

TECHNISCHE UNIVERSITÄT MÜNCHEN
TUM School of Engineering and Design

**System zur energieorientierten und kostenbasierten
Produktionssteuerung mittels Reinforcement Learning**

Martin Walter Rösch

Vollständiger Abdruck der von der TUM School of Engineering and Design
der Technischen Universität München zur Erlangung eines

Doktors der Ingenieurwissenschaften (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr.-Ing. Harmut Spliethoff

Prüfer der Dissertation:

1. Prof. Dr.-Ing. Gunther Reinhart
2. Prof. Dr.-Ing. Birgit Vogel-Heuser

Die Dissertation wurde am 25.03.2021 bei der Technischen Universität
München eingereicht und durch die TUM School of Engineering and De-
sign am 26.07.2021 angenommen.

Vorwort

Die vorliegende Dissertation entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Fraunhofer-Institut für Gießerei-, Composite- und Verarbeitungstechnik IGCV in Augsburg sowie am Institut für Werkzeugmaschinen und Betriebswissenschaften (iwb) der Technischen Universität München.

Mein besonderer Dank gilt zunächst meinem Doktorvater Herrn Prof. Dr.-Ing. Gunther Reinhart für die wohlwollende Unterstützung dieser Arbeit sowie für die kontinuierliche Förderung meiner Tätigkeit am Institut. Frau Professor Dr.-Ing. Birgit Vogel-Heuser, der Leiterin des Lehrstuhls für Automatisierung und Informationssysteme an der Technischen Universität München und Herrn Prof. Dr.-Ing. Harmut Spliethoff, dem Leiter des Lehrstuhls Energiesysteme ebenfalls an der Technischen Universität München, danke ich für die Übernahme des Korreferats bzw. des Prüfungsvorsitzes.

Für die wertvollen Hinweise und die gründliche Durchsicht meiner Arbeit möchte ich mich besonders bei Dr.-Ing. Peter Simon, Dr.-Ing. Cedric Schultz und Dr.-Ing. Johannes Graf bedanken. Darüber hinaus gilt ein besonderer Dank allen Studierenden, die mich bei der Entstehung dieser Arbeit unterstützt haben. Außerdem möchte ich mich bei allen Kolleginnen und Kollegen bedanken, die meine Zeit am Institut zu einem ganz besonderen Lebensabschnitt für mich machten.

Ohne den bedingungslosen Rückhalt und die Unterstützung durch meine Freundin und meine Familie wäre diese Arbeit sicherlich nicht möglich gewesen, herzlichen Dank dafür! Ein letztes Dankeschön gilt zudem der deutschen Gesellschaft, die mir diese hervorragende Ausbildung ermöglicht hat, welche nun mit der vorliegenden Arbeit abgeschlossen wird. Ich hoffe, ich kann während meiner weiteren Laufbahn etwas zurückgeben und dazu beitragen, die kommenden gesellschaftlichen Herausforderungen zu meistern.

Inhaltsverzeichnis

Abkürzungsverzeichnis.....	x
Verzeichnis der Formelzeichen	ix
1 Einleitung.....	1
1.1 Ausgangssituation	1
1.2 Motivation der Arbeit	4
1.3 Zielsetzung und Aufbau der Arbeit	6
2 Grundlagen.....	11
2.1 Übersicht	11
2.2 Elektrischer Strom in produzierenden Unternehmen	11
2.2.1 Lastgang und Energieflexibilität	11
2.2.2 Stromeigenerzeugung.....	12
2.2.3 Energiespeicher	13
2.2.4 Strombeschaffung.....	14
2.3 Produktionsplanung und -steuerung	16
2.3.1 Aufgaben und Abgrenzung	16
2.3.2 Zielgrößen	19
2.4 Ansätze zur Produktionssteuerung.....	22
2.4.1 Problemformulierung Job Shop Scheduling	22
2.4.2 Steuerungsstrategien	22
2.4.3 Reaktive Produktionssteuerung.....	23
2.4.4 Prädiktiv-reaktive Produktionssteuerung	25
2.4.5 Agentensysteme	26

2.5	Reinforcement Learning	27
2.5.1	Einführung	27
2.5.2	Erlernen der optimalen Strategie	28
2.5.3	Lernverfahren	29
2.5.4	Reinforcement Learning in Agentensystemen	31
2.6	Zusammenfassung und Fazit.....	32
3	Stand der Forschung	33
3.1	Übersicht	33
3.2	Energieflexibilität in der Produktion.....	33
3.2.1	Einordnung der betrachteten Forschungsfelder.....	33
3.2.2	Einordnung der bestehenden Ansätze	35
3.2.3	Energieflexibilität in der Produktionsplanung	36
3.2.4	Energieflexibilität in der Produktionssteuerung.....	40
3.2.5	Lastmanagement im Umfeld produzierender Unternehmen	44
3.3	Gesamtkosten als Zielgröße in der PPS	46
3.4	Ableitung des Handlungsbedarfs	50
4	System zur energieorientierten und kostenbasierten Produktionssteuerung	53
4.1	Ableitung der Anforderungen	53
4.2	Auswahl des Steuerungsansatzes	54
4.3	Systemübersicht	56
4.4	Systemelemente.....	57

5	Integration des kurzfristigen Energiebezugs in die Produktionssteuerung	59
5.1	Einordnung der steuerungsrelevanten Bezugsoptionen.....	59
5.1.1	Relevante Energiebezugsoptionen	59
5.1.2	Parameter zur Beschreibung der Energiebezugsoptionen.....	60
5.1.3	Kategorisierung	61
5.2	Aggregation und Modellierung.....	62
5.2.1	Fluktuierende Eigenerzeugung.....	62
5.2.2	Regelbare Eigenerzeugung.....	64
5.2.3	Stromspeicher.....	66
5.2.4	Netzbezug.....	69
5.2.5	Stromhandel am Intraday-Markt.....	70
5.2.6	Zusammenfassende Betrachtung des Energiebezugs.....	71
5.3	Ableitung des Steuerungsschemas.....	72
5.3.1	Ad-hoc Strombezug	73
5.3.2	Kurzfristiger Stromhandel.....	76
5.4	Fazit.....	77
6	Gesamtkostenbasiertes Zielsystem.....	79
6.1	Vorauswahl der Kostenfaktoren	79
6.2	Analyse der durch die Produktionssteuerung beeinflussbaren Kostenverursacher	81
6.3	Ableiten der monetären Auswirkungen	85
6.4	Modellierung der Kostenbestandteile	87
6.4.1	Terminbezogene Kosten.....	87
6.4.2	Erhöhte Materialkosten	90

6.4.3	Zusätzlich Zinskosten	91
6.4.4	Erhöhte Personalkosten	91
6.4.5	Funktion des gesamtkostenbasierten Zielsystems	92
6.5	Datenmodell	93
6.6	Fazit	94
7	Integrierte Produktionssteuerung mittels Reinforcement Learning	95
7.1	Aufbau und Ablauf	95
7.2	Modellierung des Produktionssystems	97
7.2.1	Produktionsressourcen	97
7.2.2	Produktionsaufträge	98
7.2.3	Stochastische Ereignisse	99
7.3	RL-Agenten	100
7.3.1	Allgemeine Gestaltung des Zustandsraums	100
7.3.2	Aktions- und Zustandsraum der Ressourcenagenten	105
7.3.3	Aktions- und Zustandsraum des Speicheragenten	108
7.3.4	Aktions- und Zustandsraum des Marktagenten	110
7.4	Belohnungsfunktion und -vergabe	112
7.4.1	Globale und lokale Belohnung	112
7.4.2	Energiekostenbestandteil der Belohnung	114
7.4.3	Produktionskostenbestandteil der Belohnung	114
7.4.4	Resultierende Belohnungsvergabe	119
7.5	Lernverfahren	120
7.6	Trainingsumgebung	122
7.7	Fazit	124

8	Prototypische Umsetzung und Evaluierung.....	125
8.1	Beschreibung des Anwendungsbeispiels	125
8.1.1	Allgemeines.....	125
8.1.2	Produktionssystem	126
8.1.3	Energiebezugsoptionen	128
8.2	Technische Umsetzung	131
8.2.1	Softwarestruktur	131
8.2.2	Simulationsablauf und Training	132
8.3	Simulationstechnische Evaluierung.....	134
8.3.1	Evaluierungsszenarien.....	134
8.3.2	Ergebnisse	139
8.4	Nutzen und Wirtschaftlichkeitsbetrachtung	147
9	Zusammenfassung und Ausblick	153
9.1	Zusammenfassung	153
9.2	Ausblick	155
10	Anhang.....	157
10.1	Parameter des Batteriemodells.....	157
10.2	Konfiguration MARL	157
10.3	Produktionsdaten der Evaluierungsepisoden.....	158
10.4	Energiedaten der Evaluierungsepisoden.....	164
10.5	Regelbasierte Betriebsstrategie von Eigenerzeugung und Speicher (RBS)	166
10.6	Parametrierung von Simulated Annealing	167

11	Studienarbeiten	169
12	Literaturverzeichnis	171

Abkürzungsverzeichnis

AS.....	Agentensystem
BHKW	Blockheizkraftwerk
DDPG.....	Deep Deterministic Gradient Descent
DoD.....	Depth of Discharge
DQL	Deep Q-Learning
EDF	Earliest Deadline First
EE.....	erneuerbare Energien
EEG.....	Erneuerbare-Energien-Gesetz
EVU	Energieversorgungsunternehmen
FEE	Fluktuierende Eigenerzeugung
FIFO.....	First-In-First-Out
JSSP	Job Shop Scheduling Problem
KWK.....	Kraft-Wärme-Kopplung
MARL	Multi-Agent Reinforcement Learning
OTC.....	Over-The-Counter
PPO	Proximal Policy Optimization
PPS	Produktionsplanung und -steuerung
PV.....	Photovoltaic
RBS	regelbasierte Betriebsstrategie
REE.....	regelbare Eigenerzeugung
RL.....	Reinforcement Learning
SA.....	Simulated Annealing
SoC.....	State of Charge
TGA	technische Gebäudeausrüstung
VDI	Verein Deutscher Ingenieure e.V.
ES	Evaluierungsszenario
WKA.....	Windkraftanlagen

Verzeichnis der Formelzeichen

Große lateinische Buchstaben

Symbol	Einheit	Bedeutung
$A_{BA,t}$	-	Menge an Aktionen, die zum Zeitpunkt t nicht ausgeführt werden können
$AV_{m,r}$	kg	Anfahrverluste an Material m bei Rüstvorgang r
C_N	kWh	Nennkapazität der Batterie
ΔL_i	-	Veränderung der Alterung im Zyklus i
D_n^A	-	Dauer von Versandart A und Auftrag n
$E_{AE}(t)$	kWh	Vom Netz bezogene Ausgleichsenergiemenge in der Periode t
$E_I(t)$	kWh	Vom Intraday-Markt bezogene Strommenge im Zeitintervall t
E_π	-	Erwartungswert der Strategie π
$E_V(t)$	kWh	Vorab beschaffte Energiemenge für das Zeitintervall t
$GK(t)$	€	Gesamtkosten für die Fertigung von Auftrag n
INV_{Bat}	€	Investitionskosten der Batterie
$K_{AE}(t)$	€	Kosten für vom Netz bezogene Ausgleichsenergie in der Periode t
$K_{BHKW}(t)$	€	Stromkosten des BHKW zum Zeitpunkt t
$K_{Bat,i}$	€	Resultierende Alterungskosten der Batterie pro Zyklus i
$K_I(t)$	€	Kosten für Strom aus Intraday-Handel Zeitintervall t
$K_{LS}(t)$	€	Kosten für Lastspitze im 15-Minuten-Intervall t
$K_{P,I}$	€	Gebührenpauschale pro Order für Intraday-Handel
KR_r	€	Kosten von Rüstvorgang r

Verzeichnis der Formelzeichen

L	-	Alterungszustand der Batterie
LS_{max}	kW	aktuelle, kostenwirksame Lastspitze
$P_{batt}(t)$	kW	Lade-/Entladeleistung der Batterie zum Zeitpunkt t
$P_{BHKW,min}$	kW	minimal mögliche Teillast der BHKW-Anlage
$P_{BHKW,N}$	kW	Nennleistung der BHKW-Anlagen
$P_{BHKW}(t)$	kW	Leistung der BHKW-Anlage zum Zeitpunkt t
$P_{ES}(t)$	kW	Durchschnittlich vom externen Stromnetz bezogene Leistung im Intervall t
PK_n	€	Zusätzliche Personalkosten für die Fertigung von Auftrag n
PS_n	h	Zusätzliche geleistete Personalstunden für die Fertigung von Auftrag n
R_n	-	Für die Fertigung von Auftrag n erforderliche Rüstvorgänge
$R_{t,a}^P$	-	Produktionskostenbestandteil der Belohnung für Produktionsressource p im Zeitpunkt t für Aktion a
R_s^E	-	Energiekostenbestandteil der Belohnung für die Schicht s
R_t	-	Erwartungswert der langfristigen Belohnung zum Zeitpunkt t
$S_\delta(\delta_i)$	-	DOD-Stressmodell
$SOC(t)$	-	Ladezustand zum Zeitpunkt t
$S_\sigma(\sigma_i)$	-	SOC-Stressmodell
$S_t(d)$	-	Zeitliches Stressmodell
$S_T(T)$	-	Temperatur-Stressmodell
TK_{AE}	kWh	Toleranzkorridor für Lastgangabweichung
$TK_n(t)$	€	Terminbezogene Kosten eines Auftrags n bei Fertigstellungstermin t

TP	K	Durchschnittliche Zelltemperatur
T_{ref}	$^{\circ}C$	Referenztemperatur
TT_n	h	Terminoleranz von Auftrag n
VK_n^A	€	Kosten für Versandart A und Auftrag n
$V(t)$	V	Batteriespannung zum Zeitpunkt t
WK_{n_p}	€	Terminbezogene Kosten des Auftrags n_p in der Warteschlange

Kleine lateinische Buchstaben:

Symbol	Einheit	Bedeutung
a_t	-	Aktion zum Zeitpunkt t
d_{n_p}	-	Erforderliche Bearbeitungsdauer von Auftrag n_p
d_a	-	Tatsächliche Dauer der abgeschlossenen Aktion a
$f_{d,i}$	-	Linearisierte Degradation pro Zyklus i
k_{AE}	$\text{€}/kWh$	Strafkostensatz für Abweichung
$k_{BHKW,0}$	ct/kWh	Nennstromgestehungskosten des BHKW
$k_{\delta 1}, k_{\delta 2}, k_{\delta 3}$	-	Empirische Koeffizienten für die DOD bedingte Alterung
$k_{I,t}$	$\text{€}/kWh$	Preis am Intraday-Markt in Zeitintervall t
k_{LS}	$\text{€}/kW$	Mit Energieversorger vereinbarter Leistungspreis
$k_{rel,t}$	$\text{€}/kWh$	Anteilige Gebühr beim Intraday-Handel
k_{σ}	-	SOC-Stress Koeffizient
k_{SU}	$\text{€}/kWh$	Kostensatz für Steuern und Umlagen
k_T	-	Empirischer Koeffizient für die temperaturbedingte Alterung
k_t	$1/s$	Empirischer Koeffizient für die kalendarische Alterung

Verzeichnis der Formelzeichen

lk_n	€/h	Spezifischer Lagerkostensatz von Auftrag n
mk_m	€/kg	Kostensatz von Material m
pk	€/h	zusätzliche geleistete Personalstunden
r_{BA}	-	Strafbelohnung einer Bad Action
$r(t)$	-	direkte Belohnung im Zeitschritt t
s_t	-	Zustand zum Zeitpunkt t
$t_{F,np}$	-	Sollfertigstellungstermin von Auftrag n
$t_{V,n}$	h	zeitlicher Verzug von Auftrag n
z	-	Anzahl der für die aktuelle Order betroffenen Zeitintervalle

Griechische Buchstaben:

Symbol	Einheit	Bedeutung
$\alpha_{sei}, \beta_{sei}$	-	Koeffizienten für das Ausbilden der Passivierungsschicht der Anode
δ		DoD des aktuellen Zyklus
η	-	Wirkungsgrad beim Laden beziehungsweise Entladen der Batterie
γ	-	Discount-Faktor
φ		Wirkungsabhängiger Korrekturfaktor des BHKW
π	-	Strategie
σ	-	Durchschnittlicher SOC im aktuellen Zyklus
σ_{ref}	-	Referenz SOC-Wert
vk_n	€/h	Spezifischer Verzugskostensatz von Auftrag n
$v_\pi(s)$	-	Zustands-Wertefunktion der Strategie π im Zustand s

1 Einleitung

1.1 Ausgangssituation

Die Ressourcen des Planeten Erde sind endlich. Die Verträglichkeit anthropogener Eingriffe in Ökosysteme ist begrenzt (EICHLER 1993, WEIZSÄCKER 1997). So fiel im Jahr 2018 mit dem „Earth Overshoot Day“ der Tag des Jahres, an dem der Mensch mehr Ressourcen verbraucht hat, als die Erde innerhalb eines Jahres regenerieren kann, bereits auf den 29. Juli (GLOBAL FOOTPRINTING NETWORK 2020). In Verbindung mit der wachsenden Weltbevölkerung stellt dies eine der größten Herausforderungen für die Weltwirtschaft dar (ABELE & REINHART 2011, BAUERNHANSL ET AL. 2014). Von besonderer Bedeutung ist dabei die Begrenzung des vom Menschen verursachten Klimawandels auf ein verträgliches Maß, der insbesondere dem hohen CO₂-Ausstoß zuzurechnen ist (UN 2015B). Die Folgen führen bereits heute zu vermehrten Auftreten extremer Wetterereignisse (MANN ET AL. 2017). Die zunehmend spürbaren Auswirkungen dieser Veränderungen setzen Politiker weltweit unter Druck. Im Zuge der Pariser Klimakonferenz hat sich daher die Weltgemeinschaft zu dem Ziel bekannt, die Erwärmung auf 2° C zu begrenzen (UN 2015A). Dies kann jedoch nur mit einer Reduzierung der Treibhausgasemissionen gelingen. Da 42 Prozent der globalen CO₂-Emissionen auf die Elektrizitätserzeugung zurückzuführen sind, ist somit eine grundlegende Transformation der fossilen Stromerzeugung hin zu erneuerbaren Energien (EE) erforderlich (UN 2015B, IEA 2017).

Als eine der ersten großen Industrienationen hat Deutschland die Weichen für eine emissionsfreie Stromversorgung gestellt (BMW I 2010). Im Zuge dessen soll der CO₂-Ausstoß im Vergleich zum Jahr 1990 bis 2050 um 80 Prozent gesenkt werden. Dies geht mit einem massiven Ausbau von EE, insbesondere Wind, Biomasse und Photovoltaik (PV) einher. So hat sich deren Anteil an der Gesamtstromerzeugung in den Jahren 2007 bis 2017 von 14 auf 36 Prozent mehr als verdoppelt (UMWELTBUNDESAMT 2018).

Die dabei erforderlichen Investitionen werden in Form der sogenannten EEG-Umlage auf die Endverbraucher verteilt (BMJV 2014). Für Betriebe mit einem Jahresverbrauch bis 20 Millionen kWh sind infolgedessen die Stromkosten seit dem Jahr 2000 von unter 6 Ct/kWh kontinuierlich auf über 15 Ct/kWh im Jahr 2017 angestiegen, wobei alleine auf die EEG-Umlage 40 Prozent der Stromkosten zurückzuführen sind (BDEW 2018). In Summe bedeutet dies eine Preissteigerung

Einleitung

von ca. 170 Prozent. Für industrielle Großverbraucher mit bis zu 70 Mio. MWh im Jahr war der vorläufige Kostenhöhepunkt im Jahr 2014 mit 10,48 ct/kWh erreicht. Der Preis verharrt mit 9,3 ct/kWh im Jahr 2017 jedoch weiterhin auf einem hohen Niveau (BDEW 2020). Dieser Sachverhalt wird auch im Vergleich zu den Strompreisen in anderen EU-Ländern deutlich (vgl. Abbildung 1). Deutschland weist dabei die im Vergleich höchsten Strompreise für Unternehmen auf (BMWi 2018, EUROSTAT 2018).

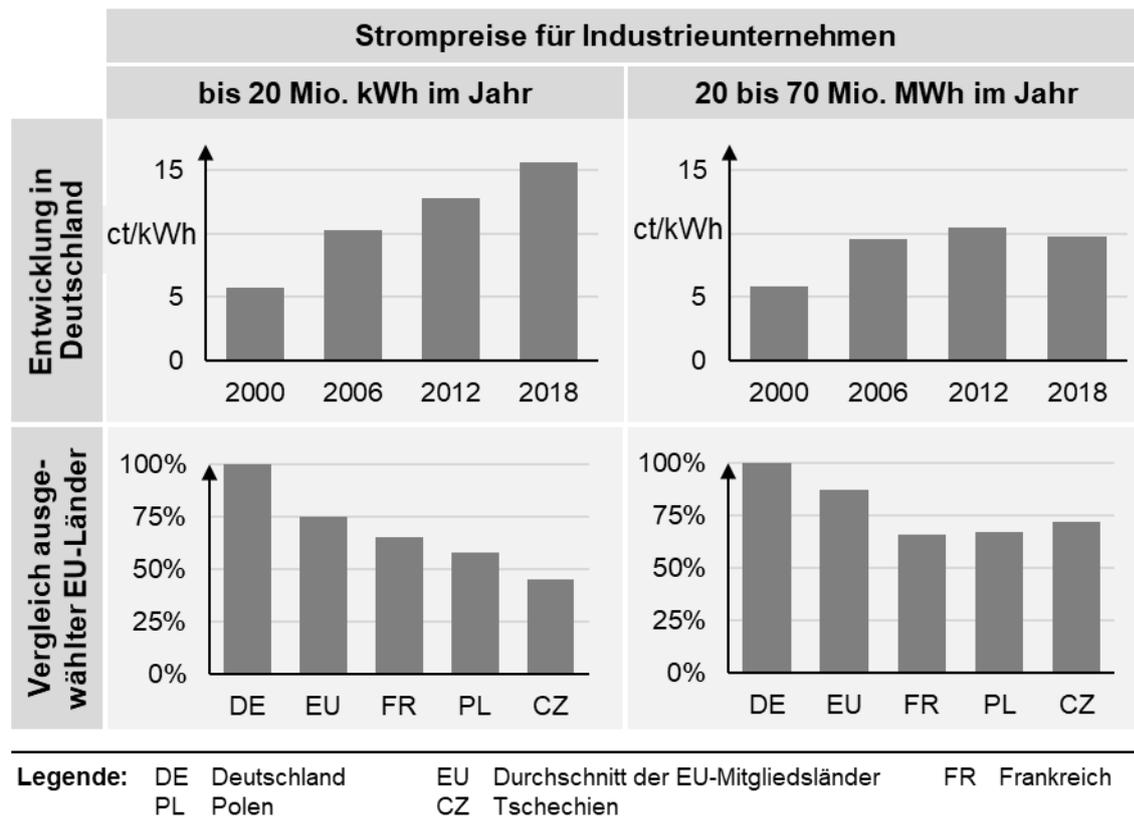


Abbildung 1: Entwicklung der Industriestrompreise in Deutschland sowie deren Höhe im internationalen Vergleich (Daten aus BDEW 2018, EUROSTAT 2018)

Ein zusätzlicher Effekt des Ausbaus der EE ist die zunehmende Schwankung der Strompreise an den Börsen, die auf die witterungsabhängige Erzeugungsleistung zurückzuführen ist (PUDLIK ET AL. 2015). So ist etwa die Schwankungsbreite der stündlichen Durchschnittspreise in Deutschland von 2010 bis 2016 um 30 Prozent gestiegen (BNETZA 2011, BNETZA 2017). In extremen Fällen können sogar negative Preise auftreten. Die Summe der jährlichen Stunden mit negativen Strompreisen in Deutschland hat sich von 2011 bis 2017 verzehnfacht (HANDELSBLATT 2017).

