der Trainingsphasen, bevor auf den anderen Agenten gewechselt wird

## A10 Beispiel für ein Hochpräzisionsgewicht nach Norm<sup>7</sup>



---

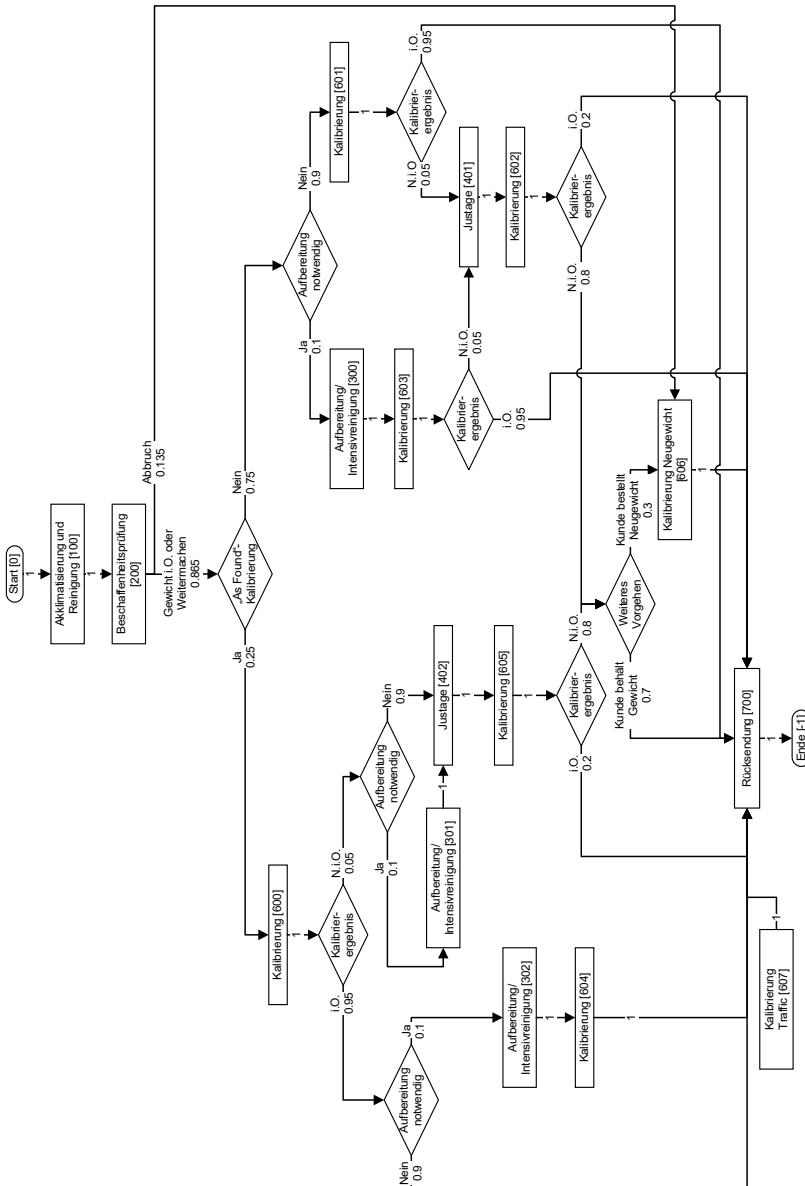
<sup>7</sup> <https://res.cloudinary.com/rsc/image/upload/Y1826541-01> [24.04.2022]

## A11 Normtabelle der OIML

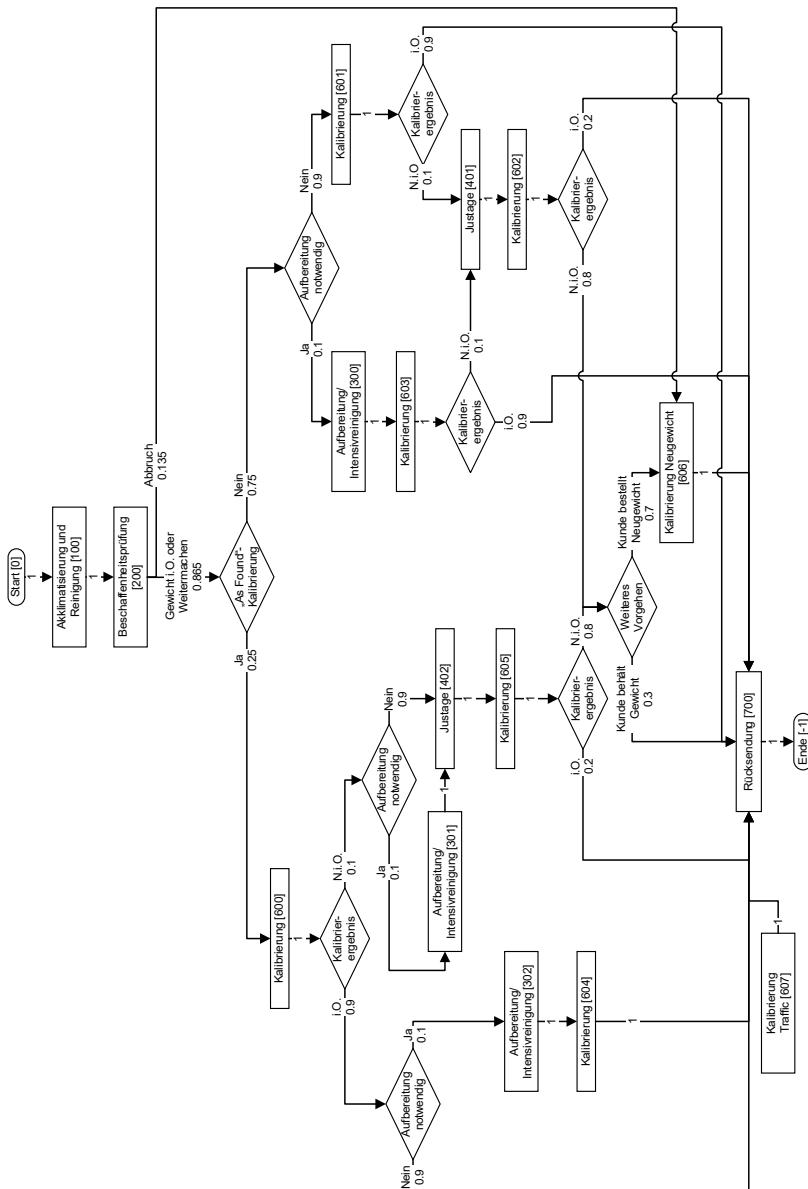
Nominal -wert	Class E1	Class E2	Class F1	Class F2	Class M1	Class M1-2	Class M2	Class M2-3	Class M3
5000 kg			25.000	80.000	250.000	500.000	800.000	1.600.000	2.500.000
2000 kg			10.000	30.000	100.000	200.000	300.000	600.000	1.000.000
1000 kg		1.600	5.000	16.000	50.000	100.000	160.000	300.000	500.000
500 kg		800	2.500	8.000	25.000	50.000	80.000	160.000	250.000
200 kg		300	1.000	3.000	10.000	20.000	30.000	60.000	100.000
100 kg		160	500	1.600	5.000	10.000	16.000	30.000	50.000
50 kg	25	80	250	800	2.500	5.000	8.000	16.000	25.000
20 kg	10	30	100	300	1.000		3.000		10.000
10 kg	5,0	16	50	160	500		1.600		5.000
5 kg	2,5	8,0	25	80	250		800		2.500
2 kg	1,0	3,0	10	30	100		300		1.000
1 kg	0,50	1,6	5,0	16	50		160		500
500 g	0,25	0,80	2,5	8,0	25		80		250
200 g	0,10	0,30	1,0	3,0	10		30		100
100 g	0,05	0,16	0,50	1,6	5,0		16		50
50 g	0,03	0,10	0,30	1,0	3,0		10		30
20 g	0,025	0,08	0,25	0,80	2,5		8,0		25
10 g	0,020	0,06	0,20	0,60	2,0		6,0		20
5 g	0,016	0,05	0,16	0,50	1,6		5,0		16
2 g	0,012	0,04	0,12	0,40	1,2		4,0		12
1 g	0,010	0,03	0,10	0,30	1,0		3,0		10
500 mg	0,008	0,025	0,08	0,025	0,80		2,5		
200 mg	0,006	0,020	0,06	0,200	0,60		2,0		
100 mg	0,005	0,016	0,05	0,160	0,50		1,6		
50 mg	0,004	0,012	0,04	0,120	0,40				
20 mg	0,003	0,010	0,03	0,100	0,30				
10 mg	0,003	0,008	0,025	0,080	0,25				
5 mg	0,003	0,006	0,020	0,060	0,20				
2 mg	0,003	0,006	0,020	0,060	0,20				
1 mg	0,003	0,006	0,020	0,060	0,20				

Alle Werte ohne Einheit sind in mg.

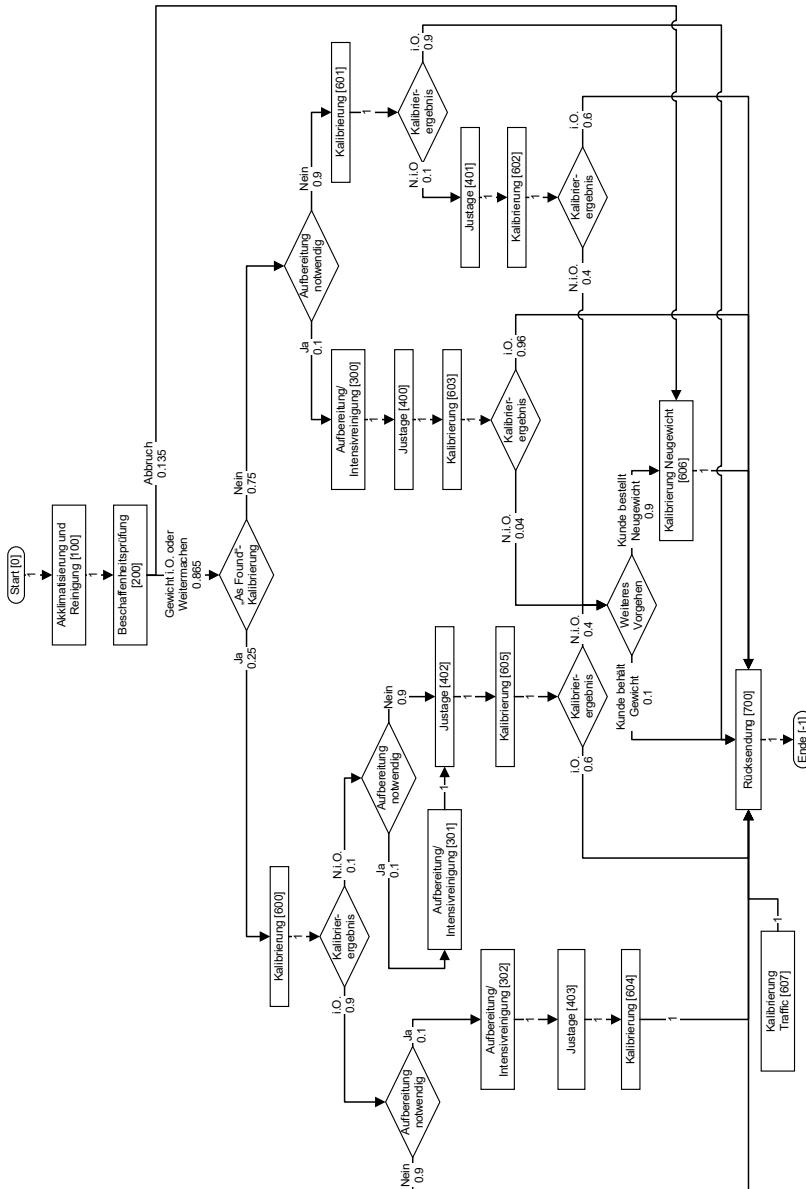
## A12 Prozessdiagramm für E1-Gewichte



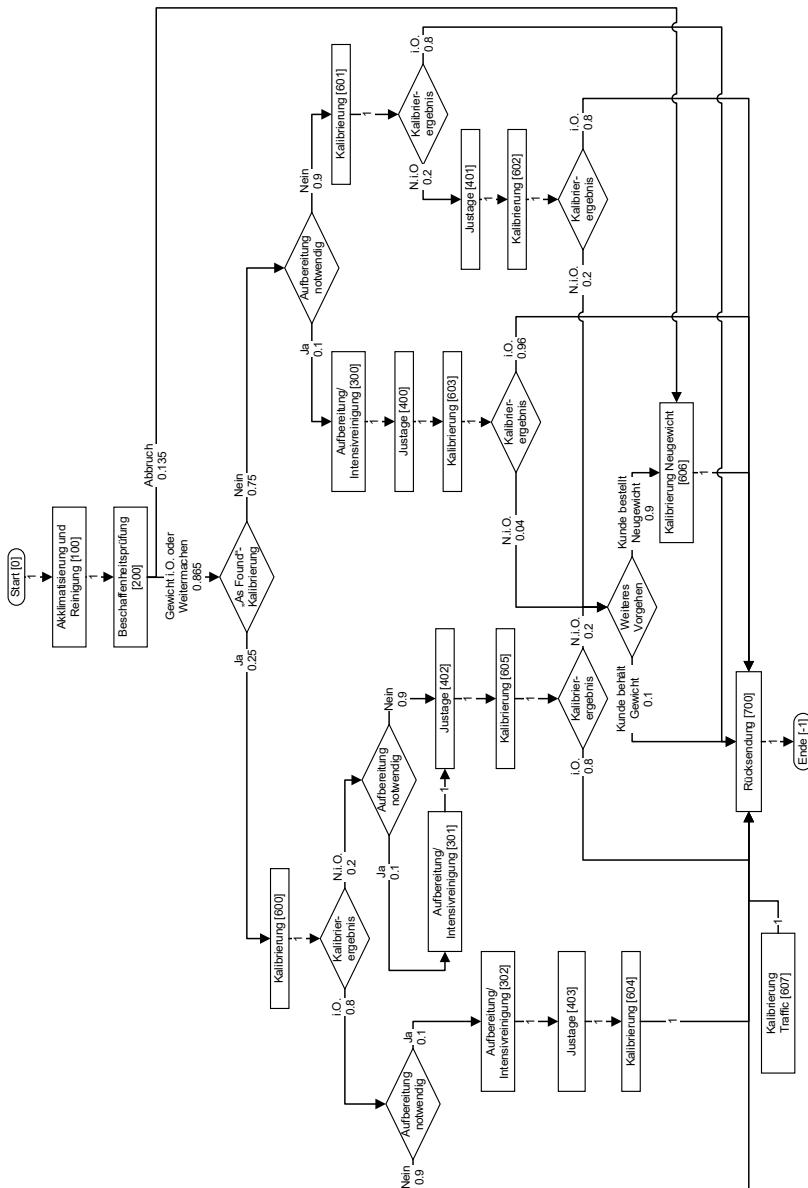
## A13 Prozessdiagramm für E2-Gewichte



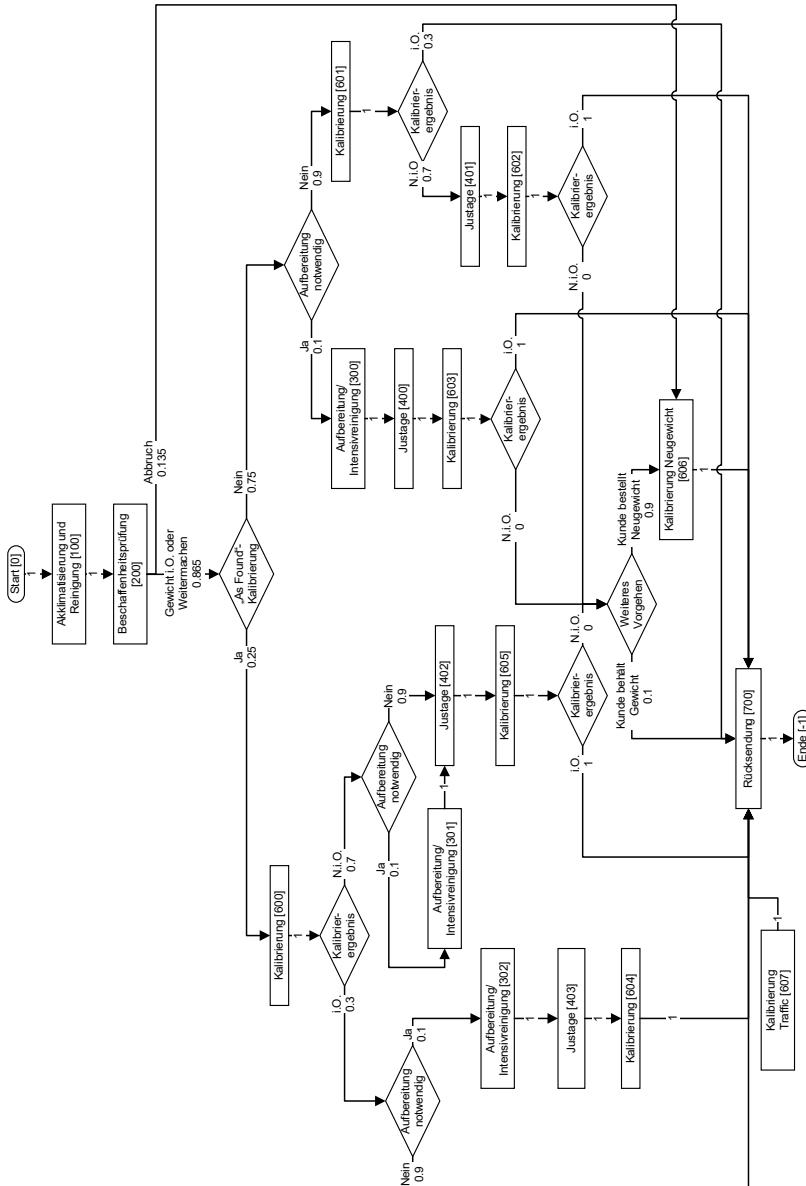
## A14 Prozessdiagramm für F1-Gewichte



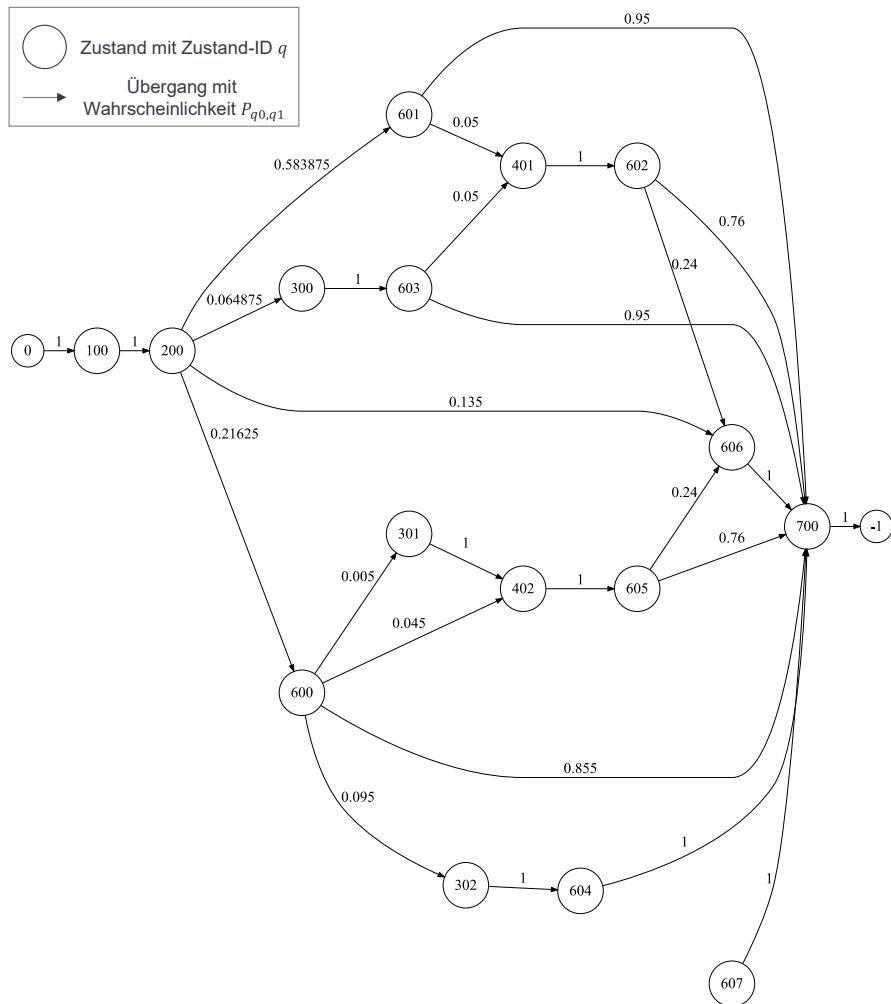
## A15 Prozessdiagramm für F2-Gewichte



## A16 Prozessdiagramm für M-Gewichte



### A17 Beispiel für einen Fertigungsplan für Anwendungsfall 1



## A18 Fertigungszeiten am Beispiel E2 für Anwendungsfall 1

In den nachfolgenden Tabellen zu den Fertigungszeiten  $f_{Z,j,q,Styp}$  werden exemplarisch die Stationen aufgeführt, die relevant für Gewichte der Klasse E2 sind. Zusätzlich zur Stationsidentifikationsnummer in Zeile 1 sind auch die Stationsbezeichnungen in Zeile 2 angegeben. Alle Prozesszeiten sind in Minuten angegeben.

Tabelle A18-1: Maschinenspezifische Prozesszeiten für Toleranzklasse E2 (Teil 1)

Tabelle A18-2: Maschinenspezifische Prozesszeiten für Toleranzklasse E2 (Teil 2)

ID	15	16	17	21	22	23	24	25	26	27	28	29
Stations- name												
Gewicht in g	M16	M17	M18	M22	M23	M24	M28	M29	M30	M31	M34	M35
0,001	80	80	50	50	80	80	90	50	80	80	80	80
0,002	80	80	50	50	80	80	90	50	80	80	80	80
0,005	80	80	50	50	80	80	90	50	80	80	80	80
0,01	80	80	50	50	80	80	90	50	80	80	80	80
0,02	80	80	50	50	80	80	90	50	80	80	80	80
0,05	80	80	50	50	80	80	90	50	80	80	80	80
0,1	80	80	50	50	80	80	90	50	80	80	80	80
0,2	80	80	50	50	80	80	90	50	80	80	80	80
0,5	80	80	50	50	80	80	90	50	80	80	80	80
1	80	80	50	50	80	80	90	50	80	80	80	80
2	80	80	50	50	80	80	90	50	80	80	80	80
5	80	80	50	50	80	80	90	50	80	80	80	80
10	80	80	50	50	80	80	90	50	80	80	80	80
20	80	80	50	50	80	80	90	50	80	80	80	80
50	80	80	50	50	80	80	90	50	80	80	80	80
100	80	80	50	50	80	80	90	50	80	80	80	80
200	80	80	50	50	80	80	90	50	80	80	80	80
500	80	80	50	50	80	80	90	50	80	80	80	80
1000	80	80	50	50	80	80	90	50	80	80	80	80
2000	80	80	50	50	80	80	90	50	80	80	80	80
5000	80	80	50	50	80	80	90	50	80	80	80	80
10000	80	80	50	50	80	80	90	50	80	80	80	80
20000	80	80	50	50	80	80	90	50	80	80	80	80
50000	80	80	50	50	80	80	90	50	80	80	80	80

## A19 Statistische Details – statische Lieferzeit – Szenario 1 – Anwendungsfall 1

		Agent	Status Quo mit statischer Lieferzeit
Anzahl Verspätungen $n_{verspätet}$ [-]	Servicegrad $sg$ [%]	$\mu$	97,75 %
		$\sigma$	1,44 %
		$min$	93,15 %
		$max$	99,80 %
Anzahl entgangener Aufträge $n_{aa}$ [-]		$\mu$	100,50
		$\sigma$	73,31
		$min$	8,00
		$max$	343,00
Liefertermin- Abweichung $t_{ia}$ [min]		$\mu$	2.265,80
		$\sigma$	593,43
		$min$	1.554,00
		$max$	4.476,00
		$\mu$	14.840.840
		$\sigma$	4.235.665
		$min$	4.496.527
		$max$	20.533.630

## A20 Statistische Details – flexible Lieferzeit – Szenario 1 – Anwendungsfall 1

		Agent	Status Quo mit flexibler Lieferzeit
Anzahl Verspätungen $n_{verp\ddot{a}t}$ [-]	$\mu$	95,56 %	78,19 %
	$\sigma$	3,68 %	3,99 %
	$min$	81,03 %	68,73 %
	$max$	99,40 %	83,29 %
Anzahl entgangener Aufträge $n_{aa}$ [-]	$\mu$	185,77	1.442,77
	$\sigma$	188,34	265,76
	$min$	19,00	1.106,00
	$max$	1.008,00	2.076,00
Liefertermin-Abweichung $t_{ta}$ [min]	$\mu$	2.637,10	0
	$\sigma$	731,95	0
	$min$	1.303,00	0
	$max$	3.861,00	0
	$\mu$	14.653.130	21.976.500
	$\sigma$	4.796.835	2.740.804
	$min$	5.879.796	19.080.060
	$max$	23.069.400	28.808.310

## A21 Statistische Details – statische Lieferzeit – Szenario 2 – Anwendungsfall 1

		Agent	Status Quo mit statischer Lieferzeit
Anzahl Verspätungen $n_{verspätet}$ [-]	Servicegrad $sg$ [%]	$\mu$	98,63 %
		$\sigma$	1,02 %
		$min$	96,16 %
		$max$	100 %
Anzahl entgangener Aufträge $n_{aa}$ [-]		$\mu$	48,07
		$\sigma$	45,01
		$min$	0,00
		$max$	150,00
Liefertermin- Abweichung $t_{ia}$ [min]		$\mu$	3.665,80
		$\sigma$	1.307,22
		$min$	2.201,00
		$max$	6.230,00
		$\mu$	9.576.250
		$\sigma$	4.781.925
		$min$	1.011.852
		$max$	17.542.510

## A22 Statistische Details – flexible Lieferzeit – Szenario 2 – Anwendungsfall 1

		Agent	Status Quo mit flexibler Lieferzeit
Anzahl Verspätungen $n_{verp\ddot{a}t}$ [-]	$\mu$	95,50 %	77,53 %
	$\sigma$	2,67 %	3,74 %
	$min$	86,68 %	71,70 %
	$max$	99,68 %	85,00 %
Anzahl entgangener Aufträge $n_{aa}$ [-]	$\mu$	132,07	1.484,07
	$\sigma$	74,99	247,31
	$min$	7,00	990,00
	$max$	267,00	1.877,00
Liefertermin-Abweichung $t_{ta}$ [min]	$\mu$	3.391,57	0
	$\sigma$	1.236,52	0
	$min$	2.030,00	0
	$max$	6.181,00	0
	$\mu$	10.450.180	17.764.850
	$\sigma$	3.694.933	1.452.177
	$min$	1.309.749	16.080.540
	$max$	16.300.070	23.458.790