Um im internationalen Wettbewerb zu bestehen, sind Unternehmen in Deutschland bestrebt, die Ausgaben für Elektrizität trotz der hohen Strompreise möglichst gering zu halten (DIHK 2017). Die generelle Steigerung der Energieeffizienz und die damit einhergehende Verbrauchsreduzierung bieten dabei einen wichtigen Stellhebel (VBW 2012). Außerdem eröffnen sich für Unternehmen durch Flexibilisierung des Strombedarfs weitere Potenziale, die Stromkosten zu reduzieren. Auf diese Weise lassen sich die zunehmenden Strompreisschwankungen gezielt zur Kostensenkung nutzen. Gleichzeitig wird dadurch auch zur Stabilisierung der Stromnetze beigetragen, indem Abweichung zwischen der verfügbaren und benötigten Energiemenge zumindest teilweise durch die Lastanpassung der Industrieunternehmen ausgeglichen werden kann (BRÄUNINGER ET AL. 2012, BRÜNDLINGER ET AL. 2018).

Dazu müssen Unternehmen in der Lage sein, die Energienachfrage in Zeiten niedriger Strompreise zu erhöhen und umgekehrt während hoher Strompreise zu reduzieren. Diese Fähigkeit, „sich schnell und mit sehr geringem finanziellen Aufwand an Änderungen des Energiemarktes anzupassen“, wurde von REINHART ET AL. (2012A) als Energieflexibilität definiert. Die flexible Anpassung des Verbrauchs kann bei Industrieunternehmen direkt in der Produktion erfolgen, da dort der größte Teil der Energie verbraucht wird (FLEITER 2013, SAUER ET AL. 2019). Daneben besteht die Möglichkeit, zur Erhöhung der Flexibilität Strom zwischenspeichern. Die Preise für Stromspeicher sind in den letzten zehn Jahren um ca. 80 Prozent gefallen. Eine Fortführung dieses Trends wird insbesondere für Lithium-Ionen-Batterien prognostiziert (MCKINSEY & COMPANY 2017, THIELMANN ET AL. 2018). Aus diesem Grund werden Stromspeicher für den Einsatz an Produktionsstandorten zunehmend attraktiver. Dies belegt auch der extreme Anstieg der stationären Speicherkapazität in Deutschland zwischen 2015 und 2018 um den Faktor 15 (STOLLE ET AL. 2018, KÖHLER ET AL. 2018).

Eine weitere Möglichkeit zur Reduzierung der Stromkosten ist die Stromeigenerzeugung. Insbesondere bei den Investitionskosten für Kleinkraftwerke, basierend auf PV, war in den letzten Jahren eine starke Kostendegression zu beobachten (WIRTH 2018). Zudem entfällt im Rahmen der aktuellen Gesetzgebung bei der Eigenerzeugung von Strom ein Teil der Umlagen und sonstigen Abgaben (BNETZA 2016). Industriebetriebe können daher durch den Verbrauch von selbsterzeugtem Strom bis zu 25 Prozent der Stromkosten einsparen (ERNST & YOUNG 2016). Abbildung 2 macht deutlich, dass Unternehmen unabhängig vom Energieverbrauch durch den Bau und Betrieb eigener Blockheizkraftwerke (BHKW) sowie PV- und Windkraftanlagen (WKA) die Strompreise signifikant reduzieren können. Dieses

Potenzial spiegelt sich auch in der Entwicklung der Stromeigenversorgung in der Industrie wider. So hat sich die Anzahl der Unternehmen mit Eigenversorgungsanlagen über die Jahre 2005 bis 2015 versechsfacht (DESTATIS 2016). Aufgrund der fluktuierenden Erzeugungsleistung von WKA und PV kann durch die Anpassung des Energiebedarfs an die Verfügbarkeit der wirtschaftlich attraktive Stromeigenverbrauch gesteigert werden (BEIER 2017). Energieflexibilität ist daher nicht nur in Anbetracht schwankender Strompreise, sondern auch im Hinblick auf die Erhöhung der Stromeigenversorgung von großer Relevanz.

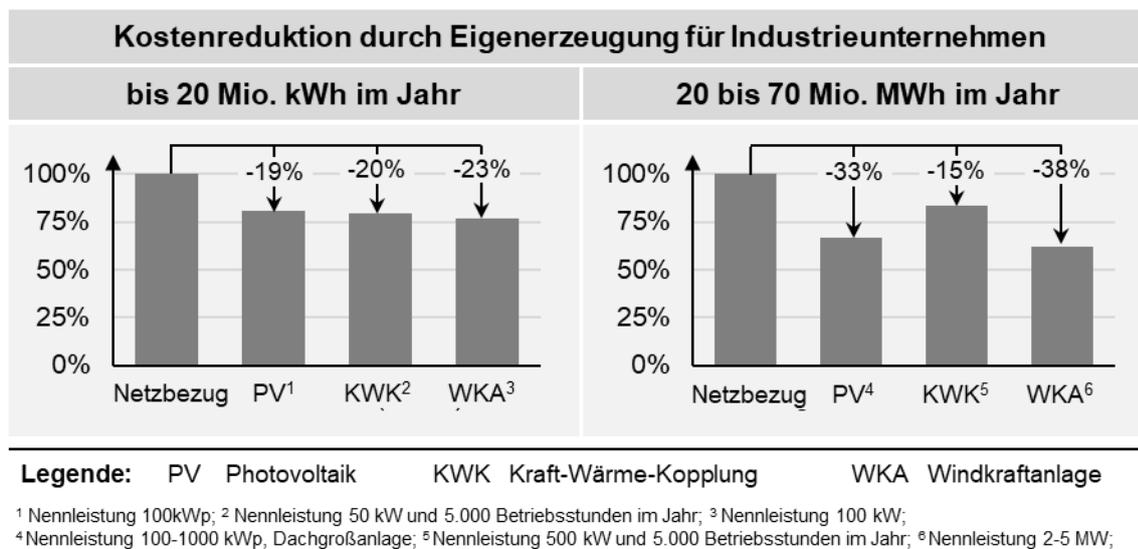


Abbildung 2: Potenzial zur Senkung der Stromkosten durch Eigenerzeugung (Daten aus (KOST ET AL. 2018, BMWI 2014, WIRTH 2018, LÜERS ET AL. 2015, BDEW 2018))

1.2 Motivation der Arbeit

Die beschriebene Ausgangssituation stellt eine Herausforderung für deutsche Unternehmen dar, die auf den Weltmärkten einem andauernden und schnelllebigen Wettbewerb ausgesetzt sind (ABELE & REINHART 2011; REINHART 2017). Im Hinblick auf Energie genießen ausländische Betriebe Wettbewerbsvorteile, da sie in ihren Heimatländern tendenziell deutlich geringere Stromkosten tragen müssen. Zudem gilt der Energiepreis als ein wichtiger Faktor bei der Standortwahl (BANKHOFER 2001, VDMA 2013, DIHK 2017). Industriebetriebe in Deutschland sind also gezwungen, diese Nachteile durch andere Faktoren auszugleichen oder aber die Stromkosten durch Flexibilisierung des Strombedarfs sowie die Erschließung kostengünstiger Energiequellen und aktiven Stromhandel zu reduzieren.

Rund die Hälfte des Stromverbrauchs produzierender Unternehmen ist direkt auf die Produktionsprozesse zurückzuführen (FLEITER 2013). Bei der Beeinflussung des Energiebedarfs in der Produktion kommt der Produktionsplanung und -steuerung (PPS) demnach eine zentrale Bedeutung zu (SCHULTZ ET AL. 2016B). Sie terminiert den Herstellungszeitpunkt der zu bearbeitenden Aufträge und die dabei zum Einsatz kommenden Ressourcen (SCHUH 2011). Während die Produktionsplanung das lang- und mittelfristige Produktionsprogramm festlegt, ist es die Aufgabe der Produktionssteuerung, die Umsetzung der Vorgaben aus der Planung kurzfristig sicherzustellen, unvorhergesehenen Ereignissen wie Maschinenausfällen zum Trotz (LÖDDING 2013). Da der Strombedarf von dem im jeweiligen Zeitschritt zu bearbeitenden Auftrag abhängt, bestimmt die PPS maßgeblich den Energieverbrauch (REINHART ET AL. 2012B, ABELE ET AL. 2012, GRABL 2015). Durch energieorientierte Anpassungen auf Ebene der PPS kann der Energiebedarf eines Produktionsstandorts zielgerichtet beeinflusst werden.

Die Kernaufgabe der PPS ist die Sicherstellung der logistischen Zielgrößen Durchlaufzeit, Termintreue, Bestände und Auslastung (WIENDAHL 2010). Energieorientierte Anpassungen können diese Zielgrößen negativ beeinflussen und ziehen somit unter Umständen zusätzliche monetäre Aufwände nach sich (BEIER 2017, BERTSCH ET AL. 2017). Energieorientierte Eingriffe sollten daher hinsichtlich des gesamtheitlichen wirtschaftlichen Nutzens geprüft werden. In dieser Abwägung ist Energie jedoch lediglich ein flexibler Kostenfaktor, den es zusammen mit allen anderen relevanten Kosteneinflüssen zu minimieren gilt. Um dies zu ermöglichen, ist die Überführung der einzelnen logistischen Zielgrößen in Kosten erforderlich (SCHULTZ 2018). Auf diese Weise lassen sich fundierte Entscheidungen fällen, die den gesamtwirtschaftlichen Nutzen energieorientierter Anpassungen sicherstellen.

Neben der gesamtheitlichen Betrachtung der Zielgrößen stellen die zahlreichen stochastischen Einflüsse im Produktionsumfeld, die beispielsweise in Form von kurzfristigen Änderungen und ungeplanten Störungen auftreten, eine weitere Herausforderung dar (BRACKEL 2009, NIEHUES 2017). Aus diesem Grund ist die Produktionssteuerung gefordert, die Plan-Vorgaben dennoch zu erreichen (SCHUH & FUB 2015). Im Zuge der Integration von erneuerbaren Energien und der Teilnahme am Energiemarkt kommen weitere Unsicherheitsfaktoren hinzu (BEIER ET AL. 2015). So können etwa die Strompreise an den Börsen kurzfristig stark schwanken. Ebenso kann die Erzeugungsleistung aufgrund unvorhergesehener Wetteränderungen von der vorhergesagten Menge abweichen (BERTSCH ET AL. 2017). Auch ziehen Planabweichungen in der Produktion direkte Auswirkungen auf den Energieverbrauch nach sich (SCHULTZ 2018). Umgekehrt kann die

Änderung der erzeugten Strommenge ebenfalls Anpassungen im Produktionsablauf erfordern, um die Stromkosten effektiv zu reduzieren (BIEL & GLOCK 2016). Ein Stromspeicher kann in diesem komplexen Spannungsfeld als zusätzlicher Energieflexibilitätsanbieter genutzt werden. Zur optimalen Steuerung muss jedoch eine Integration der Betriebsstrategie in die Steuerung des Energieverbrauchs und des Stromhandels erfolgen (BEIER ET AL. 2016).

Insgesamt wird deutlich, dass an einem energieflexiblen Produktionsstandort mit Eigenerzeugung und Stromspeicher bei gleichzeitiger Teilnahme am Strommarkt eine starke Wechselwirkung zwischen den genannten Elementen und der Produktion vorliegt. Gleichzeitig ist die Planbarkeit aufgrund von stochastischen Ereignissen beschränkt, wodurch kurzfristige Anpassungen erforderlich sind.

1.3 Zielsetzung und Aufbau der Arbeit

Als Schlussfolgerung aus der dargelegten Ausgangssituation und Motivation liegt das zentrale Ziel der vorliegenden Arbeit in der Entwicklung eines Systems zur energieorientierten und kostenbasierten Produktionssteuerung. Zur Optimierung der Stromkosten ist dabei eine integrierte Steuerung von Eigenerzeugung, Stromspeicher und kurzfristigem Stromhandel erforderlich. Die Produktionssteuerung muss gleichzeitig in der Lage sein, die Energiekosten mit den klassischen Produktionszielen abzuwägen. Da eine Vergleichbarkeit nicht direkt möglich ist, sind dazu die logistischen Zielgrößen in ein gesamtkostenbasiertes Zielsystem zu überführen.

Um dieses Ziel zu erreichen, werden die in Abbildung 3 aufgeführten vier Forschungsfragen abgeleitet, die im Zuge der vorliegenden Arbeit beantwortet werden. Durch den dadurch erzielten Erkenntnisgewinn sollen Unternehmen mit Hilfe der integrierten Steuerung sowohl des Strombezugs als auch der Produktion in die Lage versetzt werden, gleichermaßen die Strom- und die weiteren Produktionskosten zu reduzieren.

Nach ULRICH & HILL (1976) ist die vorliegende Arbeit den Handlungswissenschaften und damit den Realwissenschaften zuzuordnen. Dabei wird aus der Analyse menschlicher Handlungsalternativen ein System zur Bewältigung der oben genannten Aufgabenstellung abgeleitet. Dies erfordert eine Verknüpfung der wissenschaftlichen Themengebiete Produktionstechnik, Energiewirtschaft und aufgrund des im weiteren Verlauf ausgewählten Steuerungsansatzes in Teilen auch

künstliche Intelligenz. Die verwendete Vorgehensweise orientiert sich an der *Design Research Methodology* (DRM) nach BLESSING & CHAKRABARTI (2009), weil diese Methodik am besten von allen bekannten die Anforderungen der Handlungswissenschaften erfüllt. Die Struktur der insgesamt neun Kapitel ist an den Typ 3 dieser Methodik mit dessen vier wesentliche Schritten angelehnt (vgl. Abbildung 4), welche nachfolgend in den Kontext des Aufbaus dieser Arbeit gesetzt werden.

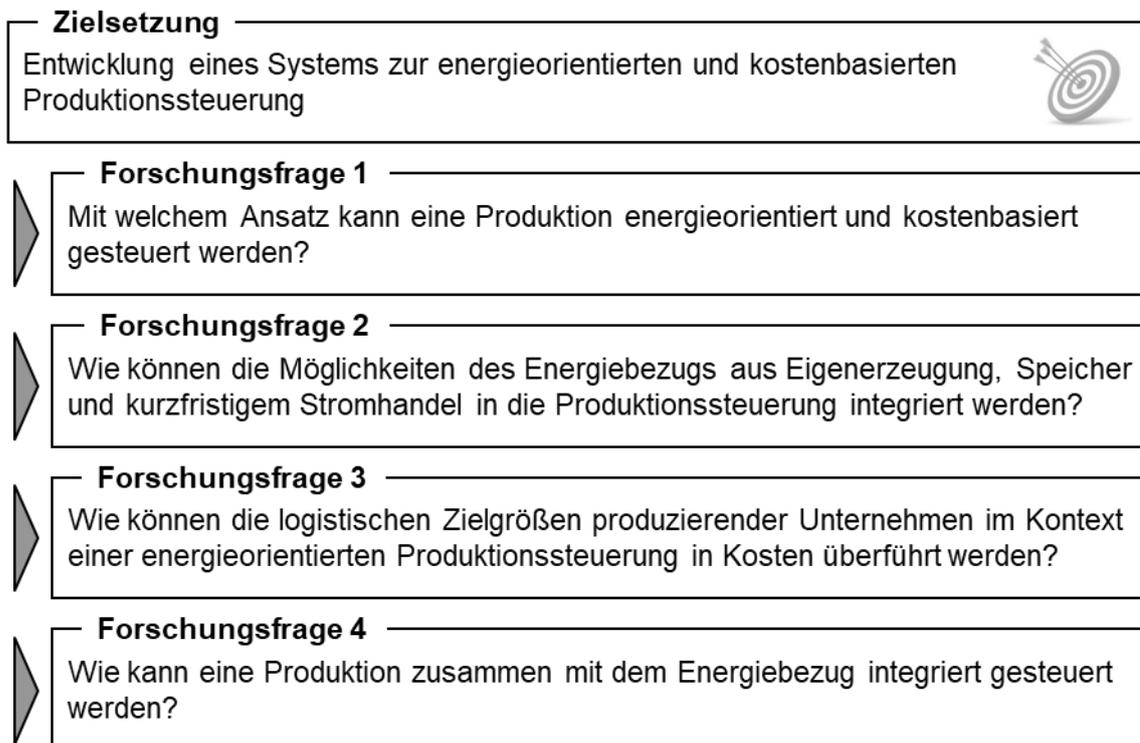


Abbildung 3: Zielsetzung und Forschungsfragen der vorliegenden Arbeit

Im ersten Schritt der Methodik, der *Research Clarification*, erfolgt die Ausarbeitung der Forschungsziele anhand von Belegen und Indikatoren. Diese wurden in der vorausgegangenen Beschreibung zur Ausgangssituation, Motivation und Zielsetzung in Kapitel 1 dargelegt. Um die Faktoren zu identifizieren, die es zur Erreichung dieser Ziele zu berücksichtigen gilt, wird im Zuge des zweiten Schrittes die *Descriptive Study I* durchgeführt, welche ebenfalls auf einer Recherche basiert. Zu diesem Zweck werden in Kapitel 2 die relevanten Grundlagen dargelegt. Dabei wird zunächst auf den Stromverbrauch und -bezug in Industrieunternehmen eingegangen, bevor das Basiswissen zur PPS und die wesentlichen Ansätze zur Produktionssteuerung aufgezeigt werden. Auf dieser Grundlage wird in Kapitel 3 der für die Zielsetzung relevante Stand der Wissenschaft zusammengefasst. Dabei stehen einerseits die Vorarbeiten zur energieorientierten PPS im Vordergrund, wobei ins-

besondere die Produktionssteuerung beleuchtet wird. Andererseits werden die Ansätze im Bereich der kostenbasierten Produktionssteuerung analysiert. Das Ergebnis der *Descriptive Study I* ist ein tiefgreifendes Verständnis, wodurch am Ende von Kapitel 3 die zentralen Handlungsempfehlungen abgeleitet werden können.

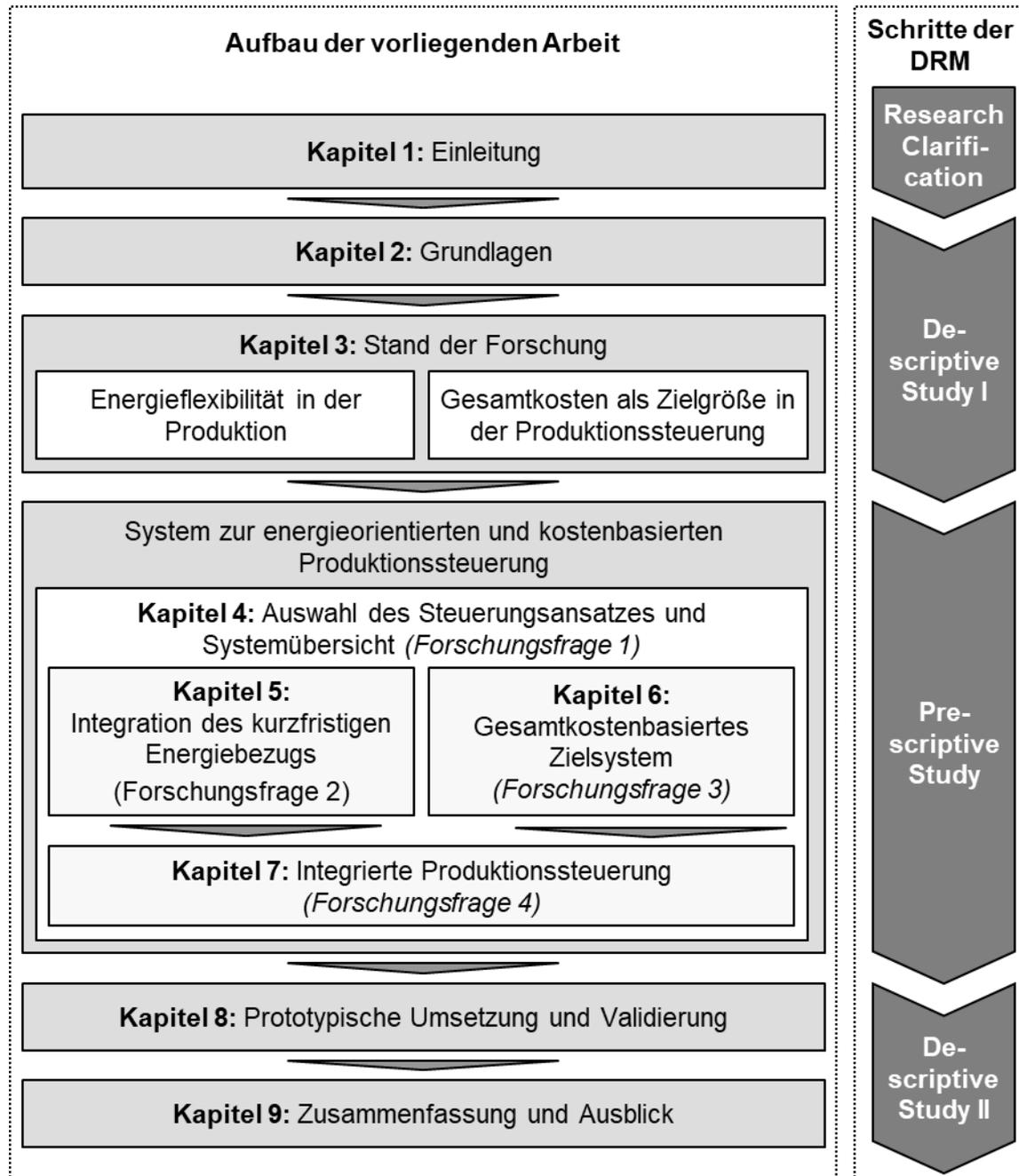


Abbildung 4: Aufbau der vorliegenden Arbeit und Einordnung in die Schritte der Design Research Methodology (DRM)

Im Zuge der *Prescriptive Study*, die den dritten Schritt der DRM darstellt, erfolgt die systematische Entwicklung des Systems zur Erreichung der Forschungsziele.

Aus diesem Grund wird in Kapitel 4 ein geeignetes Steuerungsverfahren ausgewählt und anschließend ein Überblick über das Gesamtsystem gegeben. Die Integration des Energiebezugs in Form von Eigenerzeugung, Speichernutzung und dem kurzfristigen Stromhandel steht im Fokus von Kapitel 5. Der Kernaspekt in Kapitel 6 ist die Ableitung des gesamt-kostenbasierten Zielsystems. Im Zuge dessen werden die für die Produktionssteuerung relevanten Elemente der logistischen Zielgrößen zu einer gesamtheitlichen Kostenbewertung zusammengefasst. Das aus Kapitel 5 resultierende Steuerungsschema wird in Kapitel 7 zusammen mit dem Produktionssystem in eine integrierte Steuerung überführt, wobei das in Kapitel 6 entwickelte Zielsystem ebenfalls als Input dient.

Die Phase der *Descriptive Study II* stellt den letzten Schritt der DRM dar und dient dazu, die Wirkung des entwickelten Systems zu analysieren. Aus diesem Grund erfolgt die technische Umsetzung des entwickelten Gesamtsystems und dessen Evaluierung anhand mehrerer Benchmark-Szenarien in Kapitel 8. Abschließend wird in Kapitel 9 die vorliegende Arbeit zusammengefasst. Die erzielten Ergebnisse werden eingeordnet, ebenso erfolgt ein Ausblick auf den weiteren Forschungsbedarf.