## A23 Statistische Details – statische Lieferzeit – Szenario 3 – Anwendungsfall 1

		Agent	Status Quo mit statischer Lieferzeit
Anzahl Verspätungen $n_{verspätet}$ [-]	Servicegrad $sg$ [%]	$\mu$	98,74 %
		$\sigma$	1,31 %
		$min$	94,79 %
		$max$	99,93 %
Anzahl entgangener Aufträge $n_{aa}$ [-]		$\mu$	43,20
		$\sigma$	50,22
		$min$	1,00
		$max$	199,00
Liefertermin- Abweichung $t_{ia}$ [min]		$\mu$	3550,93
		$\sigma$	1173,64
		$min$	2257,00
		$max$	6284,00
		$\mu$	10.694.050
		$\sigma$	5.213.114
		$min$	981.812
		$max$	21.589.300

## A24 Statistische Details – flexible Lieferzeit – Szenario 3 – Anwendungsfall 1

		Agent	Status Quo mit flexibler Lieferzeit
Anzahl Verspätungen $n_{verp\ddot{a}t}$ [-]	$\mu$	95,28 %	78,95 %
	$\sigma$	5,83 %	3,14 %
	$min$	72,24 %	70,73 %
	$max$	99,60 %	86,14 %
Anzahl entgangener Aufträge $n_{aa}$ [-]	$\mu$	111,40	1.390,33
	$\sigma$	93,71	208,43
	$min$	10,00	916,00
	$max$	372,00	1.935,00
Liefertermin-Abweichung $t_{ta}$ [min]	$\mu$	3.638,27	0
	$\sigma$	1.137,08	0
	$min$	1.855,00	0
	$max$	6.273,00	0
	$\mu$	11.076.110	21.348.090
	$\sigma$	4.749.982	1.904.149
	$min$	605.058	19.071.210
	$max$	22.159.480	28.109.700

## A25 Beispiel für einen Pneumatikzylinder<sup>8</sup>



## A26 Herleitung der verkaufsrelevanten Daten im zweiten Anwendungsfall

Es wird nachfolgend angenommen, dass der betrachtete Pneumatikzylinder zu den A-Teilen innerhalb einer ABC-Analyse gehört. A-Teile machen ca. 80 % des Umsatzes aus (Thommen & Achleitner et al. 2020, S. 172). Es ergibt sich, dass die A-Teile des betrachteten Unternehmens 1,91 Mrd. € ausmachen.

$$1,91 \text{ Mrd. €} = 3,07 \text{ Mrd. €} \cdot 0,97 \cdot 0,8 \cdot 0,8 \quad \text{Formel A26-1}$$

Die A-Teile machen wiederum 10 % der täglich produzierten Produkte aus, wodurch sich 220 A-Teile ergeben.

$$220 = 2200 \cdot 0,1 \quad \text{Formel A26-2}$$

Es wird angenommen, dass sich der A-Teile-Umsatz auf alle A-Teile gleichmäßig aufteilt. Hierdurch ergibt sich ein Umsatz des betrachteten Pneumatikzylinder von ca. 8,7 Mio. €.

$$8,7 \text{ Mio. €} = 1,91 \text{ Mrd. €} / 220 \quad \text{Formel A26-3}$$

Weiter wird angenommen, dass sich der Umsatz nicht auf alle Varianten des Pneumatikzylinders gleich aufteilt. Stattdessen wird der Umsatz gemäß einer Dreiecksverteilung auf die verschiedenen Durchmesservarianten aufgeteilt. Bei Varianten mit demselben Durchmesser wird angenommen, dass der Umsatz gleichverteilt ist. Hieraus ergibt sich die in Abbildung A26-1 dargestellte Umsatzanteilverteilung.

---

<sup>8</sup> [https://www.festo.com/media/pim/271/D15000100124271\\_1056x1024.jpg](https://www.festo.com/media/pim/271/D15000100124271_1056x1024.jpg) [24.04.2022]

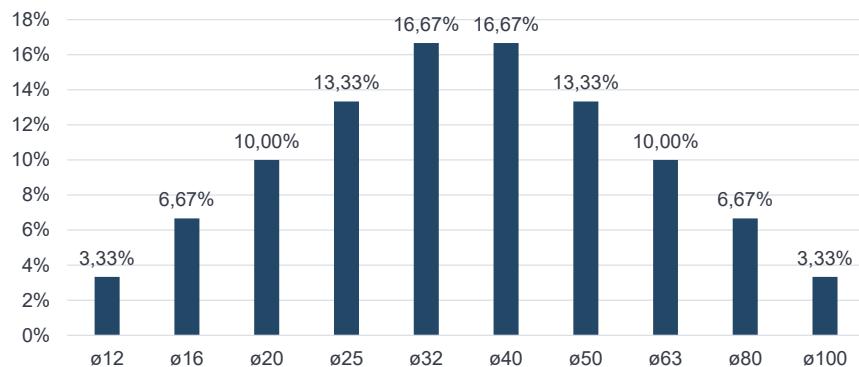


Abbildung A26-1: Dreiecksverteilung der Stückzahl verkaufter ADN-Kompaktzylinder nach Kolbendurchmesser in mm

Die Verkaufspreise der einzelnen Varianten werden aus einem Webshop entnommen<sup>9</sup>. Um das Verkaufsvolumen des Pneumatikzylinders zu berechnen, wird der Durchschnittspreis bestimmt und der Umsatz durch den Durchschnittspreis geteilt. Hierdurch ergibt sich eine Stückzahl von ca. 113.500.

$$113.500 \text{ Stk} \approx 8,7 \text{ Mio. €} / 76,68 \frac{\text{€}}{\text{Stk}} \quad \text{Formel A26-4}$$

Die Stückzahl kann nun mithilfe einer Dreiecksverteilung auf die einzelnen Durchmesser-Varianten umgelegt werden. Es wird weiter angenommen, dass die Bestellungen einer Poisson-Verteilung folgen, analog zu Kapitel 5.2.1.2.

Aus den bisher berechneten Werten und Annahmen kann nun die kolbendurchmesser-spezifische Datentabelle kalkuliert werden. Diese findet sich in Tabelle 6-24.

## A27 Bestimmung der Bestellmenge einer Bestellung in Anwendungsfall 2

Auf Basis der Bestellmenge pro Monat insgesamt und der erwarteten Bestellhäufigkeit von  $\frac{1}{\lambda_{exp}} = 1,5$  Monaten (vgl. 6.2.1.2) kann die erwartete Bestellmenge pro Bestellung eines Kunden hergeleitet werden. Es wird angenommen, dass die Bestellmenge für

<sup>9</sup> Vgl. <https://www.landefeld.de/gruppe/de/kompaktzylinder-adn/FTO-536374> [24.04.2022]

jeden Kunden poissonverteilt ist (vgl. Anhang A26). Hierdurch ergibt sich eine Möglichkeit zur Umrechnung der Gesamtnachfrage auf die einzelne Nachfrage(-verteilung) wie in Formel A27-1 gezeigt. Dies gilt nur, sofern ein Kunde  $i$  Variante  $v$  nachfragt. Diesem Umstand wird mit dem besonderen Iterator  $i^*$  Rechnung getragen.

$$\lambda_{pois,i} = \frac{\lambda_{pois,v}}{\sum_{i^* \in I}(1)} * \lambda_{exp} \quad \text{Formel A27-1}$$

## A28 Fertigungszeiten für den Pneumatikzylinder

In den nachfolgenden Ausführungen wird als Erstes dargelegt welche Referenzmaschine herangezogen wurde. Danach werden für die Bestimmung der Fertigungszeit relevante Standortgegebenheiten für die Standorte Deutschland, China und USA dargelegt. Auf Grundlage der Überlegungen werden dann zuerst die Fertigungszeiten pro Variante und dann die Gesamtfertigungszeit bestimmt.

### Bestimmung einer Referenzmaschine

Zur weiteren Herleitung des Produktionssystems wird nun als Erstes eine Referenzmaschine herangezogen. Hierfür wird ein hoher Automatisierungsgrad angenommen. Deshalb wird die Universal-Drehmaschine des Typs CTX der Firma DMG Mori als Referenz gewählt. Diese Drehmaschine ist geeignet, Komponenten des Pneumatikzylinders zu fertigen<sup>10</sup>. Zudem sind Maschinen der Firma DMG Mori als größter europäischer Werkzeugmaschinenhersteller in einer Vielzahl von produzierenden Betrieben zu finden und daher repräsentativ<sup>11</sup>. Zur Orientierung bei der Ermittlung der Fertigungszeit dient ein Kennwert von 3,8 min, den DMG Mori für die Zerspanung eines Hydraulikteils mit 44mm Durchmesser, 1 mm Aufmaß und einer Länge von 100 mm angibt<sup>10</sup>.

### Berücksichtigung von Standortgegebenheiten

Bevor nun die Fertigungszeit pro Variante ermittelt werden kann, gilt zu berücksichtigen, dass der Pneumatikzylinder an drei Standorten (Deutschland, USA, China) gefertigt wird. Für diese drei Standorte wird angenommen, dass jeder bis zu zwei Maschinen

---

<sup>10</sup> Vgl. [https://media.dmgmori.com/media/epaper/ctx\\_alpha\\_beta\\_de/index.html](https://media.dmgmori.com/media/epaper/ctx_alpha_beta_de/index.html) [24.04.2022]

<sup>11</sup> Vgl. <https://www.produktion.de/wirtschaft/das-sind-die-groessten-werkzeugmaschinen-hersteller-europas-387.html> [24.04.2022].

dieselben Typs mit unterschiedlichem Automatisierungsgrad zu Fertigung des Pneumatikzylinders besitzt. Referenzstandort ist hierbei Deutschland mit dem höchsten Automatisierungsgrad<sup>12</sup>. Danach folgen die USA und China<sup>12</sup>. Es wird angenommen, dass diese Reihenfolge analog auch für die Produktionsstandorte gilt. Entsprechend werden pro Standort ein Maschinentyp mit höherer Automatisierung und ein Maschinentyp mit niedrigerer Automatisierung angenommen. Weiter wird angenommen, dass eine niedrigere Automatisierung zu 5 % längeren Fertigungszeiten führt. In Kombination mit den Standorten ergeben sich somit sechs Maschinentypen, welche jeweils in 5 %-Inkrementen längere Fertigungszeiten besitzen. Auf Basis dieser Überlegung können die in Tabelle A28-1: zu findenden Bearbeitungszeitfaktoren aufgestellt werden.

*Tabelle A28-1: Überblick über die Bearbeitungszeitfaktoren*

Land	Stationstyp $S_{typ}$	Bearbeitungszeit-Faktor $b_f$
Deutschland	Höher automatisiert	1,00
Deutschland	Weniger automatisiert	1,05
USA	Höher automatisiert	1,10
USA	Weniger automatisiert	1,15
China	Höher automatisiert	1,20
China	Weniger automatisiert	1,25

### **Bestimmung der Fertigungszeit pro Varianteneinheit**

Da es sich bei der Fertigung des Pneumatikzylinders um einen zerspanenden Prozess handelt, sind bei der Schätzung der Fertigungszeit die Ausführungen von Klocke & König (2008, S. 375) relevant. Die Fertigungszeit einer Einheit setzt sich aus der Erholungszeit, Verteilzeit und Grundzeit zusammen (Klocke & König 2008, S. 375). Für die Fertigungszeit wird eine Erholungszeit und Verteilzeit von in Summe 0,1 min angenommen. Die Grundzeit wird zuerst in die Hauptzeit und die Nebenzeit unterschieden (Klocke & König 2008, S. 375). Die Hauptzeit beschreibt hierbei die Zeit, in der ein unmittelbarer Fertigungsfortschritt erzielt wird. Die Nebenzeit umfasst alle Zeiten, die mittelbar bei der Zerspanung anfallen, wie zum Beispiel der Werkzeugwechsel (Klocke & König 2008, S. 375). Es wird angenommen, dass für alle Varianten die Nebenzeiten konstant sind und nur die Hauptzeit aufgrund des sich ändernden Spanvolumens variiert. Es wird angenommen, dass von den 3,8 min pro Werkstück 1,5 min Nebenzeit

<sup>12</sup> Vgl. [https://ifr.org/downloads/press2018/2021-01-27\\_IFR\\_Pressemeldung\\_Roboterdichte.pdf](https://ifr.org/downloads/press2018/2021-01-27_IFR_Pressemeldung_Roboterdichte.pdf) [24.04.2022]

sind. Die Hauptzeit hingegen nimmt mit dem Zerspanvolumen linear zu. Für den Pneumatikzylinder kann ein mittleres Aufmaß von 2,5 mm gerechnet werden. Es ergibt sich folgende Formel:

$$V_j = 2\pi \cdot \left( \left( \frac{d_j + d_{aufmaß}}{2} \right)^2 - \left( \frac{d_j}{2} \right)^2 \right) \cdot l_j \quad \text{Formel A28-1}$$

Die Fertigungszeit lässt sich für den Stationstyp 1 bestimmen mit:

$$fz_{j,1,S_{typ}} = \left( \frac{t_{ref,haupt} \cdot V_j + t_{ref,neben}}{V_{ref}} \right) \cdot b_f + t_{verteil} \quad \text{Formel A28-2}$$

Das Zerspanvolumen des Referenzwerkstücks ergibt sich gemäß Formel A28-3 zu:

$$V_{ref} = V_{44mm} = 2\pi \cdot 22,25mm^2 \cdot 100mm \quad \text{Formel A28-3}$$

Mit den bekannten Daten von Referenzzyliner und Pneumatikzyliner, sowie dem Verhältnis aus Zerspanvolumen des Referenzbauteils und des Pneumatikzyliners kann nun die Formel für die Fertigungszeit eines Pneumatikzyliners bestimmt werden:

$$fz_{j,1,S_{typ}} = \left( \frac{2,3 \text{ min.}}{V_{ref}} \cdot V_j + 1,5 \text{ min.} \right) \cdot b_f + 0,1 \text{ min} \quad \text{Formel A28-4}$$

$$fz_{j,1,S_{typ}} = \left( \frac{2,3 \text{ min.}}{2225mm^3} \cdot \left( \left( \frac{d_j + 2,5 \text{ mm}}{2} \right)^2 - \left( \frac{d_j}{2} \right)^2 \right) \cdot l_j + 1,5 \text{ min.} \right) \cdot b_f + 0,1 \text{ min} \quad \text{Formel A28-5}$$

Zusammen mit den verschiedenen Kolbendurchmessern und der durchschnittlichen Länge ergeben sich die Fertigungszeiten gemäß Tabelle A28-2.