2 Grundlagen

2.1 Übersicht

Im vorliegenden Kapitel werden die zur Erreichung der dargelegten Zielstellung erforderlichen Grundlagen zusammengefasst. Als Ausgangspunkt wird zunächst die Bedeutung von Strom, dessen Verbrauch, Beschaffung und Speicherung in produzierenden Unternehmen beleuchtet (Abschnitt 2.2). Diese Grundlagen bilden eine wichtige Basis für die Betrachtung der im Kontext der Produktion möglichen Energiebezugsoptionen in den späteren Teilen der Arbeit. Darüber hinaus erfolgt in Abschnitt 2.3 eine Betrachtung der Produktionsplanung und -steuerung sowie der existierenden Lösungsansätze für die Produktionssteuerung (Abschnitt 2.4). Abschließend wird mit Reinforcement Learning einer der Lösungsansätze detaillierter vorgestellt (Abschnitt 2.5).

2.2 Elektrischer Strom in produzierenden Unternehmen

2.2.1 Lastgang und Energieflexibilität

Elektrische Energie wird in produzierenden Unternehmen für eine Vielzahl an Anwendungen von Motoren über Wärmeerzeugung bis hin zur Beleuchtung eingesetzt. In Abhängigkeit der Randbedingungen und des Zustands der einzelnen Fertigungsmaschinen schwankt der Energiebedarf einer Fabrik über den Tag. Dieser zeitliche Verlauf des Stromverbrauchs wird im Allgemeinen als *Lastgang* bezeichnet (HESSEL 2008). Die Fähigkeit, diesen schnell und ohne großen finanziellen Aufwand anzupassen, wird im Kontext der Produktion allgemein unter *Energieflexibilität* zusammengefasst (vgl. Abschnitt 1.2). Das *Lastmanagement* kann dabei in die kurzfristige Umsetzung von derartigen Maßnahmen eingeordnet werden, um beispielsweise den Energieverbrauch innerhalb von abrechnungsrelevanten 15-Minutenintervallen durch eine temporäre Abschaltung von Verbrauchern zu beschränken (TALUKDAR & GELLINGS 1987). Hinter dem besonders im US-amerikanischen Raum verbreiteten Begriff *Demand Side Management* verbirgt sich eine mittelfristige bis strategische Planung und Nutzung von Energieflexibilität (GELLINGS & CHAMBERLIN 1993).

Um Potenziale für Energieflexibilität heben zu können, hat GRAßL (2015) verschiedene Maßnahmen definiert, die auf unterschiedlichen operativen Ebenen der

Fabrik ausgeführt werden können (KÖRNER ET AL. 2019). Da der Strombedarf in der Produktion vom aktuellen Zustand der Produktionsressourcen wie Maschinen und Anlagen abhängt, wird durch die Steuerung der zu bearbeitenden Aufträge und Bearbeitungsparameter sowie Standby- und Rüstzeiten maßgeblich der Energieverbrauch festgelegt (WEINERT 2010, BEIER ET AL. 2015, GRAßL 2015). Die Modellierung des Energiebedarfs, bezogen auf einen Auftrag, wird im Kontext der diskreten Fertigung in vielen Fällen durch gemittelte Durchschnittswerte, den sogenannten „Energy Blocks“ abgebildet (WEINERT 2010, GRAßL 2015).

2.2.2 Stromeigenerzeugung

Neben der Strombeschaffung aus externen Quellen können Unternehmen Elektrizität auch selbst erzeugen. Da dabei in der Regel lediglich 40 Prozent der EEG-Umlage zu entrichten sind, können die Stromkosten auf diese Weise deutlich gesenkt werden (BNETZA 2016). Im Hinblick auf die Steuerbarkeit wird nachfolgend zwischen *fluktuierenden* und *regelbaren Erzeugungsanlagen* differenziert (GÜNTHER 2015).

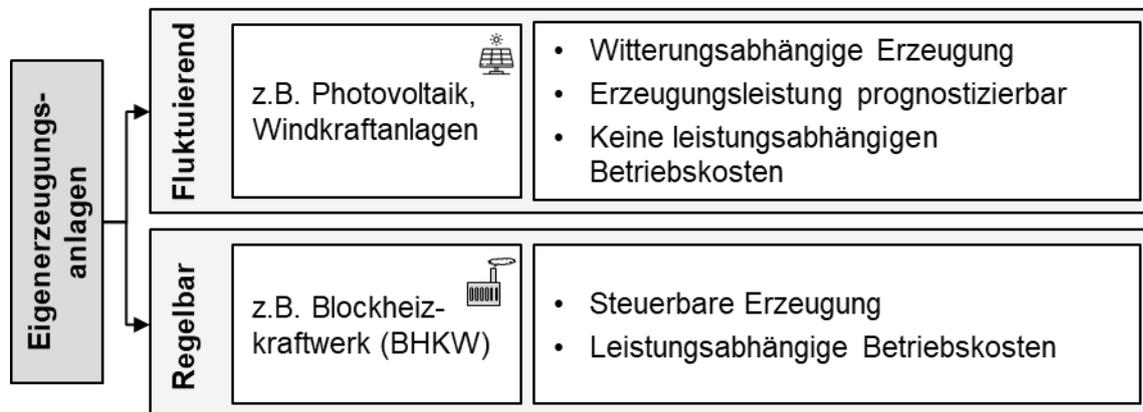


Abbildung 5: Einordnung von potenziellen Stromeigenerzeugungsanlagen

Zu den *fluktuierenden Erzeugungsanlagen* zählen Windkraft- und PV-Anlagen, deren realisierbare Erzeugungsleistung maßgeblich vom Wetter abhängt und dadurch starken Schwankungen unterworfen ist (KELLER 2018). Die Größe der Anlagen wird anhand der Nennleistung bemessen, die den maximalen Stromertrag bei optimalen Wetter- und Standortbedingungen angibt. Eine gezielte Steuerung der Stromproduktion ist lediglich möglich, indem durch die sogenannte Abregelung die Leistung reduziert wird (JARASS ET AL. 2009). Da fluktuierende Erzeuger keine nennenswerten leistungsabhängigen Betriebskosten verursachen, ist dieser Fall jedoch aus wirtschaftlicher Sicht zu vermeiden (HUNEKE ET AL. 2016). Die

Erzeugungsleistung kann durch meteorologische Prognosemodelle vorhergesagt werden. Für einen Zeitraum von 24 Stunden ist die Vorhersage mit einem mittleren Fehler von unter fünf Prozent möglich (REICH & REPPICH 2018). Bei der Stromerzeugung sind für Unternehmen vor allem PV-Anlagen relevant. WKA spielen aufgrund aufwendiger Genehmigungsverfahren eine untergeordnete Rolle (BOLAY ET AL. 2014).

Die Leistung *regelbarer Erzeugungsanlagen* kann innerhalb der technischen Grenzen frei gesteuert werden. Dabei kommen sowohl fossile Energieträger wie Kohle oder Gas als auch regeneratives Biogas zum Einsatz (GÜNTHER 2015). Im Kontext der Eigenerzeugung sind Blockheizkraftwerke (BHKW) weit verbreitet, die sowohl Strom als auch Wärme auf Basis einer Kraft-Wärme-Kopplung (KWK) erzeugen (BOLAY ET AL. 2014). Auf diese Weise kann ein Wirkungsgrad von über 90 Prozent erreicht werden (BMW I 2014). Abhängig vom Anlagentyp sinkt der Wirkungsgrad im Teillastbetrieb und ist ab einem bestimmten Schwellenwert nicht mehr sinnvoll (SCHAUMANN & SCHMITZ 2010). Aus diesem Grund können BHKW nur zu einem bestimmten minimalen Teillastbereich betrieben werden (PILGRAM & KARGER 2014, HOLDACK 2015).

2.2.3 Energiespeicher

Um Strom temporär zu speichern, gibt es eine Vielzahl von Speichertechnologien, die auf unterschiedlichen physikalischen Prinzipien basieren. Die mögliche Speicherdauer variiert dabei von unter einer Minute, beispielsweise bei Kondensatoren, bis zu Monaten, etwa im Hinblick auf Porenspeicher (STERNER & STADLER 2017). Für den stationären Einsatz bei Industrieunternehmen eignen sich Batterien besonders, da sie Strom mit geringen Verlusten über mehrere Stunden speichern können, eine hohe Ein- und Ausspeicherleistung aufweisen und gleichzeitig auch aus wirtschaftlicher Sicht attraktiv sind (THIELMANN ET AL. 2015, BOLAY ET AL. 2017, KÖHLER ET AL. 2018). Aus diesen Gründen wird im Rahmen dieser Arbeit ausschließlich auf Batteriespeicher eingegangen.

Batterien bestehen aus elektrochemischen, wiederaufladbaren Zellen, die zu Modulen zusammengeschaltet werden (ANDRESEN ET AL. 2017). Im stationären Einsatz werden dabei vor allem Systeme auf Blei-Säure-Basis und zunehmend auch Lithium-Ionen-Batterien eingesetzt (THIELMANN ET AL. 2015). Auf technischer Ebene verfügt eine Batterie über eine definierte Kapazität und eine maximale Lade- beziehungsweise Entladeleistung, die in Form der C-Rate angegeben wird.

Das Aus- und Einspeichern von Energie ist darüber hinaus jeweils mit einem bestimmten Wirkungsgrad verbunden. Einzelne Zyklen können anhand der Entladetiefe (engl. Depth of Discharge, DoD) und dem Ladezustand (engl. State of Energie; SoC) charakterisiert werden. Ein Nachteil von Batteriespeichern ist die begrenzte Lebensdauer, die maßgeblich von der Betriebsstrategie beeinflusst wird (ZAPF 2017). Je nach Typ reagieren die Batterien mehr oder weniger empfindlich auf wechselnde Ladezyklen oder Tiefenentladung (ANDRESEN ET AL. 2017).

Die Degradation von Batterien, insbesondere bei Lithium-Ionen-Speichern, wird durch die kalendarische und die zyklische Alterung verursacht (VETTER ET AL. 2005, LARESGOITTI ET AL. 2015). Die kalendarische Alterung beschreibt die belastungsunabhängige, zeitliche Degradation einer Batterie, während die zyklische Alterung den inneren mechanischen Belastungen Rechnung trägt, die beim Laden/Entladen auftreten (VETTER ET AL. 2005, KEIL ET AL. 2016). Zur Modellierung dieser Mechanismen können unter anderem theoretische Ansätze angewendet werden, die sehr feingranulare Daten erfordern (ZHANG & WHITE 2008, SAFARI ET AL. 2009). Empirische Modelle sind dagegen deutlich weniger komplex. Voraussetzung ist hier eine ausreichende Datengrundlage (KOLLER ET AL. 2013, ORTEGA-VAZQUEZ 2014). Ein Ansatz, die Vorteile beider Modelltypen zu vereinen, sind sogenannte semi-empirische Modelle, wie sie etwa von XU ET AL. (2018) oder MARTINEZ-LASERNA ET AL. (2018) eingesetzt werden.

2.2.4 Strombeschaffung

Für Unternehmen in Deutschland eröffnen sich mit der *Stromversorgung durch ein Energieversorgungsunternehmen* (EVU) oder dem *direkten Stromhandel* zwei wesentliche Beschaffungsoptionen, um ihren Strombedarf aus externen Quellen zu decken (vgl. Abbildung 6). Wie ein Unternehmen die benötigte Energiemenge beschafft, hängt stark von der jeweiligen Risikoeinstellung ab (STROHBÜCKER 2011).

Im Zuge der *Stromversorgung durch ein EVU* besteht bei der *Vollstrombeschaffung* ein Liefervertrag zwischen Stromverbraucher und -lieferanten, meist ein EVU (KONSTANTIN 2009). Der Lieferant übernimmt dabei vollständig das Volumen- und Preisrisiko für den Kunden und stellt einen Komplettstrompreis in Rechnung (BERG & BORCHERT 2014). Auszuhandelnde Vertragsbestandteile sind dabei unter anderem der Arbeitspreis pro kWh, sowie der Leistungspreis, der sich an der maximal bezogenen Leistung innerhalb einer Periode bemisst (SCHUMACHER & WÜRFEL 2015). Alternativen sind unter anderem die *Indexierte Beschaffung* und

Tranchenmodelle. Der Preis kann dabei während der Vertragslaufzeit an die Entwicklung an der Strombörse angepasst werden (BERG & BORCHERT 2014).

Beim *direkten Stromhandel* bezieht das Unternehmen den elektrischen Strom direkt von der Strombörse oder bilateral von einem Stromerzeuger. Letzteres wird als „Over the Counter“(OTC)-Handel bezeichnet (SCHUMACHER & WÜRFEL 2015). Das Unternehmen trägt bei dieser Beschaffungsform das alleinige Preisrisiko und muss zusätzlich sicherstellen, den gekauften Strom tatsächlich abzunehmen (BERG & BORCHERT 2014). Tritt eine Abweichung zwischen der Bezugsmenge und der tatsächlich verbrauchten Energie auf, werden die Kosten, die dem Netzbetreiber zum Ausgleich dieser Abweichung entstehen, auf das Verursacherunternehmen umgelegt (STROHBÜCKER 2011, SCHUMACHER & WÜRFEL 2015). Das zulässige Toleranzband liegt dabei in der Regel zwischen \pm fünf Prozent bis \pm 20 Prozent (HESSEL 2008, WEIB 2010, SCHUMACHER & WÜRFEL 2015). Neben dem Strompreis werden beim direkten Stromhandel weitere Steuern, Abgaben und Umlagen fällig (BDEW 2018).

Für den direkten Handel an der Börse stehen im Wesentlichen drei Energiemärkte zur Auswahl, die sich durch den zeitlichen Vorlauf unterscheiden (vgl. Abbildung 6). Der *Terminmarkt* und der direkte *OTC-Handel* zwischen Stromerzeuger und Verbraucher eignen sich für die mittel- bis langfristige Strombeschaffung mit einem Vorlauf von bis zu sechs Jahren. Am *Day-Ahead-Markt* können Gebote bis 12:00 Uhr für den darauffolgenden Tag abgegeben werden. Einen untertägigen Stromhandel ermöglicht hingegen der *Intraday-Markt*. Seit 2017 sind hier der An- und Verkauf von Strom in Viertelstundenintervallen mit nur fünf Minuten Vorlaufzeit möglich. Dabei fallen pauschale und anteilige Ordergebühren an. Die kleinste handelbare Einheit sind 0,1 MWh. Unternehmen sind mit Hilfe von Aggregatoren, die als Dienstleister die Flexibilität von mehreren Verbrauchern bündeln, in der Lage, auch kleinere Strommengen zu erwerben. Die am Intraday-Markt gehandelten Volumina nehmen infolge des Ausbaus der EE deutlich zu und haben sich von 2010 bis 2018 mehr als verfünffacht. (BNETZA 2011, BERTSCH ET AL. 2017, BNETZA & BKARTA 2015, EPEX SPOT 2018, BNETZA 2019)

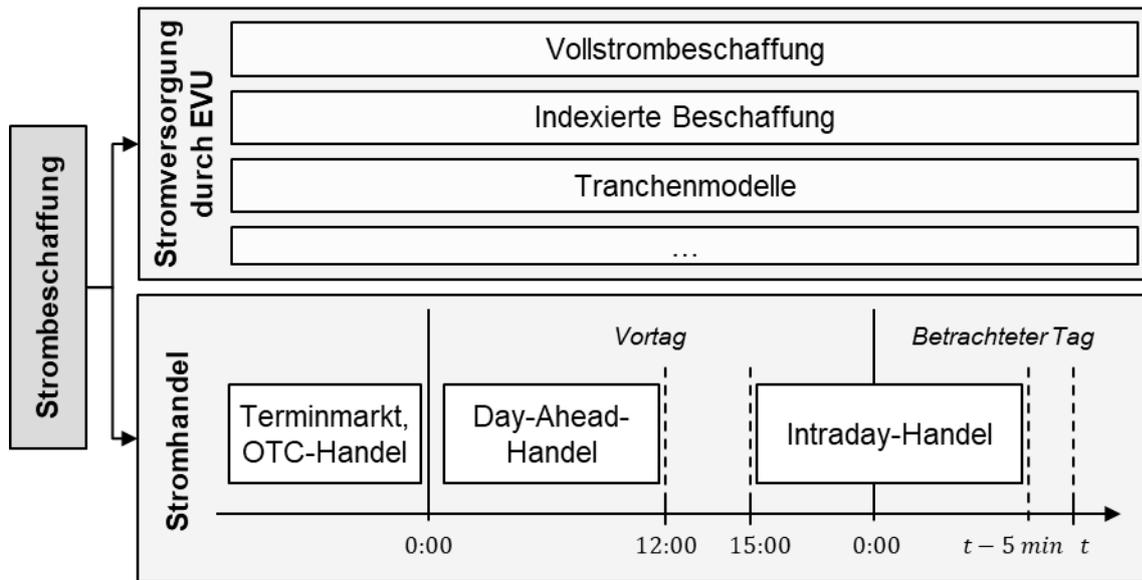


Abbildung 6: Möglichkeiten der Strombeschaffung für Industrieunternehmen

Neben dem Handel an der Strombörse bieten Netzbetreiber Industrieunternehmen zusätzliche finanzielle Anreize, um durch Systemdienstleistungen zur Netzstabilität beizutragen. Dabei kann zwischen positiver und negativer Regelleistung unterschieden werden, je nachdem, ob die vom Netz bezogene Leistung im Zuge des Abrufs erhöht oder verringert wird. Wollen Unternehmen am Regelleistungsmarkt teilnehmen, müssen sie zunächst eine Präqualifizierung durchlaufen, damit die Verfügbarkeit der Regelleistung im Falle eines Abrufs gesichert ist. Je nach Zeithorizont wird dabei in *Primärregelleistung*, *Sekundärregelleistung* und *Minutenreserve* unterschieden. (ELLER 2015, BERTSCH ET AL. 2017, REGELLEISTUNG.NET 2018)

2.3 Produktionsplanung und -steuerung

2.3.1 Aufgaben und Abgrenzung

Die zentrale Aufgabe eines produzierenden Unternehmens ist die Herstellung von Gütern in der vom Markt benötigten Menge, Qualität und Zeit (DANGELMAIER 2009). Zu diesem Zweck werden die elementaren Ressourcen Mensch, Maschine, Material und Information zur Abarbeitung von Aufträgen eingesetzt. Bei zunehmender Auslastung der Produktion konkurrieren dabei die Aufträge um diese begrenzten Ressourcen (LÖDDING 2008). Die Produktionsplanung und -steuerung (PPS) dient dazu, die sich daraus ergebende Problemstellung zu bewältigen

(NYHUIS & WIENDAHL 2009). Die Kernaufgabe der PPS ist folglich die „Koordination der konkurrierenden Aufträge unter Betrachtung des untergeordneten produktionswirtschaftlichen Zielsystems“ (VAHRENKAMP & SIEPERMANN 2008). In Anbetracht der zunehmenden Dynamik im Produktionsumfeld und der wachsenden Produktvielfalt spielt die PPS eine zentrale Rolle in der Auftragsabwicklung produzierender Unternehmen (SCHUH & STICH 2012).

Die Produktionsplanung umfasst nach VDI 2815 das „(...) systematische Suchen und Festlegen von Zielen für Fertigung [und das] Vorbereiten von Fertigungsaufgaben, deren Durchführung zum Erreichen der Ziele erforderlich ist“ (VDI-RICHTLINIE 2815). Demgegenüber weist die Produktionssteuerung als das „*Veranlassen, Überwachen und Sichern der Durchführung von Fertigungsaufgaben hinsichtlich Bedarf (Menge, Termin), Qualität, Kosten und Arbeitsbedingungen*“ (VDI-RICHTLINIE 2815) einen deutlich kürzeren Betrachtungshorizont auf. Während die Produktionsplanung demnach den Produktionsplan unter Berücksichtigung gegebener Randbedingungen erstellt, ist es Aufgabe der Produktionssteuerung, diesen trotz ungeplanter Abweichungen, wie zum Beispiel Maschinenausfälle, in die Tat umzusetzen (WIENDAHL 2010).

Einen Überblick über die zahlreichen Aufgaben der PPS gibt das Aachener PPS-Modell. Wie in Abbildung 7 dargestellt, werden dabei die Funktionen der PPS in *Kern- und Querschnittsaufgaben* unterteilt. Die Kernaufgaben umfassen die *Produktionsprogramm- und Produktionsbedarfsplanung* sowie die *Planung und Steuerung für Eigenfertigung und Fremdbezug*. Der Fremdbezug von externen Lieferanten spielt im Kontext der vorliegenden Arbeit eine untergeordnete Rolle und wird daher im weiteren Verlauf nicht näher beleuchtet. Die Querschnittsaufgaben der PPS sind das Auftrags- und Bestandsmanagement, das Controlling und die Datenverwaltung. (SCHUH & STICH 2012)

Im Zuge der *Programmplanung* wird auf Basis des Absatzplans die Menge der zu produzierenden Produkte festgelegt. Die *Produktionsbedarfsplanung* gleicht diesen im Rahmen der *Materialdisposition* mit dem Lagerbestand ab und leitet daraus den Bedarf an Produktionsgütern in Form der Beschaffungsartzuordnung her. Anschließend beginnt die *Produktionsplanung* mit der Arbeitsvorbereitung, wobei die entsprechenden Stücklisten und Arbeitspläne erstellt werden. Die *Auftragsterminierung* legt die Fertigstellungstermine der einzelnen Aufträge fest und berücksichtigt die erforderlichen Bearbeitungszeiten. Falls mehrere Aufträge zusammengefasst werden können, ist es die Aufgabe der *Losgrößenrechnung*, die optimalen Losgrößen festzulegen. Die folgende *Kapazitätsplanung* gleicht die erforderlichen

und verfügbaren Kapazitäten der Fertigungsressourcen ab und leitet gegebenenfalls eine Anpassung ein, etwa durch Überstunden.

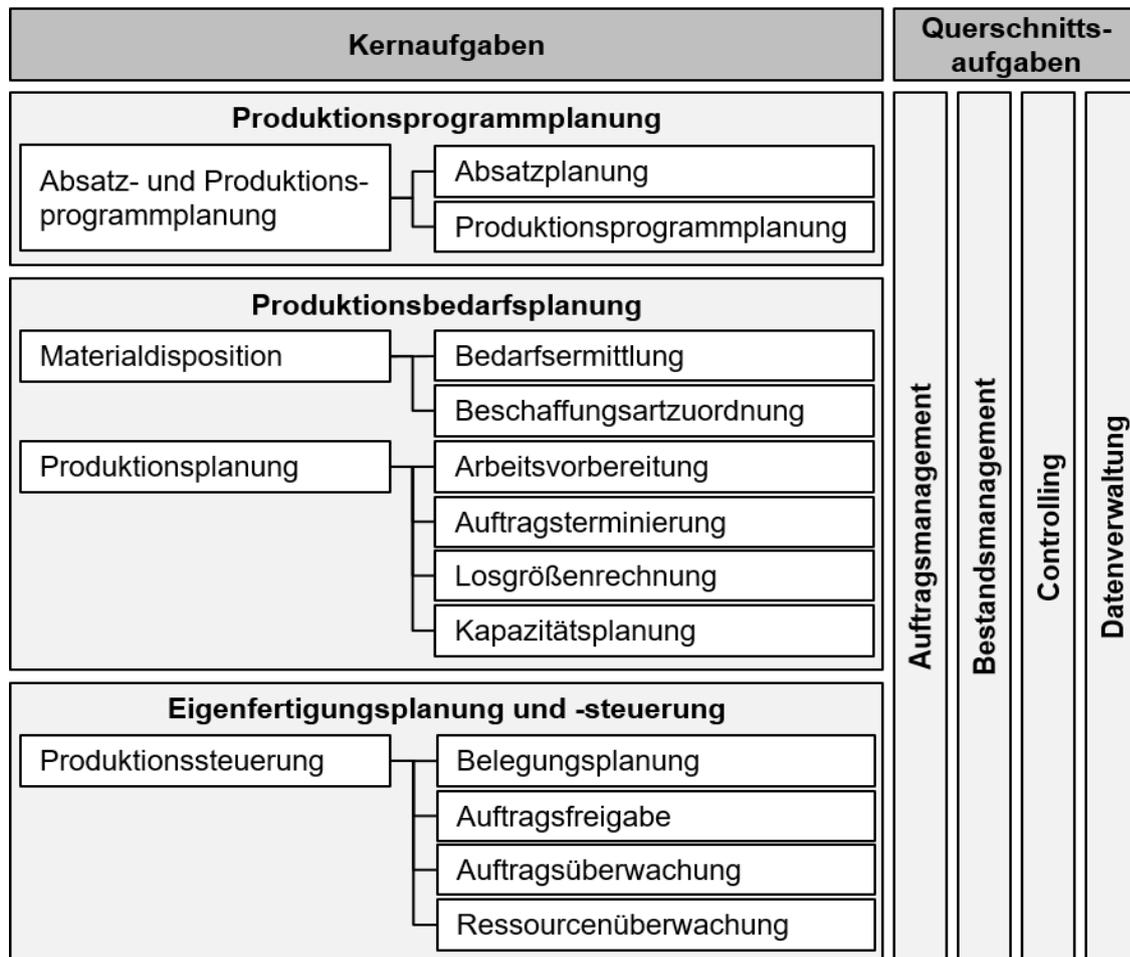


Abbildung 7: Das Aachener PPS-Modell (SCHUH & STICH 2012)

Im Falle der *Eigenfertigung* ist die Produktionssteuerung am Zug. Diese definiert zunächst im Rahmen der *Belegungsplanung* die Produktionsressource und den Zeitpunkt, zu dem jeder Auftrag zu bearbeiten ist. Im nächsten Schritt erfolgen die *Auftragsfreigabe* sowie die anschließende *Auftrags-* und *Ressourcenüberwachung*. Eine etwas andere Einordnung definiert LÖDDING (2008). Hier werden die Auftragsfreigabe, Reihenfolgebildung und Kapazitätssteuerung als die drei wesentlichen Aufgaben der Produktionssteuerung definiert.

In der Literatur findet sich keine einheitliche Aufgabenteilung zwischen Planung und Steuerung (ZÄPFEL 2001, DÖRMER & GÜNTHER 2013). Je nach Granularität der Planung können einzelne Schritte wie Auftragsterminierung oder Belegungsplanung auch der Steuerung beziehungsweise der Planung zugeordnet werden

(ZÄPFEL 2001, HANSMANN, WIENDAHL 2010). Im weiteren Verlauf der vorliegenden Arbeit werden die Auftragsterminierung und Belegungsplanung nicht mehr als Teil der Produktionssteuerung verstanden. Im Allgemeinen sind die resultierende Belastung der Produktionsressourcen und dadurch die Terminierung der Aufträge umso genauer zu bestimmen, je näher der geplante Bearbeitungszeitpunkt rückt (WIENDAHL 2010). Aus diesem Grund sind die Planung und Steuerung von Produktionsaufträgen in zyklisch ablaufenden Stufen mit zunehmender Genauigkeit üblich (WIENDAHL 2010).

Zur Differenzierung von Planung und Steuerung kann ebenfalls der *Regelkreis der PPS* herangezogen werden (vgl. Abbildung 8). Während das Produktionssystem dabei als Regelstrecke betrachtet wird, agiert die Produktionssteuerung als Regler. Basierend auf den rückgemeldeten Ist-Betriebsdaten liegt die Aufgabe der Steuerung darin, das von der Planung vorgegebene Produktionsprogramm trotz auftretender Störeinflüsse umzusetzen, wie zum Beispiel der Ausfall von Produktionsressourcen. Es wird deutlich, dass die Produktionssteuerung im Vergleich zur Planung eine hohe Reaktionsfähigkeit aufweisen muss, um mit ausreichender Schnelligkeit auf ungeplante Ereignisse reagieren zu können. (LÖDDING 2008)

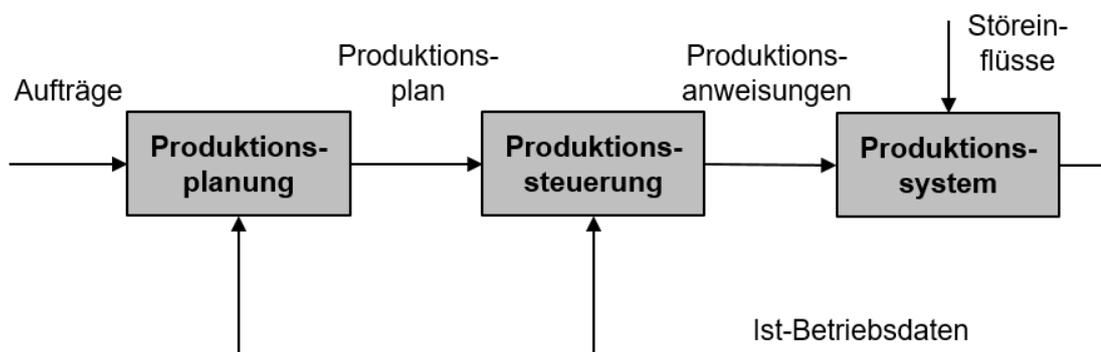


Abbildung 8: Regelkreis der PPS (in Anlehnung an HACKSTEIN 1989, WIENDAHL 2010)

2.3.2 Zielgrößen

Das zentrale Ziel marktwirtschaftlicher Unternehmen ist die Gewinnmaximierung. Aus diesem Grund steht die Wirtschaftlichkeit auch im Mittelpunkt der PPS (WIENDAHL 2010). Die PPS kann auf zwei Wegen zum wirtschaftlichen Erfolg eines Unternehmens beitragen. Zum einen trägt die Sicherstellung der marktseitigen Anforderungen im Hinblick auf den Fertigstellungszeitpunkt zur Kundenzufriedenheit und damit zum mittel- und langfristigen Unternehmenserfolg bei (LÖDDING

ET AL. 2012, SCHÄFER ET AL. 2016). Zum anderen ist die PPS ebenfalls in der Lage, die Produktionskosten von Produkten zu beeinflussen (NYHUIS & MÜNZBERG 2010). Zur Maximierung der Wirtschaftlichkeit ist die PPS also gefordert, die beeinflussbaren Faktoren zur Zielerfüllung und Kostenreduzierung zu berücksichtigen und daraus optimale Entscheidungen abzuleiten. Die erforderlichen Informationen für diese Bewertung sind zum Zeitpunkt der Entscheidung jedoch oft nicht verfügbar oder im allgemeinen schwer zu quantifizieren (BRACKEL 2009, ACKER 2011).

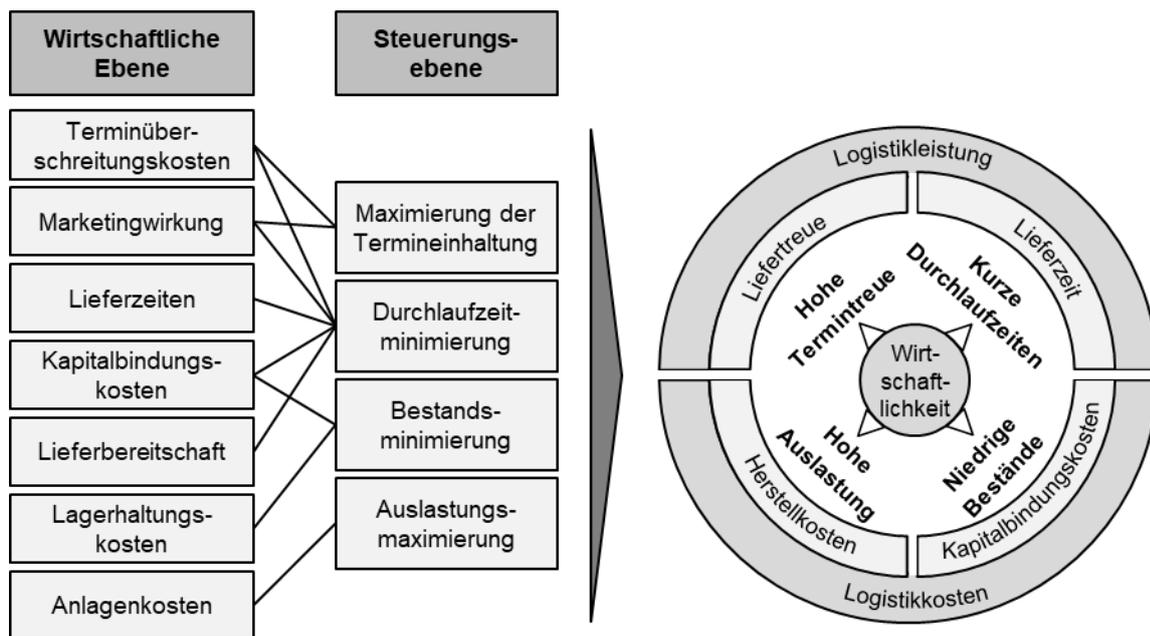


Abbildung 9: Zielgrößen der Produktionslogistik links und das daraus abgeleitete logistische Zielsystem rechts (in Anlehnung an WIENDAHL (2010))

Die komplexen wirtschaftlichen Einflussfaktoren der PPS können, wie in Abbildung 9 auf der linken Seite dargestellt, auf vier wesentliche Zielgrößen der Steuerungsebene heruntergebrochen werden. Auf Basis dieses Ansatzes lässt sich das logistische Zielsystem der PPS ableiten (vgl. Abbildung 9, rechts). Die marktseitigen, terminlichen Anforderungen werden unter der Logistikleistung in Form der beiden Zielgrößen Liefertreue und Lieferzeit zusammengefasst. Zur Minimierung der Logistikkosten sind niedrige Herstellkosten durch eine hohe Auslastung und geringe Kapitalbindungskosten im Zuge niedriger Bestände erforderlich. Da eine hohe und gleichmäßige Auslastung gleichzeitig die Verfügbarkeit von Rohmaterial und Zulieferteilen in Form von hohen Beständen erfordert, zeigt sich hier ein inhärenter Zielkonflikt. Im Zuge hoher Bestände steigt zudem die Durchlaufzeit, was wiederum die Termintreue und Lieferzeit gefährden kann. (TEMPELMEIER 2003, WIENDAHL 2010)

Der dargelegte Zielkonflikt wirft die Frage nach der Gewichtung der Zielgrößen auf (WIENDAHL & WEDEMEYER 1993, LÖDDING 2008). Während bei ungesättigten Märkten beispielsweise die Fokussierung auf eine hohe Auslastung zielführend ist, steht bei starkem Konkurrenzdruck tendenziell eine niedrige Durchlaufzeit im Vordergrund (BRACKEL 2009). Die eher technische Formulierung der Zielgrößen der PPS erlaubt zudem keine direkte Analyse des Einflusses einzelner Steuerungsentscheidungen auf die Wirtschaftlichkeit (ADAM 1998). Kritiker bemängeln daher, dass in den klassischen Zielen der PPS die klare betriebswirtschaftliche Bewertung und damit die Orientierung an den eigentlichen, übergeordneten Unternehmenszielen wie Kostenminimierung oder Rentabilitätsmaximierung fehlen (MERTENS & WIENDAHL 1992, DANNHAUSER 2000, MOSLER 2017). Trotz dieser Tatsache wird im Kontext der PPS überwiegend auf die logistischen Zielgrößen zurückgegriffen, weil die genaue monetäre Quantifizierung nur schwer möglich ist (SCHULTZ & MERTENS 2000, BRACKEL 2009, ACKER 2011).

Neben den in Abbildung 9 aufgeführten Bestandteilen auf wirtschaftlicher Ebene werden in der Literatur noch weitere Faktoren genannt, die von der PPS beeinflusst werden können. Dazu gehören etwa Rüstkosten, die durch Rüstvorgänge vor der Bearbeitung unterschiedlicher Aufträge auf Maschinen anfallen, sowie Leer- beziehungsweise Stillstand- oder Transportkosten (REHWINKEL 1978, HOITSCHE 1993, ACKER 2011). Die Material- und Betriebsstoffkosten bleiben durch die PPS unberührt, da sie lediglich den Zeitpunkt des Verbrauchs beeinflussen, jedoch nicht die benötigten Rohstoffmengen bei der Herstellung (DANNHAUSER 2000). Im Falle von flexiblen, über die Zeit variierenden Preisen, ist diese Annahme jedoch nicht mehr korrekt. In Zeiten fluktuierender Strompreise (vgl. Abschnitt 2.2.4) beeinflusst die PPS somit darüber hinaus auch die Stromkosten (BEIER ET AL. 2015, KELLER 2018, SCHULTZ 2018).

Die gleichzeitige Berücksichtigung aller Zielgrößen birgt große Herausforderungen für die Produktionssteuerung und erfordert nicht zuletzt eine Kompromissfindung. Zur Vereinfachung wird daher bei manchen Steuerungsstrategien lediglich eine bestimmte Zielgröße verfolgt. Mehrdimensionale beziehungsweise „multikriterielle“ Ansätze, bei denen gleichzeitig mehrere Ziele angestrebt werden, erfordern eine Gewichtung der Zielgrößen zueinander. (LÖDDING 2008, DOMSCHKE ET AL. 2015, NIEHUES 2017)

2.4 Ansätze zur Produktionssteuerung

2.4.1 Problemformulierung Job Shop Scheduling

Die Aufgabe der PPS ist es, den Bearbeitungszeitpunkt jedes Auftrags auf der jeweiligen Produktionsressource festzulegen und dabei die angestrebten Zielgrößen möglichst optimal zu erfüllen. In der englischen Fachliteratur wird dieser Vorgang oft *Scheduling* oder im Falle einer kurzfristigen Umplanung *Rescheduling* genannt (VIEIRA ET AL. 2003, OUELHADJ & PETROVIC 2009). Das sich daraus ableitende mathematische Optimierungsproblem wird infolgedessen als *Job Shop Scheduling Problem* (JSSP) bezeichnet und lässt sich folgendermaßen formulieren:

Eine Anzahl von n Aufträgen (Jobs) ist auf m Produktionsressourcen zu fertigen. Alle Aufträge bestehen aus jeweils einem Set an Arbeitsschritten, die es abzuarbeiten gilt. Jede Produktionsressource kann gleichzeitig nicht mehr als einen Job bearbeiten. Außerdem muss jeder Arbeitsschritt nur einmalig ausgeführt werden. Das JSSP ist ein komplexes mathematisches Problem, da die Anzahl der möglichen Lösungen mit zunehmender Problemgröße exponentiell ansteigt. (MUTH & THOMPSON 1963, EILON & CHOWDHURY 1976, FRENCH 1986)

2.4.2 Steuerungsstrategien

In der Fachliteratur existiert eine Vielzahl von Ansätzen zur Produktionssteuerung beziehungsweise zum englischen Pendant *Rescheduling*. Sie lassen sich in die beiden wesentlichen Strategien *reaktive* und *prädiktiv-reaktive Steuerung* einteilen (VIEIRA ET AL. 2003, BRACKEL 2009, OUELHADJ & PETROVIC 2009, NIEHUES 2017). Neben diesen beiden Steuerungsstrategien gibt es noch weitere Ansätze, die sich nicht eindeutig in dieses Schema einordnen lassen (NIEHUES 2017). Hervorzuheben sind *Agentensysteme* (AS), die sich durch ihre dezentrale Architektur auszeichnen und somit einen Paradigmenwechsel im Vergleich zu den reaktiven und prädiktiv-reaktiven Ansätzen darstellen (OUELHADJ & PETROVIC 2009). Die Möglichkeit zur Kombination von AS mit neuartigen Methoden der künstlichen Intelligenz, insbesondere Reinforcement Learning, stellt dabei einen vielversprechenden Ansatz dar (VEITH ET AL. 2020). In Abbildung 10 werden die drei aufgezeigten Steuerungsansätze eingeordnet und die verwendeten Lösungsverfahren skizziert. Eine detaillierte Vorstellung der Ansätze erfolgt in den nachfolgenden Unterkapiteln.

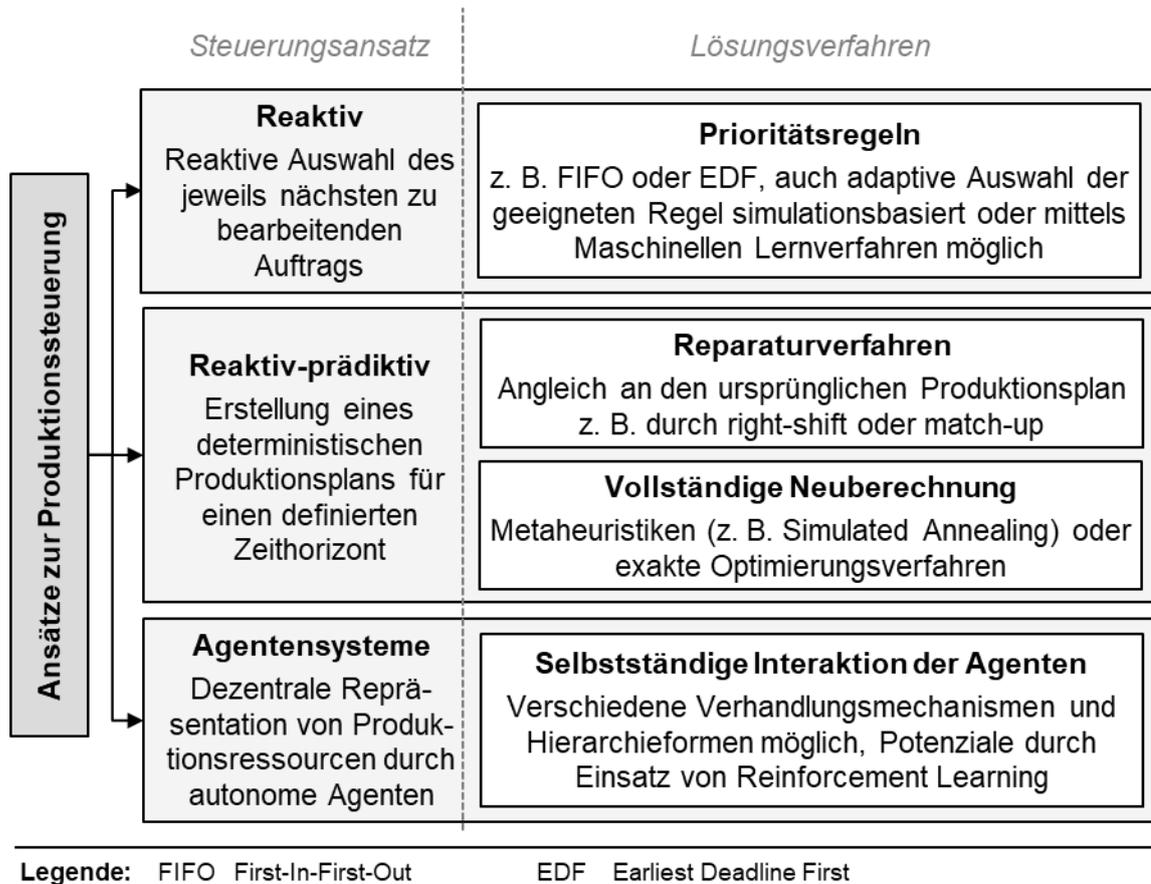


Abbildung 10: Übersicht über die verbreiteten Ansätze zur Produktionssteuerung

2.4.3 Reaktive Produktionssteuerung

Bei Anwendung einer reaktiven Produktionssteuerung wird kein definierter Produktionsplan erstellt. Vielmehr erfolgt nach Ende eines Bearbeitungsschrittes iterativ die Auswahl des nächsten zu bearbeitenden Auftrags. Der Entscheidungshorizont reicht demnach lediglich bis zur Fertigstellung des aktuellen Bearbeitungsschrittes. Es existiert keine vorab geplante Auftragsreihenfolge. Die Steuerung erfolgt in der Regel in Form von *Prioritätsregeln*, die dezentral auf der Maschinenebene angewendet werden. Verbreitete Verfahrensregeln sind beispielsweise *First in First Out (FIFO)* oder *Earliest Deadline First (EDF)*. Da kein umfangreicher Produktionsplan erstellt wird, erfordern reaktive Steuerungsansätze in der Regel einen vernachlässigbaren Rechenaufwand, wodurch eine hohe Reaktionsfähigkeit gewährleistet ist. Dagegen kann auf diese Weise kein optimaler Produktionsplan erreicht und lediglich eine sehr begrenzte Zahl an Einflüssen in der Entscheidungsfindung berücksichtigt werden. (BRACKEL 2009, OUELHADJ & PETROVIC 2009, LÖDDING 2013, NIEHUES 2017)

Reaktive Steuerungsansätze sind aufgrund der einfachen Implementierung weit verbreitet und waren in der Vergangenheit Gegenstand intensiver Forschung (LÖDDING 2008). Die dabei verwendeten Prioritätsregeln greifen in den meisten Fällen auf Bearbeitungszeiten und Fertigstellungstermine zurück, um daraus einfache Regeln abzuleiten (VIEIRA ET AL. 2003). Einen Überblick über bestehende Prioritätsregeln geben PANWALKAR & ISKANDER (1977), BLACKSTONE ET AL. (1982) oder GUPTA ET AL. (1989). Insbesondere im Kontext flexibler Produktionssysteme kommen die Vorteile von Prioritätsregeln zum Tragen (KRISHNAN ET AL. 2012).

Die Leistung von Prioritätsregeln hängt von vielen Faktoren ab: etwa von den ausgewählten Entscheidungskriterien, der Konfiguration des Produktionssystems oder der Auslastung (CHO & WYSK 1993). Unter diesen Gesichtspunkten ist es vorteilhaft, wenn die Auswahl der Prioritätsregeln adaptiv und situationsabhängig erfolgt, was entweder simulationsbasiert oder mit Hilfe maschineller Lernverfahren möglich ist (SHAW ET AL. 1992, PRIORE ET AL. 2001). In Kombination mit einer *Simulation* kann vorab die Wirkung unterschiedlicher Prioritätsregeln verglichen werden. Auf diese Weise wird die geeignete Steuerungsregel iterativ bewertet und ausgewählt (SHAFAEI & BRUNN 1999, SABUNCUOGLU & BAYIZ 2000). Der größte Nachteil ist dabei die für die Simulation erforderliche lange Rechenzeit (PRIORE ET AL. 2014).

Mit Hilfe von *maschinellen Lernverfahren* ist es möglich, die zu verwendenden Prioritätsregeln situationsbasiert vor jeder zu treffenden Entscheidung auszuwählen (PRIORE ET AL. 2001). Dadurch kann im Vergleich zu klassischen Prioritätsregeln ein deutlich größerer Betrachtungsraum in der Entscheidungsfindung berücksichtigt werden (JIMÉNEZ 2012). Ansätze mit maschinellem Lernen eignen sich daher sehr gut für dynamische und sich schnell verändernde Produktionssysteme (OUELHADJ & PETROVIC 2009). Es existieren zahlreiche statistische Verfahren, die im Zuge dessen zum Einsatz kommen. Am verbreitetsten sind Support Vector Machines, Neuronale Netze, Case-based Reasoning, Fuzzy Logic und Reinforcement Learning (RL) (SHAHZAD & MEBARKI 2016). Während alle anderen genannten Verfahren mit Hilfe bestehender Beispieldaten trainiert werden, lernt RL durch direkte Interaktion mit dem Produktionssystem (PRIORE ET AL. 2014). Zum genaueren Verständnis der Funktionsweise von RL sei an dieser Stelle auf Abschnitt 2.5 verwiesen.