*Tabelle A28-2: Übersicht über die Fertigungszeiten im zweiten Anwendungsfall*

Kolbendurchmesser $d_j$ in mm	Ø12	Ø16	Ø20	Ø25	Ø32	Ø40	Ø50	Ø63	Ø80	Ø100	
Durchschnittliche Kolbenlänge $l_j$ in mm	20,7	24,4	28,3	28,3	33,5	33,5	33,5	36,7	36,7	36,7	
$b_f$	1,0	1,95	2,14	2,38	2,56	3,04	3,39	3,82	4,64	5,45	6,40
	1,05	2,05	2,25	2,49	2,68	3,19	3,55	4,00	4,87	5,72	6,71
	1,1	2,14	2,35	2,61	2,81	3,33	3,71	4,19	5,10	5,98	7,03
	1,15	2,23	2,45	2,72	2,93	3,48	3,88	4,38	5,33	6,25	7,34
	1,2	2,33	2,55	2,83	3,05	3,63	4,04	4,56	5,55	6,52	7,66
	1,25	2,42	2,65	2,95	3,18	3,77	4,21	4,75	5,78	6,79	7,97
$\emptyset fz_{j,1,S_{typ}}$	<b>2,19</b>	<b>2,40</b>	<b>2,66</b>	<b>2,87</b>	<b>3,41</b>	<b>3,80</b>	<b>4,28</b>	<b>5,21</b>	<b>6,12</b>	<b>7,18</b>	

### Bestimmung der Gesamtfertigungszeit

Mithilfe der Fertigungszeit pro Variante, den Bearbeitungszeitfaktoren und den Verkaufsvolumina lässt sich die theoretische Gesamtfertigungszeit berechnen. Es wird angenommen, dass alle Standorte ungefähr gleich viele Maschinen haben. Für die Berechnung der Gesamtfertigungszeit werden deshalb die Bearbeitungszeitfaktoren gemittelt, sodass sich ergibt:

$$\bar{b}_f = \frac{\sum b_f}{n_{bf}} = \frac{1 + 1,05 + 1,1 + 1,15 + 1,2 + 1,25}{6} = 1,125 \quad \text{Formel A28-6}$$

Nun lässt sich durch Multiplikation der Einzelfertigungszeit mit den Verkaufsvolumen einer Variante und die anschließende Summenbildung der Ergebnisse die Gesamtfertigungszeit ermitteln. Die sich so ergebenden Werte finden sich in Tabelle 6-26.

### A29 Herleitung der Maschinenanzahl

Für die Bestimmung der Maschinenanzahl wurde angenommen, dass allgemein im Zwei-Schicht-Betrieb produziert wird. Es ergibt sich hieraus eine theoretische Produktionskapazität wie folgt:

$$350.400 \frac{\text{min}}{\text{Jahr} \& \text{Station}} = \\ 2 \frac{\text{Schichten}}{\text{Tag}} \cdot 8 \frac{\text{h}}{\text{Schicht} \& \text{Station}} \cdot 60 \frac{\text{min}}{\text{h}} \cdot 365 \frac{\text{Tage}}{\text{Jahr}} \quad \text{Formel A29-1}$$

Es kann nicht davon ausgegangen werden, dass die Maschinen einzig den Pneumatikzylinder fertigen. Maschinen werden besonders in der Variantenfertigung für mehrere Produkte eingesetzt, sodass eine saubere Isolation nicht möglich ist. Stattdessen wird angenommen, dass auf den Produktionsressourcen zusätzlich vier Produkte mit der gleichen betrachteten Komponente bedient werden. Die Gesamtfertigungszeit beträgt 433.607 min (vgl. Tabelle 6-26). Hieraus folgt, dass sich die zu berücksichtigende Gesamtfertigungszeit auf 2.168.035 min erhöht.

$$5 \cdot 433.608 \text{ min} = 2.168.040 \text{ min.} \quad \text{Formel A29-2}$$

Eine letzte Annahme betrifft die Zielauslastung der Maschinen. Um eine gewisse Flexibilität gegen Schwankungen vorzuhalten, wird angenommen, dass die Maschinen im Schnitt zu 90 % ausgelastet sein sollen. Es ergibt sich nun die Anzahl benötigter Maschinen zu:

$$\frac{2.168.040 \text{ min.}}{350.400 \text{ min.} \cdot 90\%} \approx 7 \quad \text{Formel A29-3}$$

## A30 Maschineninvestitionskosten

Eine Maschine in geeigneter, vollautomatischer Ausstattung kostet ca. 230.000 € (DMG MORI Academy GmbH 2021). Es wird jedoch angenommen, dass Maschinen mit niedrigerem Automatisierungsgrad weniger kosten. So werden die Investitionskosten gemäß folgender Formel abhängig vom Bearbeitungszeitfaktor berechnet:

$$k_{f,ctx} = k_{ctx} \cdot (2 - b_f) \quad \text{Formel A30-1}$$

## A31 Ausführungen zur Potenzialanalyse in Anwendungsfall 2

Um den zentralen Grenzwertsatz zu erfüllen, werden 30 Zufallsversuche durchgeführt und die Gewinndifferenzen als Zufallsvariable gebildet. Es gilt dann wiederum zu prüfen, ob die Nullhypothese  $H_0$  verworfen werden kann, welche besagt, dass der Erwartungswert der Gewinndifferenz nicht größer als null ist (vgl. Formel A31-1).

$$H_0: \mu_{\Delta G} \leq 0 \text{ und } H_1: \mu_{\Delta G} > 0 \quad \text{Formel A31-1}$$

In den nachfolgenden Ausführungen werden aus Gründen der Übersichtlichkeit nur die Daten für den Hypothesentest direkt gezeigt. Für alle weiteren statistischen Details zu den betrachteten Kennzahlen wird in den nachfolgenden Abschnitten jeweils auf weitere Anhänge verwiesen.

Die Potenzialermittlung basiert auf der vom Verfasser der Arbeit betreuten Abschlussarbeit A\_Ochs (2021).

### Szenario 1: Statische, heterogene Kundenpräferenzen

Aus dem ersten Szenario und einem Vergleich von Agent und Status Quo mit statischer Lieferzeit resultieren die folgenden Werte:

Tabelle A31-1: Ergebnisse – statische Lieferzeit – Szenario 1 – Anwendungsfall 2

	$\bar{G}$ [€]	$\sigma_g$ [€]	min [€]	max [€]
Agent	8.009.801	460.586	7.206.097	8.917.600
Statische Lieferzeit	5.359.357	236.056	4.858.570	5.725.493

Basierend auf diesen Werten kann der Hypothesentest wie in Tabelle A31-2 durchgeführt werden.

*Tabelle A31-2: Hypothesentest geg. statische Lieferzeit – Szenario 1 – Anwendungsfall 2*

	$\bar{\Delta G}$ [€]	$\hat{\sigma}_{\Delta G}$ [€]	$n_{stichprobe}$	$T^*$	$t_{29,95\%}$	$H_0$
Differenzwerte	2.650.444	442.155	30	32,83	1,70	abgelehnt

Es zeigt sich, dass die Nullhypothese abgelehnt werden kann, weil  $T^* > t_{29,95\%}$  gilt, und somit der Agent einen signifikant höheren Gewinn als das Vergleichssystem erzielt. Das Konfidenzintervall ist durch den Bereich [2.485.340; 2.815.548] definiert. Anhang A32 führt die restlichen Kenngrößen auf. Der Agent hat hier einen höheren Servicegrad und gleichzeitig weniger entgangene Aufträge als das Vergleichssystem. Dieser Umstand ist bemerkenswert, da sich hier erstmals eine Konstellation zeigt, in der der Agent in gegenläufigen Kennzahlen dominant besser ist als das Vergleichssystem. Es konnten jedoch keine Kapazitäten eingespart werden.

Für den Simulationslauf im Vergleich zum Status Quo mit flexibler Lieferzeit stellt sich der Gewinn wie folgt dar:

*Tabelle A31-3: Ergebnisse – flexible Lieferzeit – Szenario 1 – Anwendungsfall 2*

	$\bar{G}$ [€]	$\sigma_g$ [€]	$min$ [€]	$max$ [€]
Agent	10.069.250	474.039	8.440.648	10.734.360
Flexible Lieferzeit	8.545.226	186.196	8.016.935	8.817.330

Der zugehörige Hypothesentest ist in der nachfolgenden Tabelle durchgeführt:

*Tabelle A31-4: Hypothesentest geg. flexible Lieferzeit – Szenario 1 – Anwendungsfall 2*

	$\bar{\Delta G}$ [€]	$\hat{\sigma}_{\Delta G}$ [€]	$n_{stichprobe}$	$T^*$	$t_{29,95\%}$	$H_0$
Differenzwerte	1.524.029	450.371	30	18,53	1,70	abgelehnt

Auch hier kann die Nullhypothese verworfen werden und der Agent setzt sich durch, während sich das Konfidenzintervall zu [1.355.857; 1.692.201] ergibt. Wie Anhang A33 zeigt, konnte der Agent einen höheren Servicegrad erreichen. Gleichzeitig wurden jedoch mehr Aufträge abgelehnt, was nicht überraschend ist, da das Vergleichssystem

jede Lieferzeit zum Listenpreis annimmt. Es konnten erstmals in einigen Simulationsläufen Kapazitäten eingespart werden. Diese Ersparnis ist jedoch nicht signifikant über alle Wiederholungen, sodass insgesamt keine Kapazität eingespart werden kann.

### **Szenario 2: Heterogene Kunden mit Drift zur Kurzfristigkeit**

Für das zweite Szenario mit Agent und statischer Lieferzeit ergeben sich die Gewinnwerte gemäß Tabelle A31-5.

*Tabelle A31-5: Ergebnisse – statische Lieferzeit – Szenario 2 – Anwendungsfall 2*

	$\bar{G}$ [€]	$\sigma_g$ [€]	min [€]	max [€]
Agent	7.138.630	734.846	5.448.941	8.173.410
Statische Lieferzeit	5.354.359	246.064	4.859.762	5.813.095

Es folgt der Hypothesentest in Tabelle A31-6.

*Tabelle A31-6: Hypothesentest geg. statische Lieferzeit – Szenario 2 – Anwendungsfall 2*

	$\bar{\Delta G}$ [€]	$\hat{\sigma}_{\Delta G}$ [€]	$n_{stichprobe}$	$T^*$	$t_{29,95\%}$	$H_0$
Differenzwerte	1.784.271	693.672	30	14,08	1,70	abgelehnt

Es zeigt sich, dass der Agent einen signifikant höheren Gewinn erzielen kann. Das Konfidenzintervall liegt bei [1.525.250; 2.043.293]. Anhang A34 zeigt, dass der Agent wieder einen höheren Servicegrad und weniger Verspätungen erzeugt. In diesem Fall gibt es mehr entgangene Aufträge, welche jedoch nicht im Verhältnis zur Servicegraddifferenz stehen. In diesem Fall sticht jedoch die absolute Lieferterminabweichung hervor, die deutlich höher liegt als im Vergleichssystem.

Für den Vergleich mit dem Status Quo mit flexibler Lieferzeit können im zweiten Szenario folgende Gewinnwerte erzielt werden:

*Tabelle A31-7: Ergebnisse – flexible Lieferzeit – Szenario 2 – Anwendungsfall 2*

	$\bar{G}$ [€]	$\sigma_g$ [€]	min [€]	max [€]
Agent	8.904.947	828.374	6.242.138	9.999.909
Flexible Lieferzeit	8.527.393	148.080	8.233.124	8.779.301

Mit diesen Werten ergibt sich der Hypothesentest gemäß Tabelle A31-8.

*Tabelle A31-8: Hypothesentest geg. flexible Lieferzeit – Szenario 2 – Anwendungsfall 2*

	$\bar{\Delta G}$ [€]	$\hat{\sigma}_{\Delta G}$ [€]	$n_{stichprobe}$	$T^*$	$t_{29,95\%}$	$H_0$
Differenz-werte	377.554	790.799	30	2,61	1,70	abgelehnt

Da die Nullhypothese auch hier abgelehnt wird, kann geschlussfolgert werden, dass der Agent einen höheren Gewinn erzielt. Das Konfidenzintervall liegt im Bereich [82.264; 672.844]. Für die restlichen Kennzahlen zeigt Anhang A35, dass der Agent abermals einen besseren Servicegrad erreicht. Die Werte sind mit denen aus Szenario 1 vergleichbar. Jedoch zeigt sich eine deutlich größere Diskrepanz in der Lieferterminabweichung.

### **Szenario 3: Heterogene Kunden mit Drift zu geringerer Preisbereitschaft**

Für das dritte und letzte Szenario ergeben sich die Gewinnwerte von Agent und Status Quo mit statischer Lieferzeit wie folgt:

*Tabelle A31-9: Ergebnisse – statische Lieferzeit – Szenario 3 – Anwendungsfall 2*

	$\bar{G}$ [€]	$\sigma_g$ [€]	$min$ [€]	$max$ [€]
Agent	7.229.551	589.362	5.245.905	8.024.972
Statische Lieferzeit	4.629.224	237.467	4.195.372	5.286.056

Anschließend zeigt sich der Hypothesentest gemäß Tabelle A31-10.