2.4.4 Prädiktiv-reaktive Produktionssteuerung

Eine *prädiktiv-reaktive Produktionssteuerung* erstellt einen deterministischen Produktionsplan. Dabei wird der Startzeitpunkt aller zu bearbeitenden Fertigungsschritte innerhalb eines definierten Betrachtungshorizonts vorausschauend geplant. Die Umplanung kann periodisch oder aufgrund bestimmter Ereignisse getriggert werden. Aus mathematischer Sicht muss dazu das bereits beschriebene JSSP gelöst werden. Dies ist die größte Herausforderung von prädiktiv-reaktiven Ansätzen, da je nach Größe des zu lösenden Problems ein hoher Rechenaufwand erforderlich ist, was dem Ziel einer kurzen Reaktionszeit entgegenwirkt (VIEIRA ET AL. 2003, BRACKEL 2009, OUELHADJ & PETROVIC 2009, NIEHUES 2017)

Zur Umsetzung von prädiktiv-reaktiven Steuerungsstrategien existieren zwei wesentliche Lösungsprinzipien. Treten unerwartete Ereignisse wie etwa Störungen auf, können gültige Produktionspläne durch *Reparaturverfahren* oder durch eine *vollständige Neuberechnung* angepasst werden (VIEIRA ET AL. 2003). Unter den Reparaturverfahren sind die zeitliche Verschiebung der Aufträge („right-shift“) und die sogenannte „match-up“-Strategie verbreitet (OUELHADJ & PETROVIC 2009). Bei Letzterer werden die Aufträge so eingeplant, dass das Produktionsprogramm zu einem späteren Zeitpunkt wieder mit dem ursprünglichen Plan übereinstimmt (AKTURK & GORGULU 1999). Derartige Verfahren sind leicht zu implementieren und erfordern eine geringe Rechenzeit. Gleichzeitig ist die Lösungsgüte vergleichsweise gering (OUELHADJ & PETROVIC 2009).

Eine Neuberechnung erfordert eine kurzfristige Optimierung der Feinplanung (FATTAHI & FALLAHI 2010). Zu diesem Zweck existieren zahlreiche Lösungsalgorithmen, deren zentrale Herausforderung darin liegt, die Neuplanung innerhalb einer vertretbaren Zeit abzuschließen. Aus diesem Grund können optimierende Lösungsverfahren, wie etwa Lineare Programmierung oder der Branch-and-Bound-Algorithmus, nur begrenzt angewendet werden (BIERWIRTH 1993, FISCHER 2007). In den meisten Fällen werden heuristische Ansätze bevorzugt, die eine gute Lösung berechnen, jedoch nicht garantiert das Optimum finden (BRACKEL 2009). Verbreitete Verfahren sind dabei Simulated Annealing, genetische Algorithmen oder Tabu-Suche (HERRMANN 2006, OUELHADJ & PETROVIC 2009, SHEN ET AL. 2014, NIEHUES 2017). Dennoch sind derartige Ansätze mit hohem Rechenaufwand verbunden, was die Reaktionsfähigkeit begrenzt (VIEIRA ET AL. 2003, SHAHZAD & MEBARKI 2012, KHODKE & BHONGADE 2013).

2.4.5 Agentensysteme

Eine Möglichkeit zur Bewältigung von Komplexität ist das Zerlegen der Problemstellung in Einzelteile. Um eine derartige Dekomposition zu erreichen, können dezentrale AS eingesetzt werden. Agenten werden im Kontext der Automatisierungstechnik nach VDI-RICHTLINIE 2653 definiert als „(...) eine abgrenzbare (Hardware- und/ oder Software-) Einheit mit definierten Zielen. Ein Agent ist bestrebt, diese Ziele durch selbständiges Verhalten zu erreichen und interagiert dabei mit seiner Umgebung und anderen Agenten.“ Zu den charakteristischen Eigenschaften eines Agenten zählt die Kapselung von Zustand und Verhalten, wodurch ein Agent das Gesamtsystem nicht durch eine globale, sondern lediglich durch eine lokale Sicht wahrnimmt (WAGNER ET AL. 2003). Bei Anwendungen im Bereich der Informatik und Künstlichen Intelligenz weicht die Auffassung zu Agenten teilweise an dieser Stelle ab. Vielmehr wird den Agenten eine gesamtheitliche, globale Sicht auf den Zustand des Gesamtsystems eingeräumt (GABEL 2009), was ebenfalls als Verständnisgrundlage im Kontext der vorliegenden Arbeit dient.

Im Zuge von AS agieren mehrere derartige Agenten innerhalb eines Systems, die ihre eigenen Aktionen und Zustände kontrollieren können (SHEN ET AL. 2006). Um ihre vorgegebenen Ziele zu erreichen, interagieren und kooperieren die Agenten miteinander (BARBATI ET AL. 2012). Je nach Anwendungsfall existieren zahllose verschiedene Systemarchitekturen, die sich unter anderem durch ihre Organisationsstruktur unterscheiden (LEITÃO 2009, VEITH ET AL. 2020). Während in nicht-hierarchischen Systemen alle Agenten gleichberechtigt agieren, existieren in hierarchischen und zentralisierten Systemen klare Rangordnungen. Die Wege der Entscheidungsfindung reichen dabei von starren Hierarchien bis hin zu Vertragssystemen, auf deren Basis die Agenten miteinander verhandeln (CARIDI & CAVALLIERI 2004). AS eignen sich besonders für die Lösung komplexer und hochdynamischer Steuerungsprobleme, wie etwa cyber-physische Produktionssysteme (VOGEL-HEUSER ET AL. 2015) oder verteilte Energienetze (NEBEL-WENNER ET AL. 2019, LUDWIG ET AL. 2018). In den meisten Anwendungsfällen von AS in der PPS repräsentieren die Agenten physische Produktionsressourcen wie Maschinen oder Werkzeuge (SHEN ET AL. 2006), die auch mit einer Komponenten des Energiesystems kombiniert werden können (REHBERGER ET AL. 2017).

Zu den wesentlichen Vorteilen von Agentensystemen zählen Fehlerrobustheit, Flexibilität und Reaktionsfähigkeit. Große Nachteile sind das Agieren nach Einzelzielen und die geringe Nachvollziehbarkeit. Da die Agenten lediglich über eine lokale Sicht verfügen, finden Agentensysteme in der Regel keine optimale Lösung.

Zudem ist eine Einbeziehung von globalen Auswirkungen in die Entscheidungsfindung nur begrenzt möglich. (HENSELER 1998, CARIDI & SIANESI 2000, SHEN ET AL. 2001, LEITÃO 2009, BARBATI ET AL. 2012, THEISS 2015)

Methoden aus dem Bereich der künstlichen Intelligenz und des maschinellen Lernens stellen eine Möglichkeit dar, um die Kooperation zwischen den Agenten zu verbessern und den genannten Nachteilen von AS entgegenzuwirken (JIMÉNEZ 2012). So konnten GABEL & RIEDMILLER (2007) in ihrer Arbeit durch den Einsatz von RL mit einem Agentensystem ein nahezu ideales Systemverhalten erreichen. Im Kontext der Produktionssteuerung wurden derartige Ansätze mit Multi-Agent Reinforcement Learning (MARL) etwa in GABEL (2009), JIMÉNEZ (2012), WASCHNECK ET AL. (2018), WANG ET AL. (2019C) oder KUHNLE ET AL. (2019) erfolgreich eingesetzt. Im nachfolgenden Unterkapitel wird näher auf die theoretischen Grundlagen von Reinforcement Learning und dessen Anwendung im Bereich von Agentensystemen eingegangen.

2.5 Reinforcement Learning

2.5.1 Einführung

Die Herangehensweise von RL ist an das menschliche Lernverhalten angelehnt. Die grundlegende Funktionsweise ist in Abbildung 11 dargestellt. Ein Agent lernt dabei selbständig, innerhalb einer Umgebung zu agieren und durch sein Verhalten langfristig eine Belohnung zu maximieren. Die Wahrnehmung der Umgebung erfolgt auf Basis eines Zustands s_t , der die aktuelle Situation der Umgebung widerspiegelt. Der Agent wählt auf dieser Grundlage die Aktion a_t aus, die im aktuellen Zustand optimal erscheint. Die Auswirkung der Aktion innerhalb der Umgebung zieht einen neuen Zustand s_{t+1} und eine bestimmte Belohnung r_t nach sich. Durch das Sammeln von Erfahrungen lernt der Agent, über die Zeit bessere Entscheidungen zu treffen und die Belohnung zu maximieren (SUTTON & BARTO 2018).

Die mathematische Formalisierung von RL-Problemen erfolgt durch den Markov-Entscheidungsprozess. Dabei werden in einer Sequenz Zuständen bestimmte Aktionen zugeordnet, um langfristig eine Belohnung zu maximieren. Die Markov-Eigenschaft muss dabei zwangsläufig erfüllt sein und besagt, dass der Übergang von einem Zustand zum nächsten und die zugehörige Belohnung ausschließlich vom Ausgangszustand der ausgewählten Aktion, nicht jedoch von vorherigen Entscheidungen abhängig ist. (KAEBLING ET AL. 1996, PUTERMAN 2009)

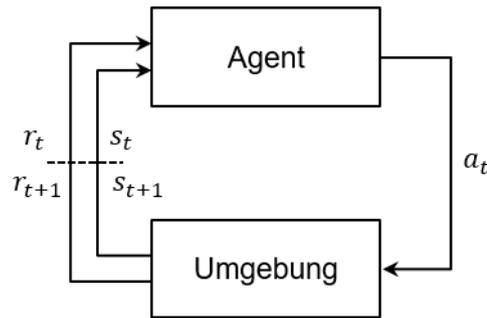


Abbildung 11: Funktionsprinzip von RL (SUTTON & BARTO (2018))

2.5.2 Erlernen der optimalen Strategie

Das Ziel von Reinforcement Learning ist, eine Strategie (engl. policy) π zu finden, die eine Maximierung der Belohnung ermöglicht. Die theoretischen Grundlagen wurden dabei in SUTTON & BARTO (1998) gelegt. Während der Agent agiert und von Zustand zu Folgezustand gelangt, erhält er eine Reihe von Belohnungen $R_t = \sum_{k=0}^{\infty} \gamma^k r(t+k)$. Es gilt, den Erwartungswert $E[R_t | s_0, \pi]$ der nachfolgenden Summe aller Belohnungen zu maximieren, welche für den aktuellen Zeitschritt t und alle nachfolgenden Zeitschritte k zu erwarten sind:

$$R_t = \sum_{k=0}^{T-t-1} \gamma^k r_{t+k+1} \quad (1)$$

- R_t Erwartungswert der Belohnung zum Zeitschritt t
- γ^t Discount-Faktor
- $r(t)$ direkte Belohnung im Zeitschritt t
- T Dauer der Trainingsepisode
- k Zählvariable für alle nach Zeitschritt t folgenden Zeitschritte

Der Discount-Faktor γ kann einen Wert zwischen 0 und 1 annehmen und bestimmt, wie stark zukünftige Belohnungen in die Gesamtberechnung einfließen. Um nun einem bestimmten Zustand einen Erwartungswert zuordnen zu können, wird eine sogenannte Wertefunktion (engl. value-function) berechnet:

$$v_{\pi}(s) = E_{\pi}[R_t | s_t = s] = E_{\pi} \left[\sum_{k=0}^{T-t-1} \gamma^k (r_{t+k+1} | s_t = s) \right] \quad (2)$$

- $v_{\pi}(s)$ Zustands-Wertefunktion der Strategie π im Zustand s
- E_{π} Erwartungswert der Strategie π
- R_t Erwartungswert der Belohnung zum Zeitpunkt t
- s_t Zustand zum Zeitpunkt t

r_t	Belohnung zum Zeitpunkt t
γ^t	Discount-Faktor
T	Dauer der Trainingsepisode
k	Zählvariable für alle nach Zeitschritt t folgenden Zeitschritte

Analog dazu kann diese Zustands-Wertefunktion $v_\pi(s)$ auch in eine Aktions-Wertefunktion $q^\pi(s)$ umformuliert werden. Auf dem Weg, eine optimale Strategie zu entwickeln, unterstützt die Wertefunktion nun dabei, diese quantifizieren zu können. Eine Strategie π' ist besser als eine Strategie π , wenn der Wert der Zustands-Wertefunktion jedes Zustands für Strategie π' größer als jener der Zustands-Wertefunktion von π ist. Die optimale Zustands-Wertefunktion beschreibt gleichzeitig auch die optimale Strategie π^* . Gleiches gilt damit für die optimale Aktions-Wertefunktion.

Wird eine optimale Wertefunktion gefunden, lässt sich daraus die optimale Strategie ableiten. Zur Ermittlung der optimalen Wertefunktion kann die sogenannte Bellmann-Gleichung herangezogen werden, die eine schrittweise Aktualisierung der Wertefunktion für alle Zustände durchführt:

$$v_\pi(s) = E_\pi[r_{t+1} + \gamma v_\pi(s_{t+1} | s_t = s)] \quad (3)$$

v_π	Zustands-Wertefunktion der Strategie π
E_π	Erwartungswert der Strategie π
r_t	Reward zum Zeitpunkt t
γ	Discount-Faktor
s_t	Zustand zum Zeitpunkt t

Eine analoge Umformulierung kann auch für die Aktions-Wertefunktion vorgenommen werden. Insgesamt lässt sich auf diese Weise der Erwartungswert berechnen und damit die optimale Strategie durch diverse Iterationen bestimmen, ohne dass alle einzelnen Belohnungen r_i bekannt sind. (BELLMAN & DREYFUS 2015, SUTTON & BARTO 2018)

2.5.3 Lernverfahren

Aufbauend auf dem in Abschnitt 2.5.2 hergeleiteten Prinzip wurden bisher zahlreiche RL-Algorithmen entwickelt. Abbildung 12 gibt einen Überblick über die wichtigsten Verfahren. Diese können grundlegend in modellbasierte und -freie Ansätze unterteilt werden. Während modellbasierte Algorithmen ein Modell ihrer Umgebung nutzen, um verschiedene Aktionen zu simulieren, ist bei modellfreien

Algorithmen das Umgebungsmodell nicht bekannt. Letztere betrachten die Umgebung als Black Box und verbessern sich nur anhand der gesammelten Erfahrung, da bei Ausführung dieses Ansatzes weniger Rechenzeit zur Simulation verschiedener Aktionen erforderlich ist. Dieser Aspekt hat zur Folge, dass insbesondere bei Systemen mit mehreren RL-Agenten, wie in der vorliegenden Arbeit, modellbasierte Ansätze eine untergeordnete Rolle spielen. Im weiteren Verlauf dieser Arbeit wird daher lediglich auf modellfreie Verfahren eingegangen. Modellbasierte Ansätze werden in Abbildung 12 ausgeklammert. (DAW ET AL. 2005, DOYA ET AL. 2002, DAYAN & NIV 2008)

Modellfreie Algorithmen können in die beiden Kategorien Policy Optimierungsalgorithmen und Q-Learning basierte Ansätze eingeteilt werden. Zudem existieren Mischformen, die eine Hybridform beider Ansätze bilden. Verfahren auf Basis von Q-Learning leiten implizit aus der Aktions-Wertefunktion eine Strategie ab, indem diese während des Trainings iterativ optimiert wird. Beim weitverbreiteten Deep-Q-Learning (DQL) wird wie bei vielen anderen vergleichbaren Algorithmen auch ein neuronales Netz zur Approximation der Wertefunktion verwendet. (MNIH ET AL. 2013)

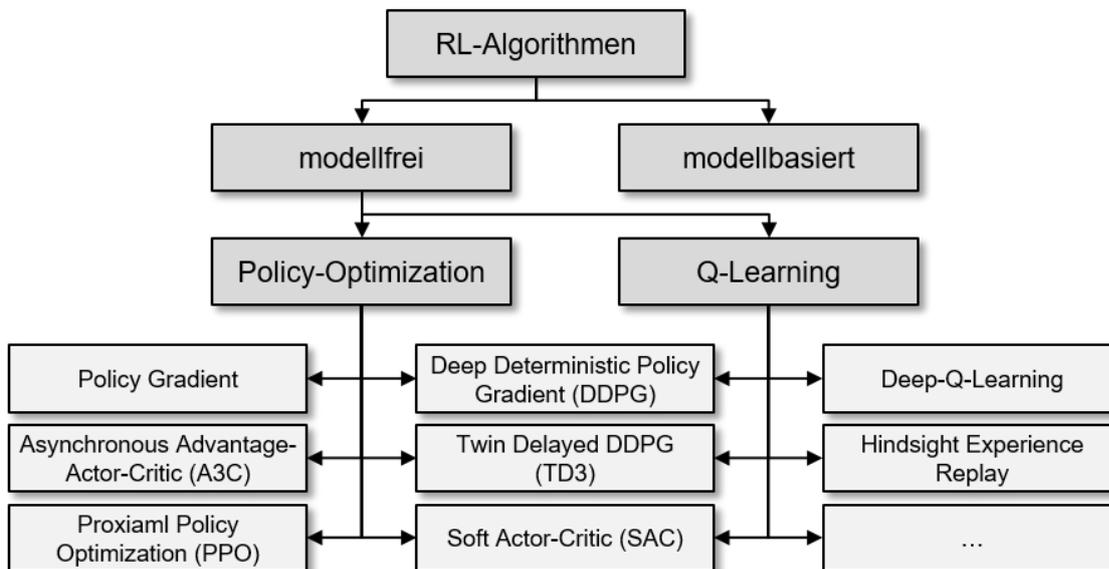


Abbildung 12: Auswahl der wesentlichen RL-Algorithmen
(in Anlehnung an OPENAI 2018)

Im Kontext der Policy-basierten Ansätze ist Proximal Policy Optimization (PPO) aufgrund der breiten Anwendung in der Literatur hervorzuheben (MINNE ET AL. 2019). Wie alle Policy-Ansätze stellt auch dieses Verfahren einen direkten Zusammenhang zwischen einem Zustand und der auszuführenden Aktion her, der

über den Gradienten-Abstieg anhand einer Kostenfunktion optimiert wird. Die dadurch erlernte Strategie hat somit ein stochastisches Verhalten. Das Ergebnis ist eine Wahrscheinlichkeitsverteilung aller Aktionen (SCHULMAN ET AL. 2017). Hybride Strategien versuchen die besten Eigenschaften von Q-Learning und Policy-Optimierung zu vereinen. So ist etwa Deep Deterministic Gradient Descent (DDPG) eine deterministische Strategie, die ein direktes Mapping zwischen Zustand und Aktion in Kombination mit einer Aktion-Wertefunktion wiedergibt (LILLICRAP ET AL. 2015).

Die Performance der einzelnen Verfahren hängt maßgeblich von der Parametrierung des jeweiligen Algorithmus ab. Um die optimale Einstellung der Parameter zu finden, werden in der Regel iterativ Versuche mit unterschiedlichen Parametersätzen durchgeführt und diese miteinander verglichen. Dieser zeitintensive Schritt wird auch Hyperparametertuning genannt. Neben den einzelnen Parametern der Lernalgorithmen können auch die Vergabe und Skalierung der Belohnung einen maßgeblichen Einfluss auf das Lernverhalten ausüben. Insgesamt können somit zur Lösung eines Problems unterschiedliche RL-Algorithmen eingesetzt werden. Eine pauschale Aussage vorab, mit welchem Algorithmus das beste Ergebnis erzielt werden kann, ist kaum möglich. (HENDERSON ET AL. 2018, SCHULMAN ET AL. 2017, HAARNOJA ET AL. 2018, BERNER ET AL. 2019)

2.5.4 Reinforcement Learning in Agentensystemen

Über die Anwendung mit einem Agenten hinaus kann RL auch in Systemen mit mehreren Agenten eingesetzt werden. Dabei kann es sich um konkurrierende oder kooperierende Agenten handeln. Auch Mischformen sind denkbar. Systeme mit kooperierenden Agenten basieren dabei auf zwei wesentlichen Säulen. Zum einen erhalten alle Agenten eine gemeinsame Belohnung. Auf diese Weise können die Agenten lernen, durch ihre individuellen Aktionen gemeinsam eine Zielfunktion zu maximieren. Zum anderen ist ein globaler Beobachtungsraum förderlich für das gemeinsame Lernen. Agenten sind dadurch in der Lage, in ihren Entscheidungen auch den aktuellen Zustand der anderen Agenten zu berücksichtigen, deren Verhalten abzuschätzen und auf diese Weise aufeinander abgestimmt zu agieren. Da alle Agenten während des Trainings ihre eigene Strategie kontinuierlich verbessern, wird die Umgebung aus Sicht der einzelnen Agenten somit instationär. Dennoch konnten mit MARL in der Vergangenheit bereits komplexe Probleme nah-optimal gelöst werden. (TAN 1997, GABEL 2009, ZHANG ET AL. 2019).

2.6 Zusammenfassung und Fazit

Aufbauend auf der in Kapitel 1 beschriebenen Ausgangssituation und Zielsetzung wurden in diesem Kapitel die relevanten Grundlagen beleuchtet. Dazu wurde zunächst in Unterkapitel 2.2 ein Überblick über die Möglichkeiten zur Anpassung des Energiebedarfs in produzierenden Unternehmen gegeben. Die Energiebezugsoptionen Strombeschaffung, Eigenerzeugung und Nutzung von Stromspeichern wurden betrachtet und die Begriffe Lastgang und Energieflexibilität eingeführt. In Abschnitt 2.3 wurden die Grundlagen der Produktionsplanung und -steuerung vorgestellt sowie deren Bedeutung und Aufgaben eingeordnet. Ein zweiter Fokus lag dabei auf den zu berücksichtigenden Zielgrößen und deren Ausprägung. Darauf aufbauend erfolgte im Unterkapitel 2.4 eine Zusammenfassung der existierenden Lösungsansätze zur Produktionssteuerung, wobei die wichtigsten Lösungsansätze vorgestellt und deren Vor- und Nachteile analysiert wurden. Mit Reinforcement Learning wurde abschließend auf einen dieser Ansätze in Abschnitt 2.5 detaillierter eingegangen und die grundlegende Funktionsweise aufgezeigt. Zudem wurde die Übertragung dieser Herangehensweise auf Agentensysteme skizziert. Die dargestellten Grundlagen bilden im Hinblick auf die Energiebezugsoptionen von produzierenden Unternehmen sowie auf die Möglichkeiten zur Steuerung von Produktion und Energieverbrauch die Basis für die Betrachtungen im weiteren Verlauf dieser Arbeit.

3 Stand der Forschung

3.1 Übersicht

Im folgenden Kapitel wird der für die vorliegende Arbeit relevante Stand der Forschung analysiert. Im Hinblick auf die in Abschnitt 1.3 abgeleitete Zielsetzung gliedert sich der Stand der Forschung in die Themenblöcke „Energieflexibilität in der Produktion“ (Abschnitt 3.2) und „bestehende Ansätze zur gesamtkostenbasierten Produktionssteuerung“ (Abschnitt 3.3). Auf dieser Basis erfolgt die Ableitung des Handlungsbedarfs sowie der resultierenden Anforderungen für eine energieorientierte und kostenbasierte Produktionssteuerung (Abschnitt 3.4).

3.2 Energieflexibilität in der Produktion

3.2.1 Einordnung der betrachteten Forschungsfelder

Das Gebiet der intelligenten Verteilung und Nutzung von Energie ist weltweit Gegenstand intensiver Forschung. Dementsprechend existieren zahlreiche wissenschaftliche Arbeiten zu diesem Thema.

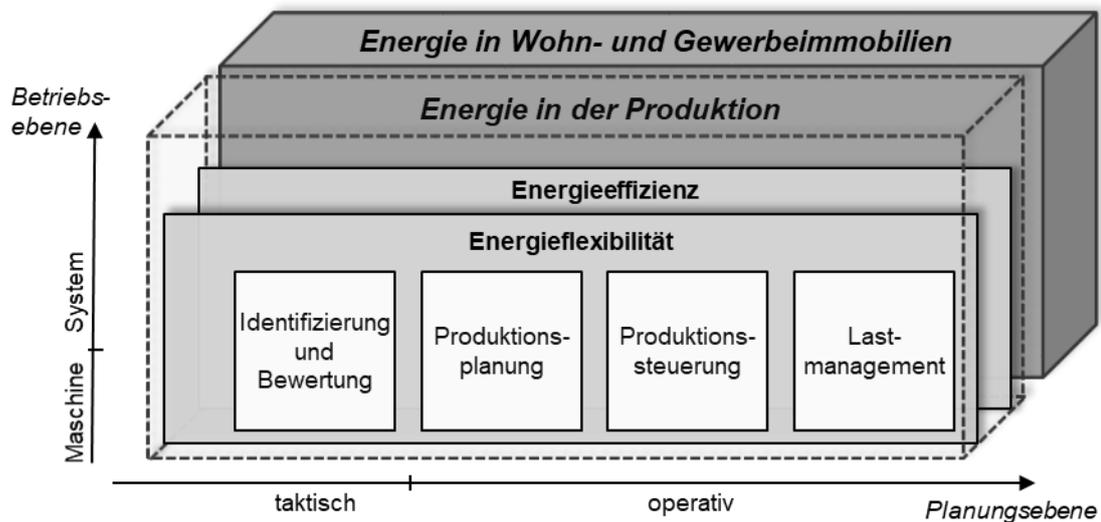


Abbildung 13: Einordnung der betrachteten Forschungsfelder

Um in Anbetracht der großen Zahl an Veröffentlichungen den Fokus auf die wesentlichen Aspekte zu fokussieren, wird der Betrachtungsraum eingegrenzt. Die bei der weiteren Analyse betrachteten Forschungsfelder können wie in Abbildung 13 eingeordnet werden. Eine erste Unterteilung erfolgt im Hinblick auf den Anwendungsbereich von *Energie in Wohn- und Gewerbeimmobilien* beziehungsweise *in der Produktion*. Eine weitere Einordnung ist anhand der betrachteten Zielgrößen *Energieeffizienz* und *Energieflexibilität* möglich. Die analysierten Ansätze können weiterhin anhand der *Planungsebene* in taktisch und operativ sowie anhand der betrachteten *Betriebsebene* in *Maschine* und *System* gruppiert werden. Der Fokus der wissenschaftlichen Arbeiten im Bereich der *Identifizierung und Bewertung* liegt auf einer taktischen Betrachtung. Im Zuge dieser wird das Energieflexibilitätspotenzial analysiert. Dieses Potenzial durch geeignete Maßnahmen zu heben, ist die Aufgabe der zeitlich nachgelagerten operativen *Produktionsplanung* und *Produktionssteuerung*. Darüber hinaus existieren weitere Ansätze, die durch geeignete Maßnahmen, wie die temporäre Abschaltung von Verbrauchern, den Energiebedarf kurzfristig für wenige Minuten reduzieren, ohne die Auftragsreihenfolge oder Kapazitätsteuerung zu verändern. Da derartige Ansätze keine umfassende Funktionalität einer Produktionssteuerung aufweisen, werden diese Arbeiten in den Bereich des kurzfristigen *Lastmanagements* (vgl. Abschnitt 2.2.1) eingeordnet.

Die nachfolgend aufgeführten Forschungsgebiete und Themenschwerpunkte sind daher explizit kein Gegenstand der weiteren Analyse:

- **Energie in Wohn- und Gewerbeimmobilien:** Die optimale Nutzung von Strom spielt auch im Kontext von Wohn- und Gewerbeimmobilien eine große Rolle in der aktuellen Forschung. Unter dem Stichwort „Smart-Grid“ werden Methoden zur Verknüpfung von Speichern, Stromerzeugern und Verbrauchern beispielsweise in GAO ET AL. (2015), NGUYEN & NGUYEN (2016) oder LU ET AL. (2018) entwickelt. Aufgrund der spezifischen Anforderungen und Randbedingungen ist eine Übertragbarkeit auf die Industrie sehr begrenzt möglich. Ansätze aus diesem Bereich werden im Zuge der vorliegenden Arbeit nicht näher analysiert.
- **Energieeffizienz in der Produktion:** Die Erhöhung der EE in der Produktion, insbesondere durch die PPS, thematisieren zahlreiche Beiträge. Aktuelle Arbeiten im Kontext der energieeffizienten Produktionsplanungen sind etwa LIU ET AL. (2014), MEZIANE & TAGHEZOUT (2018), LI ET AL. (2018B) oder LI ET AL. (2018A), während sich LANGER ET AL. (2014), LE & PANG (2013), WANG ET AL. (2017), KIM ET AL. (2017) oder NOURI ET AL. (2018)

auf der Ebene der Produktionssteuerung bewegen. Die Arbeiten verfolgen als zentrales Ziel die generelle Reduzierung des Energieverbrauchs und unterscheiden sich in diesem Punkt maßgeblich von Ansätzen zur Steigerung der EF.

- **Identifizierung und Bewertung von Energieflexibilität:** Die Identifizierung und Bewertung von EF bilden die Grundlage für den energieflexiblen Produktionsbetrieb. Beispielsweise entwickelte GRAßL (2015) in seiner Arbeit eine Methode zur allgemeinen Bewertung. Im Fokus von SIMON (2020) stehen die Risikobetrachtung und Bündelung von EF-Maßnahmen. Da es sich dabei um rein strategische beziehungsweise taktische Betrachtungen handelt, werden Ansätze aus diesem Bereich für die vorliegende Arbeit, die sich auf operativer Ebene bewegt, nicht genauer beleuchtet.
- **Ansätze auf Maschinenebene:** Mehrere Autoren präsentieren Lösungen zur energieorientierten Belegungsplanung und Steuerung einzelner Maschinen, etwa indem der Bearbeitungszeitpunkt der Aufträge optimiert wird (AGHELINJAD ET AL. 2016, LEE ET AL. 2017, RUBAIEE ET AL. 2018) oder inhärente Energieflexibilitätpotenziale der Maschinen genutzt werden (POPP ET AL. 2016, BRUGGER ET AL. 2017). Da der Fokus der vorliegenden Arbeit auf der Steuerung eines Produktionssystems, bestehend aus mehreren einzelnen Produktionsressourcen, liegt, werden Lösungsansätze für einzelne Maschinen nicht näher beleuchtet.

3.2.2 Einordnung der bestehenden Ansätze

Die bisherigen Arbeiten im Bereich der Energieflexibilität in der Produktion lassen sich in die Forschungsfelder Produktionsplanung, Produktionssteuerung und Lastmanagement (vgl. Abbildung 13) einordnen. Zusätzlich ist eine Differenzierung der Ansätze nach der Dimension der Zielgröße (vgl. Abschnitt 2.3.2) möglich. In einem Großteil der Fälle werden Entscheidungen lediglich hinsichtlich einer Zielgröße, der Reduzierung der Energiekosten getroffen. Im Zuge dieser eindimensionalen Betrachtung werden andere Zielgrößen als feste Randbedingungen definiert, beispielsweise die Einhaltung eines Mindestdurchsatzes oder die Fertigstellung aller Aufträge innerhalb des Betrachtungszeitraums. Arbeiten mit einer zweidimensionalen Zielgröße versuchen den Zielkonflikt zwischen zwei Zielen auszugleichen, indem neben den Energiekosten eine weitere Zielgröße der Produktion, wie etwa Bestands- oder Verzugskosten, optimiert wird. Drei- oder auch mehrdimensionale Ansätze versuchen, diesen Ausgleich entsprechend auf drei oder mehr Zielgrößen auszuweiten.

Stand der Forschung

In Abbildung 14 sind die relevanten Arbeiten unter dem Gesichtspunkt der betrachteten Betriebsebenen und der Zielgrößendimension eingeteilt. In den nachfolgenden Unterkapiteln werden die eingetragenen Arbeiten diskutiert.

		Forschungsfeld		
		Produktionsplanung	Produktionssteuerung	Lastmanagement
Zielgröße	Eindimensional	WANG & LI 2013 PECHMANN ET AL. 2012 RAGER ET AL. 2015 BRUZZONE ET AL. 2012 ZHANG ET AL. 2017C ZHANG ET AL. 2017A,B SUN ET AL. 2018 ADDISU ET AL. 2017 BIEL 2017 GROSCH ET AL. 2019 RUIZ DUARTE 2020	NEUGEBAUER ET AL. 2012 ZHAI ET AL. 2017A,B FAZLI KHALAF & WANG 2018 BATISTA ABIKARRAM ET AL. 2019	LIN ET AL. 2012 FLOREA ET AL. 2013 ZHOU & LI 2013 WEINERT & MOSE 2016 YU ET AL. 2016 SUN ET AL. 2016 CHOOBINEH & MOHAGHEGHI 2016 WECKMANN ET AL. 2017 BEIER 2017 WEITZEL & GLOCK 2018
	Zweidimensional	FAN ET AL. 2011 ARTIGUES ET AL. 2013 KÜSTER ET AL. 2013 LUO ET AL. 2013 STOCK & SELIGER 2015 ZHOU ET AL. 2018 GONG ET AL. 2018 KELLER 2018 ISLAM ET AL. 2018 LIU ET AL. 2020	FERNANDEZ ET AL. 2013 WILLEKE ET AL. 2016 SCHULTZ ET AL. 2017	
	Dreidimensional	HAOUASSI ET AL. 2016 BUSSE ET AL. 2017 DABABNEH ET AL. 2018	GROÙE BÖCKMANN 2014 WANG ET AL. 2019	SUN & LI 2014
	...			

Abbildung 14: Einordnung der bestehenden Ansätze zur Energieflexibilität in der Produktion

3.2.3 Energieflexibilität in der Produktionsplanung

Nachfolgend werden die Ansätze zur Energieflexibilität in der Produktionsplanung vorgestellt und diskutiert. Die betrachteten Arbeiten werden anhand der Dimension der verwendeten Zielfunktionen gruppiert.

Eindimensionale Zielgröße

Alle im folgenden genannten Arbeiten haben gemein, dass sie sich zum einen auf der Ebene der Produktionsplanung bewegen und zum anderen die Planung ausschließlich mit dem Ziel der Energiekostenminimierung ausführen. Die Einhaltung anderer Zielgrößen, wie etwa Bestände oder Termintreue, wird als feste Randbedingung formuliert.

Als einige der ersten Autoren erzielen BRUZZONE ET AL. (2012) eine Anpassung des Energiebedarfs durch eine Verschiebung der Bearbeitungszeitpunkte. PECHMANN ET AL. (2012) gehen in ihrer PPS-Software darüber hinaus und präsentieren ein System zur Belegungsplanung unter Berücksichtigung der Energiekosten. Zudem wird der daraus resultierende Lastgang vorhergesagt, der an den Energieversorger weitergegeben werden kann. Einen weiteren Ansatz zur Belegungsplanung präsentieren WANG & LI (2013), wobei mit Hilfe einer Metaheuristik die Stromkosten minimiert werden und sich die Produktivität gleichzeitig oberhalb einer vorgegebenen Schwelle bewegt. Um den Energiebedarf zu glätten, formuliert RAGER (2008) eine energieorientierte Belegungsplanung, die in RAGER ET AL. (2015) um die Einbindung volatiler Energiepreise erweitert wird.

Ebenfalls durch die Fokussierung auf Stromkosten können ZHANG ET AL. (2017C) die Produktionskosten eines Stahlwerkes senken. Das identische Ziel verfolgt der Ansatz von ZHANG ET AL. (2017B), der vor allem auf die Ausnutzung flexibler Strompreise ausgerichtet ist. In ZHANG ET AL. (2017A) wird das Optimierungsmodell für die Integration von Strom aus volatiler Eigenerzeugung erweitert. Der Ansatz von ADDISU ET AL. (2017) steuert die Maschinenauslastung energieorientiert durch die Reihenfolgebildung an dem angebotenen Fördersystem. SUN ET AL. (2018) betrachten in ihrem Ansatz neben den Produktionsaufträgen auch Instandhaltungsmaßnahmen, die unter der Randbedingung variabler Strompreise mit dem Ziel der Kostenminimierung durch ein mathematisches Modell eingeplant werden.

Neben elektrischem Strom betrachtet BIEL (2017) auch Wärme mit zugehörigen Speichern in der Produktionsplanung. Der Optimierungsalgorithmus zur Belegungsplanung ist zudem in der Lage, den Strombedarf an die volatile Erzeugung von Windkraftanlagen anzupassen. Darüber hinaus entwickelt der Autor eine stochastische Optimierung, indem unterschiedliche Erzeugungsszenarien generiert werden, um so kurzfristige Schwankungen in der Energieverfügbarkeit ausgleichen zu können (BIEL ET AL. 2018). Im Ansatz von GROSCHE ET AL. (2019) werden die Energiekosten mit Hilfe eines genetischen Algorithmus unter Berücksichtigung von schwankenden Strompreisen und Eigenerzeugungsanlagen minimiert.

Einen ähnlichen Ansatz mit einer robusten Optimierung präsentieren RUIZ DUARTE ET AL. (2020), um den Strom aus EE und Anreize für Demand Response in die Produktionsplanung zu integrieren.

Zweidimensionale Zielgröße

Einen Ansatz mit einer zweidimensionalen Zielgröße präsentieren FANG ET AL. (2011). Das entwickelte mathematische Modell minimiert gleichzeitig die Energiekosten und die Zykluszeit. Als einige der ersten Autoren berücksichtigen KÜSTER ET AL. (2013) variable Strompreise in der Belegungsplanung, die mit Hilfe einer evolutionären Optimierung gelöst wird. Neben Energiekosten wird die Höhe der Bestände minimiert. LUO ET AL. (2013) diskutieren in ihrem Beitrag die Wechselwirkung zwischen der Flexibilisierung des Energiebedarfs und anderen Zielgrößen in der Produktion. Zur Ausnutzung variabler Strompreise werden anschließend mit Hilfe eines Ameisenalgorithmus die Energiekosten und die Durchlaufzeit minimiert. Um kostspielige Lastspitzen bereits in der Produktionsplanung zu vermeiden, präsentieren ARTIGUES ET AL. (2013) ein zweistufiges Optimierungsmodell, das zudem gleichzeitig versucht, die Liefertreue zu maximieren.

STOCK & SELIGER (2015) präsentieren einen weiteren Ansatz zur energieorientierten Feinplanung, um eine schwankende Energieverfügbarkeit aus Eigenerzeugungsanlagen auszugleichen und gleichzeitig die Fertigungsdauer zu minimieren. Das identische Ziel verfolgen ZHOU ET AL. (2018) in ihrem zweidimensionalen Optimierungsmodell für Batchprozesse. Der Fokus von ISLAM ET AL. (2018) liegt auf der Integration von Eigenerzeugungsanlagen, wobei zur Einhaltung der Produktionsziele zusätzlich Strafkosten für eine Durchsatzreduzierung angesetzt werden. Das mathematische Modell von GONG ET AL. (2015) zielt auf die Reduzierung der Energiekosten durch Nutzung schwankender Strompreise in der Feinplanung ab. Die rechtzeitige Bearbeitung aller Aufträge ist dabei eine feste Prämisse. Das Modell wurde in weiteren Arbeiten um die Möglichkeit erweitert, Aufträge in Nachtschichten zu verschieben und auf diese Weise Energiekosten zu reduzieren (GONG ET AL. 2017B, 2017A, GONG ET AL. 2018).

KELLER (2018) präsentiert eine umfängliche Methodik zur energieorientierten Auftragsplanung. Neben preis- und mengenflexibler Strombeschaffung werden auch Eigenerzeugungsanlagen in der Losgrößen- und Kapazitätsplanung berücksichtigt (KELLER & REINHART 2016). Das formulierte Optimierungsproblem zur Festlegung der Produktionsmengen minimiert die Energie- und Lagerkosten mit Hilfe einer Heuristik. Der Betrachtungshorizont reicht bis zum Vortag der eigent-

lichen Bearbeitung der Aufträge. Die Produktionssteuerung wird explizit ausklammert. LIU ET AL. (2020) wenden einen energieorientierten Planungsalgorithmus, der gleichzeitig auch die Durchlaufzeit optimiert, auf eine reale Glasproduktion an und können auf diese Weise die Energiekosten deutlich reduzieren.

Dreidimensionale Zielgröße

HAOUASSI ET AL. (2016) haben sich in ihrer Arbeit zum Ziel gesetzt, einen Kompromiss zwischen Energiekosteneinsparungen und den klassischen Zielgrößen Termintreue und Bestände zu finden. BUSSE ET AL. (2017) stellen eine schrittweise und regelbasierte Durchlaufterminierung im Rahmen von schwankenden Strompreisen vor und rücken dabei den entstehenden Zielkonflikt in den Fokus. Eine potenzielle energieorientierte Auftragsverschiebung wird daher mit den entstehenden Auslastungsverlusten und Verzugskosten abgewogen und nur im Falle eines resultierenden finanziellen Vorteils umgesetzt. DABABNEH ET AL. (2018) bauen auf den Ansatz von SUN ET AL. (2016) auf und stellen in ihrer Arbeit ebenfalls einen integrierten Planungsalgorithmus für Produktionsaufträge und Instandhaltungsmaßnahmen vor. Neben den Energiekosten werden zusätzlich die Instandhaltungskosten sowie fiktive Straf- und Bonuskosten für Durchsatzabweichungen einbezogen. Die resultierende dreidimensionale Zielfunktion wird mit Hilfe einer Partikelschwarmoptimierung minimiert.

Fazit aus den Vorarbeiten zu Energieflexibilität in der Produktionsplanung

Die Analyse der Arbeiten zur energieflexiblen Produktionsplanung zeigt, dass das Feinplanungsproblem in der Regel um den Faktor Energie erweitert und mit Hilfe eines Optimierungsalgorithmus gelöst wird. Aufgrund des Planungshorizontes von mehr als einem Tag kann die daraus resultierende, längere Rechenzeit in Kauf genommen werden. Einige Autoren, insbesondere HAOUASSI ET AL. (2016), BUSSE ET AL. (2017) und DABABNEH ET AL. (2018), adressieren den Zielkonflikt zwischen Energieflexibilität und den logistischen Zielgrößen durch eine mehrdimensionale Zielfunktion. Neben den Energiekosten werden Instandhaltungs-, Verzugs- oder Bestandskosten in das Optimierungsproblem integriert. Energiespeicher wie etwa Batterien finden auf der Ebene der Produktionsplanung keine Berücksichtigung. Diese eignen sich kaum für die Speicherung von Energie über den Zeithorizont der Produktionsplanung von mehreren Tagen oder Wochen, da die Speicherkosten mit abnehmender Anzahl an Speicherzyklen deutlich ansteigen (KONDIELLA ET AL. 2013, ANDRESEN ET AL. 2017). Die Steuerung einer Batterie ist somit der nachgelagerten Produktionssteuerung zuzuordnen, die deutlich kurzfristi-

ger agiert. Die analysierten Ansätze sind zudem nicht in der Lage, auf unvorhergesehene Störungen zu reagieren. Um kurzfristige Maschinenausfälle oder Preisänderung betrachten zu können, ist eine Kopplung mit einer reaktionsfähigen Produktionssteuerung erforderlich.

3.2.4 Energieflexibilität in der Produktionssteuerung

Analog zum vorausgegangenen Abschnitt der Produktionsplanung werden nachfolgend die Vorarbeiten im Bereich der Produktionssteuerung analysiert. Lösungsansätze, die auch funktionelle Komponenten für das Lastmanagement beinhalten, werden dabei alleinig der Produktionssteuerung zugeordnet.

Eindimensionale Zielgröße

Mit eniMES präsentierten NEUGEBAUER ET AL. (2012) ein Framework zur Steuerung und Überwachung des Energiebedarfs, indem Schichtpläne in Abhängigkeit variabler Strompreise erstellt und Materialpuffer ausgenutzt werden. Um eine kurzfristige Anpassung an Strompreisschwankungen zu ermöglichen, entwickeln ZHAI ET AL. (2017B) in ihrer Arbeit ein prädiktives Steuerungsmodell zur echtzeitfähigen Feinplanung auf Basis einer gemischt ganzzahligen Optimierung. Energiekosten sind dabei die alleinige Zielgröße, wobei ein vorgegebener Durchsatz aufrechterhalten wird. In ZHAI ET AL. (2017A) wird das Modell um volatile Stromeigenerzeugung durch WKA erweitert. Zufällig auftretende Störungen im Produktionssystem werden nicht betrachtet.

FAZLI KHALAF & WANG (2018) entwerfen einen integrierten Feinplanungsansatz mit Hilfe einer zweistufigen ganzzahligen Programmierung, wobei neben dem volatilen Strommarkt auch Eigenerzeugung und Stromspeicher abgebildet werden. Das System ist in der Lage, kurzfristige Preisschwankungen und Abweichungen der Eigenerzeugungsprognose einzubeziehen. Störungen werden dabei nicht betrachtet. Die Autoren weisen auf den erforderlichen Rechenaufwand zur Lösung des Problems hin. Die Zielgröße ist die Minimierung der Energiekosten, während gleichzeitig ein vorgegebener Durchsatz aufrechterhalten werden muss. Einen echtzeitfähigen Ansatz zur energieorientierten Belegungsplanung stellen BATISTA ABIKARRAM ET AL. (2019) in ihrer Arbeit vor. Die Fertigstellung aller Aufträge innerhalb einer Schicht ist eine feste Randbedingung, während die Energiekosten minimiert werden. Bei der Analyse der Rechenzeit wird deutlich, dass der erforder-

derliche Aufwand mit der Anzahl der Produktionsressourcen exponentiell zunimmt. Maschinenausfälle und andere unvorhergesehene Ereignisse in der Produktion werden nicht berücksichtigt.

Zweidimensionale Zielgröße

FERNANDEZ ET AL. (2013) stellen eine Methode zur Reduzierung von Spitzenlast durch die Verwendung von Materialpuffern vor, die in Zeiten von Spitzenlasten abgebaut werden. Auf diese Weise bleibt der Durchsatz des Produktionssystems unverändert. Gleichzeitig werden die Lager- und Energiekosten miteinander abgewogen. Die Integration von Energieaspekten in konventionelle Steuerungsverfahren verfolgt WILLEKE (2018), um auf diese Weise volatile Strompreise auszunutzen. Um den Energiekosten und der Liefertreue Rechnung zu tragen, wird eine Auftragsfreigabe und Reihenfolgebildung auf Basis von Prioritätsregeln vorgestellt, die sich aus Energieaspekten und der verbleibenden Schlupfzeit zusammensetzen (WILLEKE ET AL. 2015A, WILLEKE ET AL. 2015B, WILLEKE ET AL. 2016, WILLEKE ET AL. 2018). Auf diese Weise werden die Stromkosten bei gleichzeitiger Verschlechterung der Liefertreue reduziert (WILLEKE ET AL. 2016). Eine energiekostenorientierte Rückstandsregelung wählt zudem den Zeitpunkt von Kapazitätserhöhungen unter Berücksichtigung des Strompreises aus (WILLEKE ET AL. 2015C).

Prioritätsregeln bilden ebenfalls die Basis für die Reihenfolgebildung in SCHULTZ ET AL. (2016A), wobei erneut Schlupfzeiten und die Energieverfügbarkeit berücksichtigt werden. Der Fokus liegt auf der Einhaltung des vorgegebenen Lastgangs, unvorhergesehenen Störungen in der Produktion zum Trotz (SCHULTZ ET AL. 2015). Die entwickelten Auftragsfreigabeverfahren und die Kapazitätssteuerung tragen zudem zur Reduzierung der Stromkosten bei. Gleichzeitig wird die Liefertreue negativ beeinflusst (SCHULTZ 2018). Zudem werden auch periphere Verbraucher an die Produktionssteuerung mit Hilfe einer gemischt ganzzahligen Optimierung gekoppelt, um ein kurzfristiges Lastmanagement zu ermöglichen (SCHULTZ ET AL. 2017). Erneuerbare Energien zur Stromeigenerzeugung werden in das System genauso wie schwankende Preise am Spotmarkt integriert.