*Tabelle A31-10: Hypothesentest geg. statische Lieferzeit – Szenario 3 – Anwendungsfall 2*

	$\bar{\Delta G}$ [€]	$\hat{\sigma}_{\Delta G}$ [€]	$n_{stichprobe}$	$T^*$	$t_{29,95\%}$	$H_0$
Differenz-werte	2.600.327	650.753	30	21,88	1,70	abgelehnt

Auch im dritten und letzten Szenario setzt sich der Agent im Gewinn durch. Das Konfidenzintervall liegt im Bereich [2.357.332; 2.843.323]. Gemäß Anhang A36 liegt der Agent sehr nah am Vergleichssystem, jedoch ist der Agent in allen Mittelwerten, bis auf die Lieferterminabweichung, leicht besser. Es konnten keine Kapazitäten eingespart werden.

Für den letzten Vergleich mit flexibler Lieferzeit ergeben sich die Gewinnwerte gemäß Tabelle A31-11.

Tabelle A31-11: Ergebnisse – flexible Lieferzeit – Szenario 3 – Anwendungsfall 2

	$\bar{G}$ [€]	$\sigma_g$ [€]	min [€]	max [€]
Agent	9.515.279	869.499	6.265.226	10.678.590
Flexible Lieferzeit	8.258.437	1.808.213	1.293.336	8.877.499

Auf den Gewinnwerten aufbauend lässt sich der Hypothesentest wie folgt durchführen:

Tabelle A31-12: Hypothesentest geg. flexible Lieferzeit – Szenario 3 – Anwendungsfall 2

	$\bar{\Delta G}$ [€]	$\hat{\sigma}_{\Delta G}$ [€]	$n_{stichprobe}$	$T^*$	$t_{29,95\%}$	$H_0$
Differenzwerte	1.256.842	1.722.144	30	3,99	1,70	abgelehnt

Damit zeigt sich, dass der Agent in allen Szenarien besser als die Vergleichssysteme ist. Hier liegt das Konfidenzintervall bei [613.783; 1.899.901]. Die Situation in den restlichen Kennzahlen stellt sich in Anhang A37 als vergleichbar mit Szenario 2 dar. Der Agent erzielt einen höheren Servicegrad, weniger Verspätungen, jedoch eine höhere Lieferterminabweichung.

## A32 Statistische Details – statische Lieferzeit – Szenario 1 – Anwendungsfall 2

		Agent	Status Quo mit statischer Lieferzeit
Anzahl Verspätungen $n_{verspätet}$ [-]	$\mu$	98,78 %	94,15 %
	$\sigma$	1,30 %	9,52 %
	$min$	94,07 %	63,20 %
	$max$	100 %	100 %
Anzahl entgangener Aufträge $n_{aa}$ [-]	$\mu$	351,77	1.593,57
	$\sigma$	374,89	2.619,45
	$min$	0	0
	$max$	1.692	10.355
Liefertermin-Abweichung $t_{ia}$ [min]	$\mu$	14.392,47	16.026,57
	$\sigma$	1.265,65	463,74
	$min$	12.013	15.136
	$max$	17.318	16.870
Eingesparte Stationen $n_{schonstation}$ [-]	$\mu$	32.043.140	9.061.766
	$\sigma$	14.770.220	2.021.833
	$min$	16.452.100	7.270.063
	$max$	93.210.790	15.925.790

## A33 Statistische Details – flexible Lieferzeit – Szenario 1 – Anwendungsfall 2

		Agent	Status Quo mit flexibler Lieferzeit
Servicegrad $sg$ [%]	$\mu$	99,58 %	98,53 %
	$\sigma$	7,00 %	2,99 %
	$min$	97,46 %	86,11 %
	$max$	100 %	100 %
Anzahl Verspätungen $n_{verspätet}$ [-]	$\mu$	108,5	403,83
	$\sigma$	181,03	839,21
	$min$	0	0
	$max$	649	3.952
Anzahl entgangener Aufträge $n_{aa}$ [-]	$\mu$	1.379,87	0
	$\sigma$	451,63	0
	$min$	825	0
	$max$	2.954	0
Liefertermin-Abweichung $t_{ta}$ [min]	$\mu$	24.114.540	11.168.420
	$\sigma$	34.007.620	8.427.309
	$min$	8.800.213	8.494.451
	$max$	163.826.900	55.505.530
Eingesparte Stationen $n_{schonstation}$ [-]	$\mu$	0,03	0
	$\sigma$	0,18	0
	$min$	0	0
	$max$	1	0

## A34 Statistische Details – statische Lieferzeit – Szenario 2 – Anwendungsfall 2

		Agent	Status Quo mit statischer Lieferzeit
Anzahl Verspätungen $n_{verspätet}$ [-]	$\mu$	99,40 %	94,33 %
	$\sigma$	1,49 %	7,53 %
	$min$	92,79 %	68,70 %
	$max$	100 %	99,99 %
Anzahl entgangener Aufträge $n_{aa}$ [-]	$\mu$	167,3	1532,1
	$\sigma$	429,33	2.051,2
	$min$	0	2
	$max$	2077	8523
Liefertermin-Abweichung $t_{ia}$ [min]	$\mu$	17.139	16.042,3
	$\sigma$	2.534,12	436,2
	$min$	11.192	14.871
	$max$	23.050	16.856
Eingesparte Stationen $n_{schonstation}$ [-]	$\mu$	68.467.970	8.408.414
	$\sigma$	41.180.890	660.757
	$min$	22.498.440	7.418.803
	$max$	182.575.700	10.404.940

## A35 Statistische Details – flexible Lieferzeit – Szenario 2 – Anwendungsfall 2

		Agent	Status Quo mit flexibler Lieferzeit
Anzahl Verspätungen $n_{verspätet}$ [-]	$\mu$	99,63 %	98,74 %
	$\sigma$	0,53 %	1,80 %
	$min$	97,91 %	92,62 %
	$max$	100 %	100 %
Anzahl entgangener Aufträge $n_{aa}$ [-]	$\mu$	85,2	347,7
	$\sigma$	120,37	501,52
	$min$	0	0
	$max$	431	2.023
Liefertermin-Abweichung $t_{ia}$ [min]	$\mu$	3.814,17	0
	$\sigma$	1.654,37	0
	$min$	1.295	0
	$max$	9.130	0
Eingesparte Stationen $n_{schonstation}$ [-]	$\mu$	78.150.650	12.222.360
	$\sigma$	37.901.610	9.949.829
	$min$	10.086.980	8.080.062
	$max$	162.334.000	61.969.100

## A36 Statistische Details – statische Lieferzeit – Szenario 3 – Anwendungsfall 2

		Agent	Status Quo mit statischer Lieferzeit
Anzahl Verspätungen $n_{verspätet}$ [-]	$\mu$	98,94 %	98,66 %
	$\sigma$	2,48 %	2,37 %
	$min$	89,16 %	88,92 %
	$max$	100 %	100 %
Anzahl entgangener Aufträge $n_{aa}$ [-]	$\mu$	303,23	315,13
	$\sigma$	715	565,32
	$min$	0	0
	$max$	3.165	2.655
Liefertermin-Abweichung $t_{ia}$ [min]	$\mu$	15.670,83	19.406,37
	$\sigma$	1.828,13	519,22
	$min$	12.322	18.324
	$max$	20.599	20.502
Eingesparte Stationen $n_{schonstation}$ [-]	$\mu$	61.230.880	40.975.310
	$\sigma$	53.372.570	1.944.294
	$min$	19.896.810	38.365.060
	$max$	271.525.700	48.986.170

## A37 Statistische Details – flexible Lieferzeit – Szenario 3 – Anwendungsfall 2

		Agent	Status Quo mit flexibler Lieferzeit
Anzahl Verspätungen $n_{verspätet}$ [-]	$\mu$	99,43 %	96,70 %
	$\sigma$	1,05 %	14,82 %
	$min$	95,83 %	18,33 %
	$max$	100 %	100 %
Anzahl entgangener Aufträge $n_{aa}$ [-]	$\mu$	140,43	984,97
	$\sigma$	256,3	4.362,94
	$min$	0	0
	$max$	1.064	24.034
Liefertermin-Abweichung $t_{ia}$ [min]	$\mu$	2.397,2	0
	$\sigma$	1855,71	0
	$min$	990	0
	$max$	9.111	0
Eingesparte Stationen $n_{schonstation}$ [-]	$\mu$	58.143.230	24.957.580
	$\sigma$	56.321.490	49.286.040
	$min$	8.789.931	8.843.152
	$max$	182.705.000	216.454.200



Forschungsberichte aus dem wbk  
Institut für Produktionstechnik  
Karlsruher Institut für Technologie (KIT)

Bisher erschienene Bände:

---

Band 0

Dr.-Ing. Wu Hong-qi

**Adaptive Volumenstromregelung mit Hilfe von drehzahlgeregelten  
Elektroantrieben**

Band 1

Dr.-Ing. Heinrich Weiß

**Fräsen mit Schneidkeramik - Verhalten des System  
Werkzeugmaschine-Werkzeug-Werkstück und Prozessanalyse**

Band 2

Dr.-Ing. Hans-Jürgen Stierle

**Entwicklung und Untersuchung hydrostatischer Lager für die  
Axialkolbenmaschine**

Band 3

Dr.-Ing. Herbert Hörner

**Untersuchung des Geräuschverhaltens druckgeregelter Axialkolbenpumpen**

Band 4

Dr.-Ing. Rolf-Dieter Brückbauer

**Digitale Drehzahlregelung unter der besonderen Berücksichtigung  
von Quantisierungseffekten**

Band 5

Dr.-Ing. Gerhard Staiger

**Graphisch interaktive NC-Programmierung von Drehteilen im Werkstattbereich**

Band 6

Dr.-Ing. Karl Peters

**Ein Beitrag zur Berechnung und Kompensation von Positionierfehlern an  
Industrierobotern**

Band 7

Dr.-Ing. Paul Stauss

**Automatisierte Inbetriebnahme und Sicherung der Zuverlässigkeit und Verfügbarkeit numerisch gesteuerter Fertigungseinrichtungen**

Band 8

Dr.-Ing. Günter Möckesch

**Konzeption und Realisierung eines strategischen, integrierten Gesamtplanungs- und -bearbeitungssystems zur Optimierung der Drehteilorganisation für auftragsbezogene Drehereien**

Band 9

Dr.-Ing. Thomas Oestreicher

**Rechnergestützte Projektierung von Steuerungen**

Band 10

Dr.-Ing. Thomas Selinger

**Teilautomatisierte werkstattnahe NC-Programmerstellung im Umfeld einer integrierten Informationsverarbeitung**

Band 11

Dr.-Ing. Thomas Buchholz

**Prozessmodell Fräsen, Rechnerunterstützte Analyse, Optimierung und Überwachung**

Band 12

Dr.-Ing. Bernhard Reichling

**Lasergestützte Positions- und Bahnvermessung von Industrierobotern**

Band 13

Dr.-Ing. Hans-Jürgen Lesser

**Rechnergestützte Methoden zur Auswahl anforderungsgerechter Verbindungselemente**

Band 14

Dr.-Ing. Hans-Jürgen Lauffer

**Einsatz von Prozessmodellen zur rechnerunterstützten Auslegung von Räumwerkzeugen**

Band 15

Dr.-Ing. Michael C. Wilhelm

**Rechnergestützte Prüfplanung im Informationsverbund moderner Produktionssysteme**

Band 16

Dr.-Ing. Martin Ochs

**Entwurf eines Programmsystems zur wissensbasierten Planung und Konfigurierung**

Band 17

Dr.-Ing. Heinz-Joachim Schneider

**Erhöhung der Verfügbarkeit von hochautomatisierten Produktionseinrichtungen mit Hilfe der Fertigungsleittechnik**

Band 18

Dr.-Ing. Hans-Reiner Ludwig

**Beanspruchungsanalyse der Werkzeugschneiden beim Stirnplanfräsen**

Band 19

Dr.-Ing. Rudolf Wieser

**Methoden zur rechnergestützten Konfigurierung von Fertigungsanlagen**

Band 20

Dr.-Ing. Edgar Schmitt

**Werkstattsteuerung bei wechselnder Auftragsstruktur**

Band 21

Dr.-Ing. Wilhelm Enderle

**Verfügbarkeitssteigerung automatisierter Montagesysteme durch selbsttätige Behebung prozessbedingter Störungen**

Band 22

Dr.-Ing. Dieter Buchberger

**Rechnergestützte Strukturplanung von Produktionssystemen**

Band 23

Prof. Dr.-Ing. Jürgen Fleischer

**Rechnerunterstützte Technologieplanung für die flexibel automatisierte Fertigung von Abkantteilen**

Band 24

Dr.-Ing. Lukas Loeffler

**Adaptierbare und adaptive Benutzerschnittstellen**

Band 25

Dr.-Ing. Thomas Friedmann

**Integration von Produktentwicklung und Montageplanung durch neue rechnergestützte Verfahren**

Band 26

Dr.-Ing. Robert Zurrin

**Variables Formhonen durch rechnergestützte Hornprozesssteuerung**

Band 27

Dr.-Ing. Karl-Heinz Bergen

**Langhub-Innenrundhonen von Grauguss und Stahl mit einem elektromechanischem Vorschubsystem**

Band 28

Dr.-Ing. Andreas Liebisch

**Einflüsse des Festwalzens auf die Eigenspannungsverteilung und die Dauerfestigkeit einsatzgehärteter Zahnräder**

Band 29

Dr.-Ing. Rolf Ziegler

**Auslegung und Optimierung schneller Servopumpen**

Band 30

Dr.-Ing. Rainer Bartl

**Datenmodellgestützte Wissensverarbeitung zur Diagnose und Informationsunterstützung in technischen Systemen**

Band 31

Dr.-Ing. Ulrich Golz

**Analyse, Modellbildung und Optimierung des Betriebsverhaltens von Kugelgewindetrieben**

Band 32

Dr.-Ing. Stephan Timmermann

**Automatisierung der Feinbearbeitung in der Fertigung von Hohlformwerkzeugen**

Band 33

Dr.-Ing. Thomas Noe

**Rechnergestützter Wissenserwerb zur Erstellung von Überwachungs- und Diagnoseexpertensystemen für hydraulische Anlagen**