Dreidimensionale Zielgröße

Einen weiteren Ansatz für eine energieorientierte Produktionssteuerung stellt GROBE BÖCKMANN (2014) vor, wobei Energieeffizienz und -flexibilität berücksichtigt werden. Ein mathematisches Modell zur erweiterten Maschinenbelegungsplanung minimiert neben Energiekosten auch Herstellungs- und Verzugskosten

und ist ein wesentlicher Bestandteil der prädiktiven Produktionssteuerung, die innerhalb eines Planungshorizontes von wenigen Stunden bis zu einigen Tagen agiert. Die nachgelagerte Produktionsregelung ist dagegen in der Lage, kurzfristige Abweichungen im Sinne eines Lastmanagements innerhalb von Minuten regelbasiert auszugleichen, etwa durch einen kurzfristigen Lastabwurf in Form einer Unterbrechung von Bearbeitungsprozessen. Änderungen der Reihenfolge sind dann nicht mehr möglich. Der Ansatz von WANG ET AL. (2019A) basiert auf einem System aus mehreren, miteinander verhandelnden Agenten. Die Zielgrößen Durchlaufzeit, Energiekosten und Produktionskosten werden nicht gleichzeitig, sondern vielmehr bei den Entscheidungen der einzelnen Agenten rollierend verfolgt. Aus diesen Gründen lässt der Ansatz schnelle Reaktionen auf unvorhergesehene Ereignisse zu. Produktionsressourcen können jedoch nur in den Stand-by-Zustand wechseln, falls kein entsprechender Auftrag auf eine Bearbeitung wartet.

Fazit aus den Vorarbeiten zu Energieflexibilität in der Produktionssteuerung

Zur Zusammenfassung der analysierten Ansätze zur Produktionssteuerung sind diese in Abbildung 15 gegenübergestellt und anhand der für die vorliegende Arbeit elementaren Kriterien *Strombezug*, *Zielgrößen*, *Funktionalität* und des verwendeten *Steuerungsansatzes* bewertet.

Auf Seiten des *Strombezugs* wird deutlich, dass alle untersuchten Arbeiten einen *preisvariablen Strombezug* von extern mit volatilen Preisen berücksichtigen. Lediglich in den Arbeiten von ZHAI ET AL. (2017A) sowie FAZLI KHALAF & WANG (2018) kann überschüssiger Strom kurzfristig in das Netz eingespeist werden. Obwohl alle Autoren von variablen Strompreisen ausgehen, ist mit den untersuchten Ansätzen kein aktiver Stromhandel unter den an der Strombörse gegebenen Randbedingungen (vgl. 2.2.4) möglich. Der Grund hierfür liegt in der erforderlichen Vorlaufzeit von mindestens fünf Minuten. Zur Ausnutzung variabler Strompreise muss die Energie bereits vor dem eigentlichen Verbrauch gekauft werden. Diese vorweggenommene Kaufentscheidung als zentraler Schritt beim Stromhandel wurde bisher bei keiner der untersuchten Arbeiten umgesetzt. Darüber hinaus wurden *Eigenerzeugung* und die Nutzung von *Stromspeichern* jedoch in den meisten Arbeiten nicht einbezogen.

Im Hinblick auf die betrachteten *Zielgrößen* werden neben Energiekosten nur in wenigen Fällen weitere Kostenfaktoren der Produktion – wie Termintreue, Bearbeitungskosten und Bestände – berücksichtigt. Hervorzuheben ist hier die Arbeit von GROBE BÖCKMANN (2014), die insgesamt drei Zielgrößen abbildet. Das Agen-

tensystem in WANG ET AL. (2019A) ist dagegen nicht in der Lage, alle drei modellierten Kostenfaktoren gleichzeitig zu verfolgen. Eine umfangreiche Analyse zu den relevanten Zielgrößen neben den Energiekosten wurde bisher in keiner der vorliegenden Arbeiten durchgeführt.

		NEUGEBAUER ET AL. 2012	FERNANDEZ ET AL. 2013	GROBE BÖCKMANN 2014	SCHULTZ ET AL. 2017	WILLEKE ET AL. 2016	ZHAI ET AL. 2017A,B	FAZLI KHALAF & WANG 2018	BATISTA ABIKARRAM ET AL. 2019	WANG ET AL. 2019
Strombezug	Preisvariabler Netzbezug	●	●	●	●	●	●	●	●	●
	Kurzfristiger Stromhandel	○	○	○	○	○	○	○	○	○
	Eigenerzeugung	○	○	○	●	○	●	●	○	○
	Stromspeicher	○	○	○	○	○	○	●	○	○
Zielgrößen	Energiekosten	●	●	●	●	●	●	●	●	◐
	Termintreue	○	○	●	●	●	○	○	○	○
	Bearbeitungskosten	○	○	●	○	○	○	○	○	◐
	Durchlaufzeit	○	○	○	○	○	○	○	○	◐
	Bestände	○	●	○	○	○	○	○	○	○
Funktionalitäten	Integriertes Lastmanagement	●	●	●	●	○	○	●	●	●
	Reaktion auf Störungen	●	○	●	●	●	○	○	○	●
Steuerungsansatz	Reaktiv	●	●	●	●	●	○	○	○	○
	Prädiktiv	○	○	●	○	◐	●	●	●	○
	Agentensystem	○	○	○	○	○	○	○	○	●

Abbildung 15: Bisherige Arbeiten im Bereich von Energieflexibilität in der Produktionssteuerung

Aus der Untersuchung zu den *Funktionalitäten* der betrachteten Systeme geht hervor, dass die meisten Arbeiten im Kontext der Produktionssteuerung auch in der Lage sind, eine kurzfristige Energieverbrauchsanpassung im Sinne eines Lastmanagements zu erreichen. Die Integration eines intelligenten Lastmanagements in die Produktionssteuerung wurde somit von zahlreichen Autoren für sinnvoll erachtet und umgesetzt. Demgegenüber haben sich einige Arbeiten mit den Auswir-

kungen von Störungen im Kontext variabler Strompreise auseinandergesetzt. Auffällig ist allerdings, dass dies lediglich bei reaktiven *Steuerungsansätzen* der Fall ist. In prädiktiven Ansätzen, die vorausschauend einen optimalen Produktionsablauf suchen, wurden Störungen bisher nicht integriert. Dies kann unter anderem auf die deutlich geringere Reaktionsfähigkeit von prädiktiven Steuerungsansätzen zurückgeführt werden (vgl. 2.4.2).

3.2.5 Lastmanagement im Umfeld produzierender Unternehmen

Gezieltes Lastmanagement zur Vermeidung von Lastspitzen wird in vielen Unternehmen bereits heute eingesetzt und basiert hauptsächlich auf einfachen, regelbasierten Ansätzen. Die Entwicklung effektiverer und komplexerer Verfahren ist aktuell Gegenstand intensiver Forschung. Die relevanten Arbeiten in diesem Kontext sind nachfolgend zusammengefasst.

Eindimensionale Zielgröße

Die Arbeit von LIN ET AL. (2012) zeigt ein kurzfristiges Lastmanagement auf, wobei der Durchsatz des Produktionssystems nicht beeinträchtigt werden soll. Maßnahmen zur temporären Abschaltung von Ressourcen werden dazu simulationsbasiert ausgewählt. Änderungen in der Auftragsreihenfolge werden nicht vorgenommen. Ein System zum Spitzenlastmanagement wird in FLOREA ET AL. (2013) skizziert. Ein Agentensystem ist kurzfristig in der Lage, den Strombedarf anzupassen. Die Verbraucher der technischen Gebäudeausrüstung (TGA) werden durch jeweils einen Energieagenten repräsentiert, der auf Anfrage den Strombedarf reduzieren kann. Das System von ZHOU & LI (2013) detektiert automatisch den Engpass des Produktionssystems und ist dadurch in der Lage, Ressourcen temporär abzuschalten, ohne den Durchsatz zu beeinflussen.

WEINERT & MOSE (2016) präsentieren einen weiteren regelbasierten und dezentralen Ansatz für das industrielle Lastmanagement, indem Produktionsprozesse punktuell verlängert werden. Die Reduzierung der Spitzenlast geht damit zu Lasten der Fertigungsdauer. Einen zentralen und robusten Optimierungsansatz für das Lastmanagement von Produktionsbetrieben beschreiben YU ET AL. (2016). Erneut wird auf Materialpuffer zurückgegriffen, damit der Durchsatz unverändert bleibt. Ebenfalls vorausschauend agiert der Ansatz von SUN ET AL. (2016) der den Strombedarf der Produktion und den der technischen Gebäudeausrüstung steuert. Ein zentraler Optimierer schaltet Produktionsressourcen und Teile der TGA in Abhängigkeit der Stromkosten temporär ab, während ein erforderlicher Mindestdurchsatz

aufrechterhalten wird. CHOUBINEH & MOHAGHEGHI (2016) betrachten in ihrer Lösung neben schwankenden Strompreisen auch Stromeigenerzeugung und Batteriespeicher. Ziele der nichtlinearen Optimierung sind die Minimierung der Strombezugskosten und ein möglichst hoher Erlös aus dem Verkauf von eigenerzeugtem Strom. Der Bearbeitungszeitpunkt von Maschinen kann dabei verzögert werden, wenn bestimmte Mindestproduktionsvorgaben erfüllt sind.

WECKMANN ET AL. (2017) formulieren einen dezentralen Ansatz für das kurzfristige Lastmanagement zur Erhöhung des Stromeigenverbrauchs. Erneut gelten die Terminerfüllung und der Produktionsdurchsatz als feste Randbedingungen. Die dezentralen Entitäten prüfen und entscheiden selbstständig über temporäre Strombedarfsanpassungen. Auftragsreihenfolge und Maschinenbelegung bleiben unverändert. WEITZEL & GLOCK (2018) stellen ein Modell für kurzfristiges Lastmanagement vor, indem in der Produktionsplanung bereits vorab spezifische Lastkorridore berechnet werden. Diese vordefinierte Flexibilität kann beispielsweise ein Energieversorger bei Bedarf abrufen.

BEIER (2017) präsentiert in seiner Dissertation einen simulationsbasierten Ansatz, um den Energieverbrauch kurzfristig und ohne Prognosen an das Energieangebot anpassen zu können. Fluktuierende Eigenerzeugungsanlagen und Batteriespeicher werden in das Produktionssystem integriert (BEIER ET AL. 2016). Zur Flexibilisierung des Energiebedarfs werden zudem Verbraucher der TGA herangezogen (BEIER ET AL. 2015, 2017). Damit der Gesamtdurchsatz unverändert bleibt, werden im Zuge eines integrierten Lastmanagements lediglich die Ressourcen temporär abgeschaltet, welche nicht den Engpass im Materialfluss darstellen, ohne den Gesamtdurchsatz zu verändern. Eine Änderung der Auftragsreihenfolge oder -freigabe im Sinne einer Produktionssteuerung im Verständnis der vorliegenden Arbeit findet nicht statt.

Dreidimensionale Zielgröße

SUN & LI (2013) stellen in ihren Arbeiten ebenfalls ein echtzeitfähiges Lastmanagement für Produktionssysteme durch die Nutzung von Pufferbeständen vor. Der Durchsatz wird dabei nicht beeinflusst. Diese Prämisse wird in einer Weiterentwicklung des Systems aufgegeben (SUN ET AL. 2014). Die eingesparten Energiekosten werden den entstehenden Bestandskosten und Produktivitätseinbußen gegenübergestellt.

Fazit aus den Vorarbeiten zu Energieflexibilität und Lastmanagement

Die diskutierten Arbeiten zum industriellen Lastmanagement versuchen überwiegend eine kurzfristige Verbrauchsreduzierung zu erreichen, ohne dabei die logistischen Ziele zu beeinflussen. Eine Ausnahme ist die Arbeit von SUN ET AL. (2014), in der neben Energie auch andere Kostenfaktoren berücksichtigt werden. Verbreitete Maßnahmen im industriellen Lastmanagement sind temporäre Abschaltungen von Produktionsressourcen. Zudem werden periphere Verbraucher und die technische Gebäudeausrüstung einbezogen. Zur Maßnahmenauswahl werden sowohl reaktive und prädiktive Ansätze als auch Agentensysteme angewendet. Da die Produktionsressourcen dabei lediglich die Zustände „Produzierend“ oder „Stand-by“ einnehmen können, ist der Lösungsraum bei der prädiktiven Optimierung im Vergleich zu einer Produktionssteuerung, die zusätzlich die Auftragsreihenfolge berücksichtigt, deutlich kleiner. Das Optimierungsproblem kann daher schneller gelöst werden.

Da Auswirkungen auf die logistischen Zielgrößen weitestgehend ausgeschlossen werden, ergeben sich ein limitiertes Maßnahmenportfolio und – folglich – eine beschränkte Wirksamkeit. Eine Integration des Lastmanagements in die Produktionssteuerung ist aus diesem Grund anzustreben.

3.3 Gesamtkosten als Zielgröße in der PPS

Wie in Abschnitt 2.3.2 dargelegt, werden in der PPS in der Regel die logistischen Zielgrößen zur Maximierung der Logistikleistung und Minimierung der Logistikkosten herangezogen. Um dabei den beschriebenen Konflikt gegenläufiger Zielgrößen zu lösen, eignen sich kostenbasierte Ansätze, die eine gesamtheitliche Betrachtung und Abwägung ermöglichen. Nachfolgend wird ein Überblick über die bisherigen Ansätze gegeben, die auf Ebene der PPS eine Integration aller relevanten Kosten verfolgen und auf diese Weise eine monetäre Bewertung ermöglichen.

In einer der ersten Arbeiten zur kostenbasierten Produktionssteuerung entwarfen AGGARWAL ET AL. (1973) eine Prioritätsregel zur reaktiven Produktionssteuerung. Im Zuge dessen wurden Kosten für Maschinen, Rüstvorgänge und Bestände berücksichtigt sowie gleichzeitig Terminüberschreitungen monetär quantifiziert. SHAFAEI & BRUNN (1999) setzen sich in ihrer Arbeit zum Ziel, die wesentlichen für eine Umplanung relevanten Kosten bei einer Entscheidung einzubeziehen. Dieser Ansatz verspricht demnach deutliche monetäre Vorteile im Vergleich zu zeitbasierten Zielfunktionen wie Durchlaufzeit und Liefertreue, da diese Zielgrößen

konträr sein können und eventuell nicht alle entscheidungsrelevanten Aspekte beinhalten. Das Resultat ist eine Zielfunktion, die sich aus drei Kostenfaktoren zusammensetzt und aus dem Kapitelwert des jeweiligen Auftrags abgeleitet wird.

Der Fokus von OU-YANG & LIOU (1998) liegt dagegen auf der Kostenreduzierung im Falle von unvorhergesehenen Ereignissen in der Produktion. Dazu wird eine Methode entwickelt, wodurch die Auswirkungen auf alle nachfolgenden Aufträge in Form von Kosten quantifiziert werden können. Insgesamt werden dabei mit Verzugs-, Lohn- und Maschinenkosten drei Faktoren einbezogen. Ebenfalls, um das Problem der Abwägung von gegenläufigen Zielgrößen zu umgehen, baut BRINZER (2005) in seinem Ansatz zur Produktionsregelung eine vierdimensionale Zielfunktion auf, deren Elemente in Anlehnung an die betriebswirtschaftlich relevanten Kostenarten gewählt werden. Neben Rüst-, Bestands- und Verzugskosten werden dabei Kapazitätskosten formuliert, die beim Leerlauf einer Produktionsmaschine anfallen. Im Zuge des Ansatzes wird von nichtlinearen Abhängigkeiten der Kosten bei einer Änderung des Bestandes ausgegangen. Das Hauptaugenmerk von GUO ET AL. (2006) liegt darauf, die Termintreue zu verbessern. Dazu werden neben den Bestands- und Verzugskosten auch Lagerkosten modelliert, die für den Fall einer verfrühten Fertigstellung von Produktionsaufträgen anfallen. Durch die Modellierung der beiden zeitabhängigen Kostenbestandteile kann insgesamt auch die Durchlaufzeit deutlich minimiert werden, wenn gleichzeitig eine entsprechend terminierte Auftragsfreigabe durch die übergeordnete Planung erfolgt.

Die Besonderheit des Ansatzes von KUMAR & RAJOTIA (2006), im Vergleich zu den anderen Arbeiten, ist die Betrachtung von Werkzeugkosten, die neben drei weiteren Kostenfaktoren in die Entscheidung einbezogen werden. Gleichzeitig werden keine Aspekte der Logistikleistung integriert. Eine umfangreiche Zielfunktion präsentiert BRACKEL (2009) in seinem System zur simulationsbasierten Produktionssteuerung für die Werkstattfertigung. Das hat zum Ziel, die Produktionskosten insgesamt zu minimieren. Zu diesem Zweck werden sechs relevante Kostenfaktoren modelliert, die die Zielfunktion des entwickelten Optimierungsmodells darstellen.

Einen weiteren kostenbasierten Ansatz, um auf Störungen optimal zu reagieren, präsentieren NASERI & KUZGUNKAYA (2010). Dabei werden die anfallenden maschinen-, material- und auftragsbezogenen Kosten modelliert, wobei Letztere in Form von Nichtnutzungskosten zu Buche schlagen. Diese Zielfunktion wird dazu verwendet, bereits vorab berechnete Planungsalternativen im Falle eines stochastischen Ereignisses bewerten zu können. Die vierdimensionale Zielgröße von

ZHANG ET AL. (2015) ist hingegen speziell auf die Auftragsplanung in Stahlwerken zugeschnitten und wird auf Basis einer Partikelschwarmoptimierung minimiert. WANG & LUO (2016) integrieren die Losgrößenplanung in die kostenbasierte Betrachtung und entwerfen dazu jeweils eine vierdimensionale Zielfunktion.

NIEHUES (2017) entwickelt in seiner Dissertation ebenfalls eine kostenbasierte Zielfunktion, um das Problem gegenläufiger Zielgrößen innerhalb der Werkstattfertigung zu lösen. Insgesamt werden dazu sieben verschiedene Kostenfaktoren auf Basis einer Literaturrecherche abgeleitet. Im Zuge einer anschließenden Analyse werden Bestands- und Beschleunigungskosten als nicht relevant eingestuft, da die Kosten vernachlässigbar gering ausfallen beziehungsweise die Möglichkeit einer beschleunigten Abarbeitung ausgeschlossen wird. Die Zielfunktion resultiert somit in insgesamt fünf Kostenfaktoren. Im Zentrum des Ansatzes von POLOTSKI ET AL. (2017) steht die rüstopoptimale PPS, wobei die Rüstkosten mit Verzugs-, Bestands- und Maschinenkosten mit Hilfe eines Optimierungsmodells abgewogen werden. Die Losgrößenplanung ist ebenfalls ein Teil des Planungsansatzes von HU & HU (2018), der auf Nachfrageschwankungen zugeschnitten ist. Zur Minimierung der Gesamtsystemkosten werden neben vier weiteren Faktoren auch Kosten für Überstunden modelliert.

Alle Arbeiten haben gemein, dass die Seite der Erlöse nicht weiter betrachtet wird. Die der Steuerung zeitlich vorgelagerte Produktionsplanung gibt ein Produktionsprogramm mit einer bestimmten Produktionsmenge vor, wodurch ebenfalls bereits potenziellen Erlöse festgelegt werden (ACKER 2011). Aus diesem Grund kann auf eine Betrachtung der Erlöse im Zuge der Produktionssteuerung verzichtet werden. Eine Zusammenfassung der untersuchten Arbeiten und der dabei berücksichtigten Kostenfaktoren ist in Abbildung 16 dargestellt. Die identifizierten Kosten werden gemäß dem logistischen Zielsystem nach WIENDAHL (2010) in die Kategorien Logistikleistung und Logistikkosten eingeteilt. Im Bereich der Logistikleistung bilden die *Terminüberschreitungskosten* den am häufigsten verwendeten Kostenfaktor. Sie werden in den betrachteten Arbeiten als lineare Funktion modelliert. Zudem werden andererseits teilweise auch zu früh fertiggestellte Produkte mit *Lagerkosten* beaufschlagt. Die etwa in REUTER (1995, S. 142) angesprochenen, für die Lieferung des fertigen Produktes an den Kunden anfallenden *Transportkosten* wurden bisher nicht als Teil der Zielgröße in der PPS modelliert.

		AGGARWAL ET AL. 1973	SHAFAEI & BRUNN 1991	OU-YANG & LIOU 1998	BRINZER 2005	GUO ET AL. 2006	KUMAR ET AL. 2006	BRACKEL 2009	NASERI ET AL. 2010	ZHANG ET AL. 2015	WANG & LUO 2016	NIEHUES 2016	POLOTSKI 2017	HU & HU 2018
Logistikleistung	Terminüberschreitungskosten	●	●	●	●	●	○	●	●	●	●	●	●	●
	Terminunterschreitungskosten	○	○	○	○	●	○	●	○	●	●	○	○	○
Logistikkosten	Nichtnutzungskosten	○	○	○	●	○	●	○	●	○	○	●	○	○
	Maschinenkosten	●	●	●	○	○	●	●	●	●	○	●	●	●
	Rüstkosten	●	○	○	●	○	○	●	○	○	●	●	●	●
	Werkzeugkosten	○	○	○	○	○	●	○	○	○	○	○	○	○
	Bestandskosten	●	●	○	●	●	○	●	●	●	○	○	●	●
	Kosten für Intralogistik	○	○	○	○	○	○	●	○	○	○	●	○	○
	Lohnkosten	○	○	●	○	○	●	○	○	○	●	○	○	●
	Energiekosten	○	○	○	○	○	○	○	○	○	○	○	○	○

Abbildung 16: Berücksichtigte Kostenfaktoren in den bisherigen Arbeiten zur kostenbasierten PPS

Auf Seiten der Logistikkosten werden teilweise *Nichtnutzungskosten* modelliert, die als fiktive Strafzahlung im Falle geplanter Stillstandzeiten von Produktionsressourcen beaufschlagt werden und auf diese Weise zur Erhöhung der Maschinenauslastung beitragen sollen. Zudem werden *Maschinen-* und *Lohnkosten* berücksichtigt. Die *Lohnkosten* können sich beispielsweise durch Überstunden erhöhen, die zum Ausgleichen eines vorausgegangenen Maschinenausfalls kurzfristig erforderlich sind (OU-YANG & LIOU 1998). Die *Werkzeugkosten* werden lediglich von KUMAR & RAJOTIA (2006) berücksichtigt, während sie etwa von NIEHUES (2017) aufgrund der vergleichsweise geringen Höhe bewusst vernachlässigt werden. Einen weiteren verbreiten Kostenfaktor bilden die *Rüstkosten*, die maßgeblich von der Bearbeitungsreihenfolge beeinflusst werden. *Kosten für Intralogistik* werden hingegen in vergleichsweise wenigen Fällen als relevant erachtet. *Bestandskosten*, die die Produktionssteuerung insbesondere in Form von Pufferbeständen zwischen einzelnen Bearbeitungsschritten beeinflussen, bilden ebenfalls im Großteil der Arbeiten einen relevanten Kostenfaktor. *Energiekosten* finden hingegen in den untersuchten Ansätzen ausnahmslos keine Berücksichtigung.

Fazit aus den Vorarbeiten zur kostenbasierten PPS

Aus der Analyse der bisherigen Arbeiten zur kostenbasierten PPS wird deutlich, dass kein einheitliches Bild zu den entscheidungsrelevanten Kostenfaktoren existiert. Die Autoren bauen die Zielfunktion aus drei bis sechs einzelnen Kostenfaktoren auf, wobei in den betrachteten Arbeiten insgesamt neun verschiedene Kostenfaktoren berücksichtigt wurden. Eine breite Überschneidung lässt sich lediglich bei den Verzugskosten feststellen, die bis auf eine Ausnahme bei allen betrachteten Kostenmodellen zum Einsatz kommen. Gleichmaßen werden Energiekosten in keiner der untersuchten Arbeiten in die jeweilige Zielfunktion integriert.