Band 34

Dr.-Ing. Ralf Lenschow

**Rechnerintegrierte Erstellung und Verifikation von Steuerungsprogrammen als Komponente einer durchgängigen Planungsmethodik**

Band 35

Dr.-Ing. Matthias Kallabis

**Räumen gehärteter Werkstoffe mit kristallinen Hartstoffen**

Band 36

Dr.-Ing. Heiner-Michael Honeck

**Rückführung von Fertigungsdaten zur Unterstützung einer fertigungsgerechten Konstruktion**

Band 37

Dr.-Ing. Manfred Rohr

**Automatisierte Technologieplanung am Beispiel der Komplettbearbeitung auf Dreh-/Fräszellen**

Band 38

Dr.-Ing. Martin Steuer

**Entwicklung von Softwarewerkzeugen zur wissensbasierten Inbetriebnahme von komplexen Serienmaschinen**

Band 39

Dr.-Ing. Siegfried Beichter

**Rechnergestützte technische Problemlösung bei der Angebotserstellung von flexiblen Drehzellen**

Band 40

Dr.-Ing. Thomas Steitz

**Methodik zur marktorientierten Entwicklung von Werkzeugmaschinen mit Integration von funktionsbasierter Strukturierung und Kostenschätzung**

Band 41

Dr.-Ing. Michael Richter

**Wissensbasierte Projektierung elektrohydraulischer Regelungen**

Band 42

Dr.-Ing. Roman Kuhn

**Technologieplanungssystem Fräsen. Wissensbasierte Auswahl von Werkzeugen, Schneidkörpern und Schnittbedingungen für das Fertigungsverfahren Fräsen**

Band 43

Dr.-Ing. Hubert Klein

**Rechnerunterstützte Qualitätssicherung bei der Produktion von Bauteilen mit frei geformten Oberflächen**

Band 44

Dr.-Ing. Christian Hoffmann

**Konzeption und Realisierung eines fertigungsintegrierten Koordinatenmessgerätes**

Band 45

Dr.-Ing. Volker Frey

**Planung der Leittechnik für flexible Fertigungsanlagen**

Band 46

Dr.-Ing. Achim Feller

**Kalkulation in der Angebotsphase mit dem selbsttätig abgeleiteten Erfahrungswissen der Arbeitsplanung**

Band 47

Dr.-Ing. Markus Klaiber

**Produktivitätssteigerung durch rechnerunterstütztes Einfahren von NC-Programmen**

Band 48

Dr.-Ing. Roland Minges

**Verbesserung der Genauigkeit beim fünfachsigen Fräsen von Freiformflächen**

Band 49

Dr.-Ing. Wolfgang Bernhart

**Beitrag zur Bewertung von Montagevarianten: Rechnergestützte Hilfsmittel zur kostenorientierten, parallelen Entwicklung von Produkt und Montagesystem**

Band 50

Dr.-Ing. Peter Ganghoff

**Wissensbasierte Unterstützung der Planung technischer Systeme: Konzeption eines Planungswerkzeuges und exemplarische Anwendung im Bereich der Montagesystemplanung**

Band 51

Dr.-Ing. Frank Maier

**Rechnergestützte Prozessregelung beim flexiblen Gesenkbiegen durch Rückführung von Qualitätsinformationen**

Band 52

Dr.-Ing. Frank Debus

**Ansatz eines rechnerunterstützten Planungsmanagements für die Planung in verteilten Strukturen**

Band 53

Dr.-Ing. Joachim Weinbrecht

**Ein Verfahren zur zielorientierten Reaktion auf Planabweichungen in der Werkstattregelung**

Band 54

Dr.-Ing. Gerd Herrmann

**Reduzierung des Entwicklungsaufwandes für anwendungsspezifische Zellenrechnersoftware durch Rechnerunterstützung**

Band 55

Dr.-Ing. Robert Wassmer

**Verschleissentwicklung im tribologischen System Fräsen: Beiträge zur Methodik der Prozessmodellierung auf der Basis tribologischer Untersuchungen beim Fräsen**

Band 56

Dr.-Ing. Peter Uebelhoer

**Inprocess-Geometriemessung beim Honen**

Band 57

Dr.-Ing. Hans-Joachim Schelberg

**Objektorientierte Projektierung von SPS-Software**

Band 58

Dr.-Ing. Klaus Boes

**Integration der Qualitätsentwicklung in featurebasierte CAD/CAM-Prozessketten**

Band 59

Dr.-Ing. Martin Schreiber

**Wirtschaftliche Investitionsbewertung komplexer Produktionsysteme unter Berücksichtigung von Unsicherheit**

Band 60

Dr.-Ing. Ralf Steuernagel

**Offenes adaptives Engineering-Werkzeug zur automatisierten Erstellung von entscheidungsunterstützenden Informationssystemen**

Band 62

Dr.-Ing. Uwe Schauer

**Qualitätsorientierte Feinbearbeitung mit Industrierobotern: Regelungsansatz für die Freiformflächenfertigung des Werkzeug- und Formenbaus**

Band 63

Dr.-Ing. Simone Loeper

**Kennzahlengestütztes Beratungssystem zur Verbesserung der Logistikleistung in der Werkstattfertigung**

Band 64

Dr.-Ing. Achim Raab

**Räumen mit hartstoffbeschichteten HSS-Werkzeugen**

Band 65,

Dr.-Ing. Jan Erik Burghardt

**Unterstützung der NC-Verfahrenskette durch ein bearbeitungs-elementorientiertes, lernfähiges Technologieplanungssystem**

Band 66

Dr.-Ing. Christian Tritsch

**Flexible Demontage technischer Gebrauchsgüter: Ansatz zur Planung und (teil-)automatisierten Durchführung industrieller Demontageprozesse**

Band 67

Dr.-Ing. Oliver Eitrich

**Prozessorientiertes Kostenmodell für die entwicklungsbegleitende Vorkalkulation**

Band 68

Dr.-Ing. Oliver Wilke

**Optimierte Antriebskonzepte für Räummaschinen - Potentiale zur Leistungs-steigerung**

Band 69

Dr.-Ing. Thilo Sieth

**Rechnergestützte Modellierungsmethodik zerspantechnologischer Prozesse**

Band 70

Dr.-Ing. Jan Linnenbuerger

**Entwicklung neuer Verfahren zur automatisierten Erfassung der geometri-schen Abweichungen an Linearachsen und Drehschwenkköpfen**

Band 71

Dr.-Ing. Mathias Klimmek

**Fraktionierung technischer Produkte mittels eines frei beweglichen Wasserstrahlwerkzeuges**

Band 72

Dr.-Ing. Marko Hartel

**Kennzahlenbasiertes Bewertungssystem zur Beurteilung der Demontage- und Recyclingeignung von Produkten**

Band 73

Dr.-Ing. Jörg Schaupp

**Wechselwirkung zwischen der Maschinen- und Hauptspindelantriebsdynamik und dem Zerspanprozess beim Fräsen**

Band 74

Dr.-Ing. Bernhard Neisius

**Konzeption und Realisierung eines experimentellen Telemanipulators für die Laparoskopie**

Band 75

Dr.-Ing. Wolfgang Walter

**Erfolgsversprechende Muster für betriebliche Ideenfindungsprozesse. Ein Beitrag zur Steigerung der Innovationsfähigkeit**

Band 76

Dr.-Ing. Julian Weber

**Ein Ansatz zur Bewertung von Entwicklungsergebnissen in virtuellen Szenarien**

Band 77

Dr.-Ing. Dipl. Wirtsch.-Ing. Markus Posur

**Unterstützung der Auftragsdurchsetzung in der Fertigung durch Kommunikation über mobile Rechner**

Band 78

Dr.-Ing. Frank Fleissner

**Prozessorientierte Prüfplanung auf Basis von Bearbeitungsobjekten für die Kleinserienfertigung am Beispiel der Bohr- und Fräsbearbeitung**

Band 79

Dr.-Ing. Anton Haberkern

**Leistungsfähigere Kugelgewindetriebe durch Beschichtung**

Band 80

Dr.-Ing. Dominik Matt

**Objektorientierte Prozess- und Strukturinnovation (OPUS)**

Band 81

Dr.-Ing. Jürgen Andres

**Robotersysteme für den Wohnungsbau: Beitrag zur Automatisierung des Mauerwerkabaus und der Elektroinstallation auf Baustellen**

Band 82

Dr.-Ing. Dipl.Wirtschaftsing. Simone Riedmiller

**Der Prozesskalender - Eine Methodik zur marktorientierten Entwicklung von Prozessen**

Band 83

Dr.-Ing. Dietmar Tilch

**Analyse der Geometrieparameter von Präzisionsgewinden auf der Basis einer Least-Squares-Estimation**

Band 84

Dr.-Ing. Dipl.-Kfm. Oliver Stiebold

**Konzeption eines reaktionsschnellen Planungssystems für Logistikketten auf Basis von Software-Agenten**

Band 85

Dr.-Ing. Ulrich Walter

**Einfluss von Kühlshmierstoff auf den Zerspanprozess beim Fräsen: Beitrag zum Prozessverständniss auf Basis von zerspantechnischen Untersuchungen**

Band 86

Dr.-Ing. Bernd Werner

**Konzeption von teilautonomer Gruppenarbeit unter Berücksichtigung kultureller Einflüsse**

Band 87

Dr.-Ing. Ulf Osmers

**Projektieren Speicherprogrammierbarer Steuerungen mit Virtual Reality**

Band 88

Dr.-Ing. Oliver Doerfel

**Optimierung der Zerspantechnik beim Fertigungsverfahren Wälzstossen: Analyse des Potentials zur Trockenbearbeitung**

Band 89

Dr.-Ing. Peter Baumgartner

**Stufenmethode zur Schnittstellengestaltung in der internationalen Produktion**

Band 90

Dr.-Ing. Dirk Vossmann

**Wissensmanagement in der Produktentwicklung durch Qualitätsmethodenverbund und Qualitätsmethodenintegration**

Band 91

Dr.-Ing. Martin Plass

**Beitrag zur Optimierung des Honprozesses durch den Aufbau einer Honprozessregelung**

Band 92

Dr.-Ing. Titus Konold

**Optimierung der Fünfachsfräsbearbeitung durch eine kennzahlenunterstützte CAM-Umgebung**

Band 93

Dr.-Ing. Jürgen Brath

**Unterstützung der Produktionsplanung in der Halbleiterfertigung durch risikoberücksichtigende Betriebskennlinien**

Band 94

Dr.-Ing. Dirk Geisinger

**Ein Konzept zur marktorientierten Produktentwicklung**

Band 95

Dr.-Ing. Marco Lanza

**Entwurf der Systemunterstützung des verteilten Engineering mit Axiomatic Design**

Band 96

Dr.-Ing. Volker Hüntrup

**Untersuchungen zur Mikrostrukturierbarkeit von Stählen durch das Fertigungsverfahren Fräsen**

Band 97

Dr.-Ing. Frank Reinboth

**Interne Stützung zur Genauigkeitsverbesserung in der Inertialmesstechnik: Beitrag zur Senkung der Anforderungen an Inertialsensoren**

Band 98

Dr.-Ing. Lutz Trender

**Entwicklungsintegrierte Kalkulation von Produktlebenszykluskosten auf Basis der ressourcenorientierten Prozesskostenrechnung**

Band 99

Dr.-Ing. Cornelia Kafka

**Konzeption und Umsetzung eines Leitfadens zum industriellen Einsatz von Data-Mining**

Band 100

Dr.-Ing. Gebhard Selinger

**Rechnerunterstützung der informellen Kommunikation in verteilten Unternehmensstrukturen**

Band 101

Dr.-Ing. Thomas Windmüller

**Verbesserung bestehender Geschäftsprozesse durch eine mitarbeiterorientierte Informationsversorgung**

Band 102

Dr.-Ing. Knud Lembke

**Theoretische und experimentelle Untersuchung eines bistabilen elektrohydraulischen Linearantriebs**

Band 103

Dr.-Ing. Ulrich Thies

**Methode zur Unterstützung der variantengerechten Konstruktion von industriell eingesetzten Kleingeräten**

Band 104

Dr.-Ing. Andreas Schmälzle

**Bewertungssystem für die Generalüberholung von Montageanlagen –Ein Beitrag zur wirtschaftlichen Gestaltung geschlossener Facility- Management- Systeme im Anlagenbau**

Band 105

Dr.-Ing. Thorsten Frank

**Vergleichende Untersuchungen schneller elektromechanischer Vorschubachsen mit Kugelgewindetrieb**

Band 106

Dr.-Ing. Achim Agostini

**Reihenfolgeplanung unter Berücksichtigung von Interaktionen: Beitrag zur ganzheitlichen Strukturierung und Verarbeitung von Interaktionen von Bearbeitungsobjekten**

Band 107

Dr.-Ing. Thomas Barrho

**Flexible, zeitfenstergesteuerte Auftragseinplanung in segmentierten Fertigungsstrukturen**

Band 108

Dr.-Ing. Michael Scharer

**Quality Gate-Ansatz mit integriertem Risikomanagement**

Band 109

Dr.-Ing. Ulrich Suchy

**Entwicklung und Untersuchung eines neuartigen Mischkopfes für das Wasser Abrasivstrahlschneiden**

Band 110

Dr.-Ing. Sellal Mussa

**Aktive Korrektur von Verlagerungsfehlern in Werkzeugmaschinen**

Band 111

Dr.-Ing. Andreas Hühsam

**Modellbildung und experimentelle Untersuchung des Wälzschälprozesses**

Band 112

Dr.-Ing. Axel Plutowsky

**Charakterisierung eines optischen Messsystems und den Bedingungen des Arbeitsraums einer Werkzeugmaschine**

Band 113

Dr.-Ing. Robert Landwehr

**Konsequent dezentralisierte Steuerung mit Industrial Ethernet und offenen Applikationsprotokollen**

Band 114

Dr.-Ing. Christoph Dill

**Turbulenzreaktionsprozesse**

Band 115

Dr.-Ing. Michael Baumeister

**Fabrikplanung im turbulenten Umfeld**

Band 116

Dr.-Ing. Christoph Gönnheimer

**Konzept zur Verbesserung der Elektromagnetischen Verträglichkeit (EMV) in Produktionssystemen durch intelligente Sensor/Aktor-Anbindung**

Band 117

Dr.-Ing. Lutz Demuß

**Ein Reifemodell für die Bewertung und Entwicklung von Dienstleistungsorganisationen: Das Service Management Maturity Modell (SMMM)**

Band 118

Dr.-Ing. Jörg Söhner

**Beitrag zur Simulation zerspanungstechnologischer Vorgänge mit Hilfe der Finite-Element-Methode**

Band 119

Dr.-Ing. Judith Elsner

**Informationsmanagement für mehrstufige Mikro-Fertigungsprozesse**

Band 120

Dr.-Ing. Lijing Xie

**Estimation Of Two-dimension Tool Wear Based On Finite Element Method**

Band 121

Dr.-Ing. Ansgar Blessing

**Geometrischer Entwurf mikromechatronischer Systeme**

Band 122

Dr.-Ing. Rainer Ebner

**Steigerung der Effizienz mehrachsiger Fräspannprozesse durch neue Planungsmethoden mit hoher Benutzerunterstützung**

Band 123

Dr.-Ing. Silja Klinkel

**Multikriterielle Feinplanung in teilautonomen Produktionsbereichen – Ein Beitrag zur produkt- und prozessorientierten Planung und Steuerung**

Band 124

Dr.-Ing. Wolfgang Neithardt

**Methodik zur Simulation und Optimierung von Werkzeugmaschinen in der Konzept- und Entwurfsphase auf Basis der Mehrkörpersimulation**

Band 125

Dr.-Ing. Andreas Mehr

**Hartfeinbearbeitung von Verzahnungen mit kristallinen diamantbeschichteten Werkzeugen beim Fertigungsverfahren Wälzstoßen**

Band 126

Dr.-Ing. Martin Gutmann

**Entwicklung einer methodischen Vorgehensweise zur Diagnose von hydraulischen Produktionsmaschinen**

Band 127

Dr.-Ing. Gisela Lanza

**Simulative Anlaufunterstützung auf Basis der Qualitätsfähigkeiten von Produktionsprozessen**

Band 128

Dr.-Ing. Ulf Dambacher

**Kugelgewindetrieb mit hohem Druckwinkel**

Band 129

Dr.-Ing. Carsten Buchholz

**Systematische Konzeption und Aufbau einer automatisierten Produktionszelle für pulverspritzgegossene Mikrobauteile**

Band 130

Dr.-Ing. Heiner Lang

**Trocken-Räumen mit hohen Schnittgeschwindigkeiten**

Band 131

Dr.-Ing. Daniel Nesges

**Prognose operationeller Verfügbarkeiten von Werkzeugmaschinen unter Berücksichtigung von Serviceleistungen**

---

Im Shaker Verlag erschienene Bände:

Band 132

Dr.-Ing. Andreas Bechle

**Beitrag zur prozesssicheren Bearbeitung beim Hochleistungs-fertigungsverfahren Wälzschälen**

Band 133

Dr.-Ing. Markus Herm

**Konfiguration globaler Wertschöpfungsnetzwerke auf Basis von Business Capabilities**

Band 134

Dr.-Ing. Hanno Tritschler

**Werkzeug- und Zerspanprozessoptimierung beim Hartfräsen  
von Mikrostrukturen in Stahl**

Band 135

Dr.-Ing. Christian Munzinger

**Adaptronische Strebe zur Steifigkeitssteigerung  
von Werkzeugmaschinen**

Band 136

Dr.-Ing. Andreas Stepping

**Fabrikplanung im Umfeld von Wertschöpfungsnetzwerken und  
ganzheitlichen Produktionssystemen**

Band 137

Dr.-Ing. Martin Dyck

**Beitrag zur Analyse thermische bedingter Werkstückdeformationen  
in Trockenbearbeitungsprozessen**

Band 138

Dr.-Ing. Siegfried Schmalzried

**Dreidimensionales optisches Messsystem für eine effizientere  
geometrische Maschinenbeurteilung**

Band 139

Dr.-Ing. Marc Wawerla

**Risikomanagement von Garantieleistungen**

Band 140

Dr.-Ing. Ivesa Buchholz

**Strategien zur Qualitätssicherung mikromechanischer Bauteile  
mittels multisensorieller Koordinatenmesstechnik**

Band 141

Dr.-Ing. Jan Kotschenreuther

**Empirische Erweiterung von Modellen der Makrozerspanung  
auf den Bereich der Mikrobearbeitung**

Band 142

Dr.-Ing. Andreas Knödel

**Adaptronische hydrostatische Drucktascheneinheit**

Band 143

Dr.-Ing. Gregor Stengel

**Fliegendes Abtrennen räumlich gekrümmter Strangpressprofile mittels Industrierobotern**

Band 144

Dr.-Ing. Udo Weismann

**Lebenszyklusorientiertes interorganisationelles Anlagencontrolling**

Band 145

Dr.-Ing. Rüdiger Pabst

**Mathematische Modellierung der Wärmestromdichte zur Simulation des thermischen Bauteilverhaltens bei der Trockenbearbeitung**

Band 146

Dr.-Ing. Jan Wieser

**Intelligente Instandhaltung zur Verfügbarkeitssteigerung von Werkzeugmaschinen**

Band 147

Dr.-Ing. Sebastian Haupt

**Effiziente und kostenoptimale Herstellung von Mikrostrukturen durch eine Verfahrenskombination von Bahnerosion und Laserablation**

Band 148

Dr.-Ing. Matthias Schlipf

**Statistische Prozessregelung von Fertigungs- und Messprozess zur Erreichung einer variabilitätsarmen Produktion mikromechanischer Bauteile**

Band 149

Dr.-Ing. Jan Philipp Schmidt-Ewig

**Methodische Erarbeitung und Umsetzung eines neuartigen Maschinenkonzeptes zur produktflexiblen Bearbeitung räumlich gekrümmter Strangpressprofile**

Band 150

Dr.-Ing. Thomas Ender

**Prognose von Personalbedarfen im Produktionsanlauf unter Berücksichtigung dynamischer Planungsgrößen**

Band 151

Dr.-Ing. Kathrin Peter

**Bewertung und Optimierung der Effektivität von Lean Methoden  
in der Kleinserienproduktion**

Band 152

Dr.-Ing. Matthias Schopp

**Sensorbasierte Zustandsdiagnose und -prognose von Kugelgewindetrieben**

Band 153

Dr.-Ing. Martin Kipfmüller

**Aufwandsoptimierte Simulation von Werkzeugmaschinen**

Band 154

Dr.-Ing. Carsten Schmidt

**Development of a database to consider multi wear mechanisms  
within chip forming simulation**

Band 155

Dr.-Ing. Stephan Niggeschmidt

**Ausfallgerechte Ersatzteilbereitstellung im Maschinen- und Anlagenbau  
mittels lastabhängiger Lebensdauerprognose**

Band 156

Dr.-Ing. Jochen Conrad Peters

**Bewertung des Einflusses von Formabweichungen in der  
Mikro-Koordinatenmesstechnik**

Band 157

Dr.-Ing. Jörg Ude

**Entscheidungsunterstützung für die Konfiguration  
globaler Wertschöpfungsnetzwerke**

Band 158

Dr.-Ing. Stefan Weiler

**Strategien zur wirtschaftlichen Gestaltung der globalen Beschaffung**

Band 159

Dr.-Ing. Jan Rühl

**Monetäre Flexibilitäts- und Risikobewertung**

Band 160

Dr.-Ing. Daniel Ruch

**Positions- und Konturerfassung räumlich gekrümmter Profile auf Basis bauteilimmanenter Markierungen**

Band 161

Dr.-Ing. Manuel Tröndle

**Flexible Zuführung von Mikrobauteilen mit piezoelektrischen Schwingförderern**

Band 162

Dr.-Ing. Benjamin Viering

**Mikroverzahnungsnormal**

Band 163

Dr.-Ing. Chris Becke

**Prozesskraftrichtungsangepasste Frästrategien zur schädigungsarmen Bohrungsbearbeitung an faserverstärkten Kunststoffen**

Band 164

Dr.-Ing. Patrick Werner

**Dynamische Optimierung und Unsicherheitsbewertung der lastabhängigen präventiven Instandhaltung von Maschinenkomponenten**

Band 165

Dr.-Ing. Martin Weis

**Kompensation systematischer Fehler bei Werkzeugmaschinen durch self-sensing Aktoren**

Band 166

Dr.-Ing. Markus Schneider

**Kompensation von Konturabweichungen bei gerundeten Strangpressprofilen durch robotergestützte Führungswerzeuge**

Band 167

Dr.-Ing. Ester M. R. Ruprecht

**Prozesskette zur Herstellung schichtbasierter Systeme mit integrierten Kavitäten**

Band 168

Dr.-Ing. Alexander Broos

**Simulationsgestützte Ermittlung der Komponentenbelastung für die Lebensdauerprognose an Werkzeugmaschinen**

Band 169

Dr.-Ing. Frederik Zanger

**Segmentspanbildung, Werkzeugverschleiß, Randschichtzustand und Bauteileigenschaften: Numerische Analysen zur Optimierung des Zerspanungsprozesses am Beispiel von Ti-6Al-4V**

Band 170

Dr.-Ing. Benjamin Behmann

**Servicefähigkeit**

Band 171

Dr.-Ing. Annabel Gabriele Jondral

**Simulationsgestützte Optimierung und Wirtschaftlichkeitsbewertung des Lean-Methodeneinsatzes**

Band 172

Dr.-Ing. Christoph Ruhs

**Automatisierte Prozessabfolge zur qualitätssicheren Herstellung von Kavitäten mittels Mikrobahnerosion**

Band 173

Dr.-Ing. Steven Peters

**Markoffsche Entscheidungsprozesse zur Kapazitäts- und Investitionsplanung von Produktionssystemen**

Band 174

Dr.-Ing. Christoph Kühlewein

**Untersuchung und Optimierung des Wälzschälverfahrens mit Hilfe von 3D-FEM-Simulation – 3D-FEM Kinematik- und Spanbildungssimulation**

Band 175

Dr.-Ing. Adam-Mwanga Dieckmann

**Auslegung und Fertigungsprozessgestaltung sintergefügter Verbindungen für µMIM-Bauteile**

Band 176

Dr.-Ing. Heiko Hennrich

**Aufbau eines kombinierten belastungs- und zustandsorientierten Diagnose- und Prognosesystems für Kugelgewindetriebe**

Band 177

Dr.-Ing. Stefan Herder

**Piezoelektrischer Self-Sensing-Aktor zur Vorspannungsregelung in adaptronischen Kugelgewindetrieben**

Band 178

Dr.-Ing. Alexander Ochs

**Ultraschall-Strömungsgreifer für die Handhabung textiler Halbzeuge bei der automatisierten Fertigung von RTM-Bauteilen**

Band 179

Dr.-Ing. Jürgen Michna

**Numerische und experimentelle Untersuchung zerspanungsbedingter Gefügeumwandlungen und Modellierung des thermo-mechanischen Lastkollektivs beim Bohren von 42CrMo4**

Band 180

Dr.-Ing. Jörg Elser

**Vorrichtungsfreie räumliche Anordnung von Fügepartnern auf Basis von Bauteilmarkierungen**

Band 181

Dr.-Ing. Katharina Klimscha

**Einfluss des Fügespalts auf die erreichbare Verbindungsqualität beim Sinterfügen**

Band 182

Dr.-Ing. Patricia Weber

**Steigerung der Prozesswiederholbarkeit mittels Analyse akustischer Emissionen bei der Mikrolaserablation mit UV-Pikosekundenlasern**

Band 183

Dr.-Ing. Jochen Schädel

**Automatisiertes Fügen von Tragprofilen mittels Faserwickeln**

Band 184

Dr.-Ing. Martin Krauß

**Aufwandsoptimierte Simulation von Produktionsanlagen durch Vergrößerung der Geltungsbereiche von Teilmodellen**

Band 185

Dr.-Ing. Raphael Moser

**Strategische Planung globaler Produktionsnetzwerke**

Bestimmung von Wandlungsbedarf und Wandlungszeitpunkt mittels multikriterieller Optimierung

Band 186

Dr.-Ing. Martin Otter

**Methode zur Kompensation fertigungsbedingter Gestaltabweichungen für die Montage von Aluminium Space-Frame-Strukturen**

Band 187

Dr.-Ing. Urs Leberle

**Produktive und flexible Gleitförderung kleiner Bauteile auf phasenflexiblen Schwingförderern mit piezoelektrischen 2D-Antriebselementen**

Band 188

Dr.-Ing. Johannes Book

**Modellierung und Bewertung von Qualitätsmanagementstrategien in globalen Wertschöpfungsnetzwerken**

Band 189

Dr.-Ing. Florian Ambrosy

**Optimierung von Zerspanungsprozessen zur prozesssicheren Fertigung nanokristalliner Randschichten am Beispiel von 42CrMo4**

Band 190

Dr.-Ing. Adrian Kölmel

**Integrierte Messtechnik für Prozessketten unreifer Technologien am Beispiel der Batterieproduktion für Elektrofahrzeuge**

Band 191

Dr.-Ing. Henning Wagner

**Featurebasierte Technologieplanung zum Preforming von textilen Halbzeugen**

Band 192

Dr.-Ing. Johannes Gebhardt

**Strukturoptimierung von in FVK eingebetteten metallischen Lasteinleitungselementen**

Band 193

Dr.-Ing. Jörg Bauer

**Hochintegriertes hydraulisches Vorschubsystem für die Bearbeitung kleiner Werkstücke mit hohen Fertigungsanforderungen**

Band 194

Dr.-Ing. Nicole Stricker

**Robustheit verketteter Produktionssysteme**

Robustheitsevaluation und Selektion des Kennzahlensystems der Robustheit

Band 195

Dr.-Ing. Anna Sauer

**Konfiguration von Montagelinien unreifer Produkttechnologien am Beispiel der Batteriemontage für Elektrofahrzeuge**

Band 196

Dr.-Ing. Florian Sell-Le Blanc

**Prozessmodell für das Linearwickeln unrunder Zahnpulpen**

Ein Beitrag zur orthozyklischen Spulenwickeltechnik

Band 197

Dr.-Ing. Frederic Förster

**Geregeltes Handhabungssystem zum zuverlässigen und energieeffizienten Handling textiler Kohlenstofffaserzuschnitte**

Band 198

Dr.-Ing. Nikolay Boev

**Numerische Beschreibung von Wechselwirkungen zwischen Zerspanprozess und Maschine am Beispiel Räumen**

Band 199

Dr.-Ing. Sebastian Greinacher

**Simulationsgestützte Mehrzieloptimierung schlanker und ressourceneffizienter Produktionssysteme**

Band 200

Dr.-Ing. Benjamin Häfner

**Lebensdauerprognose in Abhängigkeit der Fertigungsabweichungen bei Mikroverzahnungen**

Band 201

Dr.-Ing. Stefan Klotz

**Dynamische Parameteranpassung bei der Bohrungsherstellung in faserverstärkten Kunststoffen unter zusätzlicher Berücksichtigung der Einspannsituation**

Band 202

Dr.-Ing. Johannes Stoll

**Bewertung konkurrierender Fertigungsfolgen mittels Kostensimulation und stochastischer Mehrzieloptimierung**

Anwendung am Beispiel der Blechpaketfertigung für automobile Elektromotoren

Band 203

Dr.-Ing. Simon-Frederik Koch

**Fügen von Metall-Faserverbund-Hybridwellen im Schleuderverfahren ein Beitrag zur fertigungsgerechten intrinsischen Hybridisierung**

Band 204

Dr.-Ing. Julius Ficht

**Numerische Untersuchung der Eigenspannungsentwicklung für sequenzielle Zerspanungsprozesse**

Band 205

Dr.-Ing. Manuel Baumeister

**Automatisierte Fertigung von Einzelblattstapeln in der Lithium-Ionen-Zellproduktion**

Band 206

Dr.-Ing. Daniel Bertsch

**Optimierung der Werkzeug- und Prozessauslegung für das Wälzschälen von Innenverzahnungen**

Band 207

Dr.-Ing. Kyle James Kippenbrock

**Deconvolution of Industrial Measurement and Manufacturing Processes  
for Improved Process Capability Assessments**

Band 208

Dr.-Ing. Farboud Bejnoud

**Experimentelle Prozesskettenbetrachtung für Räumbauteile am Beispiel  
einer einsatzgehärteten PKW-Schiebemuffe**

Band 209

Dr.-Ing. Steffen Dosch

**Herstellungsübergreifende Informationsübertragung zur effizienten Produktion  
von Werkzeugmaschinen am Beispiel von Kugelgewindetrieben**

Band 210

Dr.-Ing. Emanuel Moser

**Migrationsplanung globaler Produktionsnetzwerke**

Bestimmung robuster Migrationspfade und risiko-effizienter Wandlungsbefähiger

Band 211

Dr.-Ing. Jan Hochdörffer

**Integrierte Produktallokationsstrategie und Konfigurationssequenz in  
globalen Produktionsnetzwerken**

Band 212

Dr.-Ing. Tobias Arndt

**Bewertung und Steigerung der Prozessqualität in globalen  
Produktionsnetzwerken**

Band 213

Dr.-Ing. Manuel Peter

**Unwuchtmiminale Montage von Permanentmagnetrotoren durch modellbasierte  
Online-Optimierung**

Band 214

Dr.-Ing. Robin Kopf

**Kostenorientierte Planung von Fertigungsfolgen additiver Technologien**

Band 215

Dr.-Ing. Harald Meier

**Einfluss des Räumens auf den Bauteilzustand in der Prozesskette  
Weichbearbeitung – Wärmebehandlung – Hartbearbeitung**

Band 216

Dr.-Ing. Daniel Brabandt

**Qualitätssicherung von textilen Kohlenstofffaser-Preforms mittels  
optischer Messtechnik**

Band 217

Dr.-Ing. Alexandra Schabunow

**Einstellung von Aufnahmeparametern mittels projektionsbasierter Qualitäts-  
kenngrößen in der industriellen Röntgen-Computertomographie**

Band 218

Dr.-Ing. Jens Bürgin

**Robuste Auftragsplanung in Produktionsnetzwerken**

Mittelfristige Planung der variantenreichen Serienproduktion unter Unsicherheit  
der Kundenauftragskonfigurationen

Band 219

Dr.-Ing. Michael Gerstenmeyer

**Entwicklung und Analyse eines mechanischen Oberflächenbehandlungs-  
verfahrens unter Verwendung des Zerspanungswerkzeuges**

Band 220

Dr.-Ing. Jacques Burtscher

**Erhöhung der Bearbeitungsstabilität von Werkzeugmaschinen durch  
semi-passive masseneinstellbare Dämpfungssysteme**

Band 221

Dr.-Ing. Dietrich Berger

**Qualitätssicherung von textilen Kohlenstofffaser-Preforms mittels prozess-  
integrierter Wirbelstromsensor-Arrays**

Band 222

Dr.-Ing. Fabian Johannes Ballier

**Systematic gripper arrangement for a handling device in lightweight production processes**

Band 223

Dr.-Ing. Marielouise Schäferling, geb. Zaiß

**Development of a Data Fusion-Based Multi-Sensor System for Hybrid Sheet Molding Compound**

Band 224

Dr.-Ing. Quirin Spiller

**Additive Herstellung von Metallbauteilen mit dem ARBURG Kunststoff-Freiformen**

Band 225

Dr.-Ing. Andreas Spohrer

**Steigerung der Ressourceneffizienz und Verfügbarkeit von Kugelgewinde-trieben durch adaptive Schmierung**

Band 226

Dr.-Ing. Johannes Fisel

**Veränderungsfähigkeit getakteter Fließmontagesysteme**

Planung der Fließbandabstimmung am Beispiel der Automobilmontage

Band 227

Dr.-Ing. Patrick Bollig

**Numerische Entwicklung von Strategien zur Kompensation thermisch bedingter Verzüge beim Bohren von 42CrMo4**

Band 228

Dr.-Ing. Ramona Pfeiffer, geb. Singer

**Untersuchung der prozessbestimmenden Größen für die anforderungsgerechte Gestaltung von Pouchzellen-Verpackungen**

Band 229

Dr.-Ing. Florian Baumann

**Additive Fertigung von endlosfaserverstärkten Kunststoffen mit dem ARBURG Kunststoff-Freiform Verfahren**

Band 230

Dr.-Ing. Tom Stähr

**Methodik zur Planung und Konfigurationsauswahl skalierbarer Montagesysteme – Ein Beitrag zur skalierbaren Automatisierung**

Band 231

Dr.-Ing. Jan Schwennen

**Einbringung und Gestaltung von Lasteinleitungsstrukturen für im RTM-Verfahren hergestellte FVK-Sandwichbauteile**

Band 232

Dr.-Ing. Sven Coutandin

**Prozessstrategien für das automatisierte Preforming von bebinderten textilen Halbzeugen mit einem segmentierten Werkzeugsystem**

Band 233

Dr.-Ing. Christoph Liebrecht

**Entscheidungsunterstützung für den Industrie 4.0-Methodeneinsatz**  
Strukturierung, Bewertung und Ableitung von Implementierungsreihenfolgen

Band 234

Dr.-Ing. Stefan Treber

**Transparenzsteigerung in Produktionsnetzwerken**

Verbesserung des Störungsmanagements durch verstärkten Informationsaustausch

Band 235

Dr.-Ing. Marius Dackweiler

**Modellierung des Fügewickelprozesses zur Herstellung von leichten Fachwerkstrukturen**

Band 236

Dr.-Ing. Fabio Echsler Minguillon

**Prädiktiv-reaktives Scheduling zur Steigerung der Robustheit in der Matrix-Produktion**

Band 237

Dr.-Ing. Sebastian Haag

**Entwicklung eines Verfahrensablaufes zur Herstellung von Batteriezellstapeln mit großformatigem, rechteckigem Stapelformat und kontinuierlichen Materialbahnen**

Band 238

Dr.-Ing. Raphael Wagner

**Strategien zur funktionsorientierten Qualitätsregelung in der Serienproduktion**

Band 239

Dr.-Ing. Christopher Ehrmann

**Ausfallfrüherkennung von Ritzel-Zahnstangen- Trieben mittels Acoustic Emission**

Band 240

Dr.-Ing. Janna Hofmann

**Prozessmodellierung des Fünf-Achs-Nadelwickelns zur Implementierung einer trajektoriebasierten Drahtzugkraftregelung**

Band 241

Dr.-Ing. Andreas Kuhnle

**Adaptive Order Dispatching based on Reinforcement Learning Application in a Complex Job Shop in the Semiconductor Industry**

Band 242

Dr.-Ing. Andreas Greiber

**Fertigung optimierter technischer Oberflächen durch eine Verfahrenskombination aus Fliehkraft-Tauchgleitschleifen und Laserablation**  
Prozesseinflüsse und Prozessauslegung

Band 243

Dr.-Ing. Jan Niclas Eschner

**Entwicklung einer akustischen Prozessüberwachung zur Porenbestimmung im Laserstrahlschmelzen**

Band 244

Dr.-Ing. Sven Roth

**Schädigungsfreie Anbindung von hybriden FVK/Metall-Bauteilen an metallische Tragstrukturen durch Widerstandspunktschweißen**

Band 245

Dr.-Ing. Sina Kathrin Peukert

**Robustheitssteigerung in Produktionsnetzwerken mithilfe eines integrierten Störungsmanagements**

Band 246

Dr.-Ing. Alexander Jacob

**Hochiterative Technologieplanung**

Rekursive Optimierung produkt- und fertigungsbezogener  
Freiheitsgrade am Beispiel der hybrid-additiven Fertigung

Band 247

Dr.-Ing. Patrick Moll

**Ressourceneffiziente Herstellung von Langfaser-Preforms  
im Faserblasverfahren**

Band 248

Dr.-Ing. Eric Thore Segebade

**Erhöhung der Verschleißbeständigkeit von Bauteilen aus Ti-6Al-4V mittels  
simulationsgestützter Zerspanung und mechanischer Mikrotextrurierung**

Band 249

Dr.-Ing. Shun Yang

**Regionalized implementation strategy of smart  
automation within assembly systems in China**

Band 250

Dr.-Ing. Constantin Carl Hofmann

**Vorausschauende und reaktive Mehrzieloptimierung  
für die Produktionssteuerung einer Matrixproduktion**

Band 251

Dr.-Ing. Paul Ruhland

**Prozesskette zur Herstellung von hybriden Faser-Metall-Preforms**

Modellbildung und Optimierung des Binderauftrags und  
der Drapierung für stabförmige Bauteile

Band 252

Dr.-Ing. Leonard Schild

**Erzeugung und Verwendung von Anwendungswissen in der industriellen  
Computertomographie**

Band 253

Dr.-Ing. Benedikt Klee

**Analyse von Phaseninformationen in Videodaten zur Identifikation  
von Schwingungen in Werkzeugmaschinen**

Band 254

Dr.-Ing. Bruno Vargas

**Wälzschälen mit kleinen Achskreuzwinkeln**

Prozessgrenzen und Umsetzbarkeit

Band 255

Dr.-Ing. Lucas Bretz

**Function-oriented in-line quality assurance of hybrid sheet molding compound**

Band 256

Dr.-Ing. Bastian Rothaupt

**Dämpfung von Bauteilschwingungen durch einstellbare Werkstückdirektspannung mit Hydrodehnspanntechnik**

Band 257

Dr.-Ing. Daniel Kupzik

**Robotic Swing Folding of three-dimensional UD-tape-based Reinforcement Structures**

Band 258

Dr.-Ing. Bastian Verhaelen

**(De-)Zentralisierung von Entscheidungen in globalen Produktionsnetzwerken**

Strategie- und Komplexitätsorientierte Gestaltung der Entscheidungsautonomie

Band 259

Dr.-Ing. Hannes Wilhelm Weinmann

**Integration des Vereinzelungs- und Stapelbildungsprozesses in ein flexibel und kontinuierlich arbeitendes Anlagenmodul für die Li-Ionen Batteriezellfertigung**

Band 260

Dr.-Ing. Florian Stamer

**Dynamische Lieferzeit-Preisgestaltung in variantenreicher Produktion**

Ein adaptiver Ansatz mithilfe von Reinforcement Learning