3.4 Ableitung des Handlungsbedarfs

In den vorausgegangenen Abschnitten dieses Kapitels wurde der aktuelle Stand der Forschung sowohl zur energieflexiblen als auch zur kostenbasierten Produktionssteuerung aufgezeigt. Um die in Kapitel 1 aufgezeigten Potenziale durch die Produktionssteuerung vollständig ausschöpfen zu können, besteht jedoch weiterhin nachfolgender Handlungsbedarf:

- **Abbilden der Möglichkeiten des Strombezugs aus kurzfristigem Stromhandel, Eigenerzeugung und Stromspeicher:** In den bestehenden Arbeiten zur energieorientierten Produktionssteuerung wurden preisvariable Strompreise sowie in Teilen auch Eigenerzeugungsanlagen und Batteriespeicher berücksichtigt. Die marktseitigen Randbedingungen – wie die erforderliche Vorlaufzeit, um durch kurzfristigen Handel von fluktuierenden Strompreisen profitieren zu können – wurden bisher in keiner der Arbeiten betrachtet. Zudem existiert kein Ansatz, der die drei aufgezeigten Möglichkeiten des Strombezugs gemeinsam in eine Steuerungsstrategie integriert.
- **Überführung der logistischen Zielgrößen in Kosten:** Bisherige Ansätze zur energieorientierten Produktionssteuerung berücksichtigen lediglich in wenigen Fällen zusätzlich Zielgrößen, um diese mit den Energiekosten abzuwägen. Eine ganzheitliche Kostenbetrachtung wird nicht durchgeführt. Demgegenüber existieren auf Seiten der kostenbasierten Produktionssteuerung Vorarbeiten, die eine Überführung der logistischen Zielgrößen in Kosten anstreben. Ein Gesamtbild zu den in diesem Kontext relevanten Kostenfaktoren existiert dabei nicht. Folglich ist eine Weiterentwicklung der kos-

tenbasierten Ansätze erforderlich, um diese auf die energieorientierte Produktionssteuerung übertragen zu können. Auf diese Weise wird ein gesamt-kostenbasiertes Zielsystem geschaffen, das eine fundierte Abwägung energieorientierter Anpassung ermöglicht.

- **Energieorientierte Produktionssteuerung mit integriertem Lastmanagement:** Die Analyse der Vorarbeiten zur energieorientierten Produktionsplanung hat gezeigt, dass derartige Ansätze nicht in der Lage sind, auf kurzfristige Ereignisse zu reagieren und aus diesem Grund eine Kopplung mit einer nachgelagerten, reaktionsschnellen Steuerung erfordern. Die Wirksamkeit eines Lastmanagements lässt sich zudem durch die Integration in die Produktionssteuerung verbessern, die in der Lage ist, kurzfristig zwischen der Lastreduktion durch eine temporäre Abschaltung und deren Folgen für die logistische Zielerfüllung abzuwägen. Insgesamt ist somit eine energieorientierte Produktionssteuerung erforderlich, die das Lastmanagement genauso wie den Energiebezug integriert und gleichzeitig in der Lage ist, auf kurzfristige Ereignisse schnell zu reagieren.

4 System zur energieorientierten und kostenbasierten Produktionssteuerung

In diesem Abschnitt wird zunächst ein geeigneter Steuerungsansatz auf Basis der in Kapitel 2.4 gezeigten möglichen Strategien und der in Kapitel 3 abgeleiteten Anforderungen ausgewählt. Des Weiteren wird ein Überblick über das in den nachfolgenden Kapiteln 5, 6 und 7 hergeleitete System gegeben. Die wesentlichen Systemelemente werden kurz vorgestellt.

4.1 Ableitung der Anforderungen

Basierend auf dem in Abschnitt 3.4 aufgezeigten Handlungsbedarf für eine energieorientierte und kostenbasierte Produktionssteuerung leiten sich für den Steuerungsansatz nachfolgende Anforderungen ab:

- **Vorausschauendes Agieren:** Insbesondere zur Teilnahme am kurzfristigen Stromhandel ist ein vorausschauendes Agieren erforderlich. Die Produktionssteuerung muss beispielsweise in der Lage sein, kurzfristig höhere Energiekosten in Kauf zu nehmen, um zu einem späteren Zeitpunkt noch höhere Strafkosten zu vermeiden.
- **Mehrdimensionale Zielfunktion:** Wie im Handlungsbedarf hergeleitet, ist eine gesamtkostenbasierte Zielfunktion zu verfolgen. Die Produktionssteuerung muss folglich eine mehrdimensionale Zielfunktion mit zugehöriger Gewichtung abbilden und entsprechend lösen können.
- **Hohe Lösungsgüte:** Die Produktionssteuerung hat das zentrale Ziel, die Gesamtkosten der Zielfunktion zu minimieren, die sowohl Energie- als auch Produktionskosten enthält. Zu diesem Zweck ist eine möglichst hohe Lösungsgüte zu erzielen, die nahe an das theoretische Optimum heranreicht.
- **Großer durchsuchbarer Lösungsraum:** Der Lösungsraum für die sich aus dem Handlungsbedarf ergebende Problemstellung ist sehr groß. Der zusätzliche Handel am Strommarkt und die Batteriesteuerung vergrößern die Zahl der kombinatorischen Möglichkeiten in jedem Zeitschritt deutlich. Darüber hinaus sollte aufgrund der erforderlichen Integrationen des Lastmanagements in die Produktionssteuerung sichergestellt werden, dass Produktionsressourcen temporär den Stand-by-Zustand annehmen können.

Dies zieht eine zusätzliche Vergrößerung des Lösungsraums nach sich, den das auszuwählende Steuerungsverfahren beherrschen muss.

- **Hohe Reaktionsfähigkeit:** Um auf die im Handlungsbedarf beschriebenen Störungen reagieren zu können, ist eine hohe Reaktionsfähigkeit des Systems erforderlich. Entscheidungen müssen somit sehr schnell, innerhalb weniger Sekunden getroffen werden können.

4.2 Auswahl des Steuerungsansatzes

In Abschnitt 2.4 wurde die wesentliche Herangehensweise für die Produktionssteuerung dargelegt, wobei die drei wesentlichen Steuerungsstrategien aufgezeigt wurden. In der weiteren Betrachtung werden demnach diese drei Ansätze – reaktiv, prädiktiv-reaktiv und MARL – als Weiterentwicklung von AS berücksichtigt. Den Steuerungsverfahren gegenüber stehen insgesamt fünf zentrale Anforderungen, die in Abschnitt 4.1 als Schlussfolgerung aus der Analyse des aktuellen Stands der Forschung hervorgehen. Eine Gegenüberstellung der drei untersuchten Steuerungsansätze mit den fünf relevanten Kriterien ist in Abbildung 17 dargestellt. Die Bewertung kann auf Basis der in Kapitel 2.4 dargelegten Grundlagen durchgeführt werden.

Anforderungen	Steuerungsansätze			
	Reaktiv	Prädiktiv	MARL	
Vorausschauendes Agieren	○	●	◐	<div style="border: 1px dashed black; padding: 5px;"> <p>● Vollständig erfüllt</p> <p>◐ Überwiegend erfüllt</p> <p>◑ Teilweise erfüllt</p> <p>◒ Geringfügig erfüllt</p> <p>○ nicht erfüllt</p> </div>
Mehrdimensionale Zielfunktion	◑	●	◐	
Hohe Lösungsgüte	◑	●	◑	
Großer durchsuchbarer Lösungsraum	◐	◑	●	
Hohe Reaktionsfähigkeit	●	◑	●	

Legende: MARL: Multi-Agent Reinforcement Learning

Abbildung 17: Bewertung der Ansätze zur Produktionssteuerung hinsichtlich der Erfüllung der Anforderungen

Reaktive Steuerungsansätze überzeugen beim Abgleich mit den Anforderungen besonders hinsichtlich der Reaktionsfähigkeit. Da es sich um heuristische Entscheidungsregeln handelt, kann keine hohe Lösungsgüte gewährleistet werden.

Gleichzeitig ist vorausschauendes Agieren aufgrund der reaktiven Natur der Entscheidungsfindung nicht möglich. Darüber hinaus lassen sich „multikriterielle“ Zielfunktionen nur sehr begrenzt berücksichtigen. Die Größe des abbildbaren Lösungsraums ist jedoch theoretisch unbegrenzt. Voraussetzung ist, eine passende Heuristik zur Problembeschreibung zu finden.

Da *reaktiv-prädiktive Steuerungsansätze* basierend auf mathematischen Optimierungsverfahren einen stringenten Produktionsplan für ein definiertes Zeitintervall berechnen, erreichen diese eine hohe Lösungsgüte, können vorausschauend agieren und eine „multikriterielle“ Zielfunktion verfolgen. Aufgrund des hohen erforderlichen Rechenaufwands ist die Reaktionsfähigkeit jedoch beschränkt. Der Rechenaufwand steigt gleichzeitig auch mit der Größe des Lösungsraums exponentiell an, wodurch etwa die gezielte Einplanung von Stillstandzeiten erschwert wird und dies deutlich zu Lasten der Reaktionsfähigkeit fällt. Um auf stochastische Störungen innerhalb des komplexen Systems reagieren zu können, eignen sich demnach reaktiv-prädiktive Steuerungsansätze trotz der wesentlichen Vorteile im Hinblick auf Lösungsgüte, vorausschauendes Agieren und Mehrdimensionalität der Zielfunktion nicht.

Ansätze basierend auf *MARL* weisen dagegen im Vergleich weniger Schwächen auf. In den bisherigen Anwendungen konnten nah-optimale Lösungen erzielt werden. Gleichzeitig weisen sie eine hohe Reaktionsfähigkeit auf und können dank des Mechanismus der Belohnungsvergabe vorausschauend langfristige Ziele verfolgen. Die Formulierung der Belohnungsfunktion ist ein entscheidender und auch kritischer Schritt bei derartigen Systemen, wodurch die Abbildung einer „multikriteriellen“ Zielfunktion möglich ist. Zudem konnte die Bewältigung von Problemen mit sehr großen Lösungsräumen mittels RL bereits von verschiedenen Autoren demonstriert werden, wie beispielsweise im Kontext von Videospiele (SILVER ET AL. 2016, JADERBERG ET AL. 2019). Gleichzeitig muss an dieser Stelle betont werden, dass Produktionssteuerungsansätze mit MARL aktuell Gegenstand der Forschung sind und im Vergleich zu den anderen beiden Verfahren eine deutlich geringere technologische Reife vorweisen.

Insgesamt geht aus der durchgeführten Bewertung hervor, dass MARL von den betrachteten Verfahren die herangezogenen Kriterien am besten erfüllt, was im Wesentlichen auf die folgenden Aspekte zurückzuführen ist. Im Vergleich zu reaktiven Steuerungsansätzen sind die wesentlichen Vorteile die verbesserte Lösungsgüte sowie die Möglichkeit, vorausschauendes Verhalten zu erreichen und eine komplexe Zielfunktion abzubilden. Auch wenn mit Hilfe eines prädiktiven

Ansatzes eine vergleichsweise höhere Lösungsgüte zu erwarten ist, überwiegen die Vorteile zu Gunsten von MARL im Hinblick auf die Reaktionsfähigkeit und die Größe des lösbaren Problems. Aus diesem Grund wird das im Zuge dieser Arbeit entwickelte Gesamtsystem auf die Lösung mit MARL zugeschnitten.

4.3 Systemübersicht

Aufbauend auf die Verfahrensauswahl wird nachfolgend ein Überblick über das Gesamtsystem gegeben, das in den weiteren Kapiteln erarbeitet wird. Das in Abschnitt 4.1 ausgewählte, dezentrale und agentenbasierte Lösungsverfahren hat maßgeblichen Einfluss auf die Systemarchitektur. Die einzelnen Elemente werden so miteinander verzahnt, dass insgesamt die Energiekosten und alle anderen relevanten Kostenfaktoren der Produktion minimiert werden. Dadurch kann eine Erfüllung der logistischen Zielgrößen bei gleichzeitiger Abwägung mit den Energiekosten gewährleistet werden.

Auf technischer Ebene werden die einzelnen zu steuernden Entitäten – die Produktionsressourcen, der Batteriespeicher und der kurzfristige Stromhandel – durch individuelle Agenten repräsentiert und gesteuert. Diese interagieren mit einer Umgebung, in der die logischen Zusammenhänge des Strommarktes, der Eigenerzeugung, des Batteriespeichers und des Produktionssystems abgebildet sind. Insgesamt erlaubt dieser Systemaufbau eine hohe Reaktionsfähigkeit, da die einzelnen Agenten autonome Entscheidungen auf Basis von Reinforcement Learning treffen. Abgesehen von der Energie können die übrigen Produktionskosten ausschließlich von den Ressourcenagenten beeinflusst werden, **welche wie in anderen Vorarbeiten (z. B. in CRUZ ET AL. 2019) die einzelnen Produktionsressourcen repräsentieren.** Aus diesen Gründen wird dieser Teil der Belohnung ausschließlich den Ressourcenagenten zurückgespielt, während die Energiekosten als globale Belohnung an alle Agenten gleichermaßen verteilt werden (vgl. Abschnitt 7.4). Abbildung 18 gibt einen Überblick über das resultierende Gesamtsystem.

In Abbildung 18 sind zudem die nachfolgenden Kapitel in den Kontext der einzelnen Systembausteine eingeordnet. Mit der Modellierung des Energiebezugs gehen direkt die resultierenden Stromkosten einher. Diese Blöcke werden gemeinsam im *Kapitel 5* erarbeitet, wo die Integration des kurzfristigen Energiebezugs im Mittelpunkt steht. Das Ableiten des kostenbasierten Zielsystems, um alle ausschließlich von den Ressourcenagenten beeinflussbaren Kosten der Produktion abzubilden, wird in *Kapitel 6* behandelt. *Kapitel 7* behandelt das eigentliche Steuerungssystem

auf Basis von MARL. Die Ressourcenagenten sind dabei sehr eng mit dem Produktionssystem verknüpft, wodurch dessen Modellierung ebenfalls Kapitel 7 zugeordnet ist.

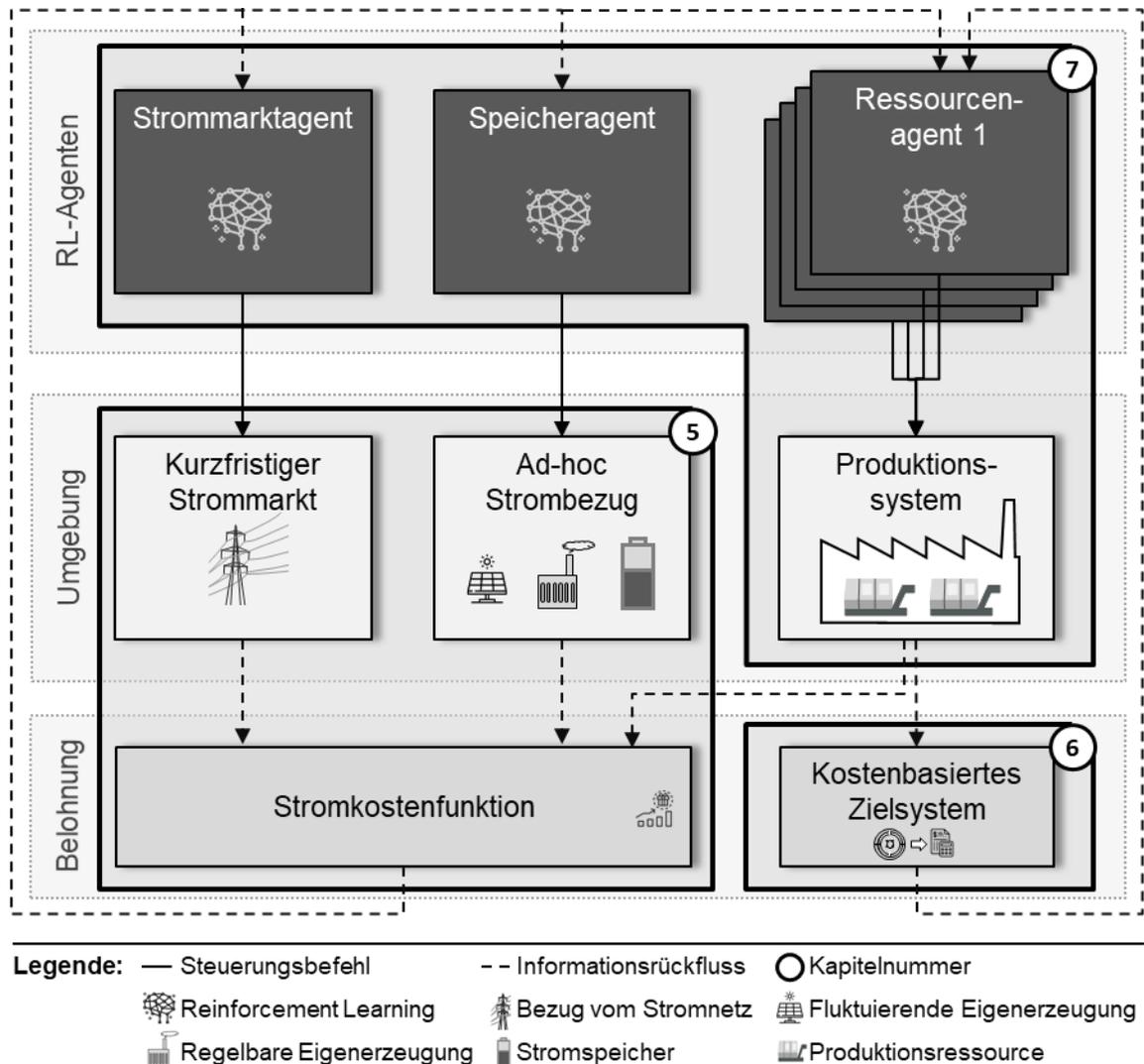


Abbildung 18: Gesamtübersicht über das System (ROESCH ET AL. 2020)

4.4 Systemelemente

Um die in Abschnitt 1.3 dargelegte Zielsetzung zu erreichen, umfasst das in Abbildung 18 skizzierte System sowohl die Produktion und dessen Steuerung als auch die Energieversorgung in Form von Stromhandel, Eigenenerzeugung und Batterie. Dabei ist das System in insgesamt drei funktionelle Ebenen unterteilt, die für das weitere Verständnis nachfolgend beleuchtet werden:

- **RL-Agenten:** Das Gesamtsystem wird auf Basis von drei unterschiedlichen Klassen von Agenten gesteuert, die alle auf RL zurückgreifen. Der Strommarktagent handelt kurzfristig Strom am Intraday-Markt unter Berücksichtigung der Marktpreise und des aktuellen Zustands der Batterie sowie der Produktion. Da der Stromhandel auf Basis von 15-Minuten-Intervallen abgerechnet wird (vgl. Abschnitt 2.2.4), tritt auch der zugehörige Agent alle 15 Minuten in Erscheinung und wird aktiv. Der Speicheragent steuert das Laden beziehungsweise Entladen der Batterie. Die Leistung der Erzeugungsanlagen wird auf dieser Basis regelbasiert festgelegt (vgl. Abschnitt 5.3). Die klassischen Aufgaben der Produktionssteuerung übernehmen mehrere Ressourcenagenten, die jeweils eine Produktionsanlage repräsentieren.
- **Umgebung:** Die Umgebung, mit der die RL-Agenten in der vorliegenden Problemstellung interagieren, setzt sich aus dem Stromhandel, den Eigenerzeugungsanlagen und dem Produktionssystem zusammen. Für das Training wird eine Simulation benötigt, damit die RL-Agenten Erfahrungen durch Interaktion mit der Umgebung sammeln können. Die Modellkomponenten der Umgebung müssen dabei alle existierenden und relevanten technischen Zusammenhänge und Randbedingungen ausreichend abbilden. Nach erfolgreichem Training können die Modelle durch die reale Umgebung ersetzt werden, um diese dann entsprechend auf Basis der in der Simulation gesammelten Erfahrungen zu steuern.
- **Belohnung:** Die RL-Agenten optimieren ihr Verhalten auf Basis der Belohnung, die sie im Zuge der Interaktion mit der Umgebung erhalten. Die Belohnung leitet sich demnach aus der Bewertung der Auswirkungen der ausgewählten Aktionen ab. Um diese quantifizieren zu können, wird zum einen eine Energiekostenfunktion und zum anderen ein kostenbasiertes Zielsystem bereitgestellt. Letzteres dient der Überführung der logistischen Zielgrößen in Kosten, um eine Vergleichbarkeit mit Energiekosten herzustellen.

5 Integration des kurzfristigen Energiebezugs in die Produktionssteuerung

Ziel des vorliegenden Kapitels ist die systematische Integration des kurzfristigen Energiebezugs in die Produktionssteuerung. Zu diesem Zweck werden im ersten Schritt alle vorhandenen Energiebezugsoptionen eingeordnet (Abschnitt 5.1). Anschließend erfolgt in Abschnitt 5.2 die Modellierung der einzelnen Elemente. Anhand dessen wird in Abschnitt 5.3 ein geeignetes Steuerungsschema abgeleitet, das sich in den Ablauf der Produktionssteuerung integrieren lässt.

5.1 Einordnung der steuerungsrelevanten Bezugsoptionen

In diesem Abschnitt sollen die unterschiedlichen Bezugsoptionen im Hinblick auf die Relevanz für die Produktionssteuerung bewertet und kategorisiert werden. Damit wird die Basis für die Auswahl und eine weiterführende Modellierung der relevanten Bezugsoptionen gelegt.

5.1.1 Relevante Energiebezugsoptionen

Eine Zusammenfassung der relevanten Energiebezugsoptionen und deren Eingruppierung ist in Abbildung 19 demonstriert. Die Strombeschaffung gliedert sich in den *OTC-Handel*, die *Terminmärkte* und den *Day-Ahead- und Intraday-Markt*. Eine Sonderrolle bei der Strombeschaffung nimmt der *Netzbezug* ein, da es sich hier um kein vorab gekauftes Produkt handelt. Selbst wenn der Stromverbrauch die bereits gekaufte Strommenge übersteigt, kann ein produzierendes Unternehmen dennoch zusätzlichen Strom aus dem öffentlichen Netz beziehen. Da dem Netzbetreiber hiermit möglicherweise ein zusätzlicher Aufwand zur Stabilisierung des Netzes entsteht, kann dies nachträglich hohe Kosten nach sich ziehen (MÜLLER ET AL. 2009). Auf Seiten der Eigenerzeuger wird in *fluktuierende und regelbare* Eigenerzeugungsanlagen unterschieden (vgl. Abschnitt 2.2.2). Im Bereich der Energiespeicher liegt der Fokus dieser Arbeit auf der Integration von Batterien, aufgrund deren Eignung für das Zwischenspeichern von Strom an Industriestandorten (vgl. Abschnitt 2.2.3).

Bei der Energiebeschaffung werden im weiteren Verlauf dieser Arbeit schwerpunktmäßig die Möglichkeiten des direkten Stromhandels betrachtet. Bei der

Integration des kurzfristigen Energiebezugs in die Produktionssteuerung

Stromversorgung durch ein Energieversorgungsunternehmen überträgt der Stromkunde die Strompreisrisiken weitestgehend auf den Stromversorger (vgl. Abschnitt 2.2.4). Aus Sicht eines produzierenden Unternehmens führt dies in der Regel zu konstanten Strompreisen, wodurch das Potenzial von Energieflexibilität weitestgehend entfällt. Lediglich die kostspieligen Lastspitzen sind im Hinblick auf die Energieflexibilität relevant.

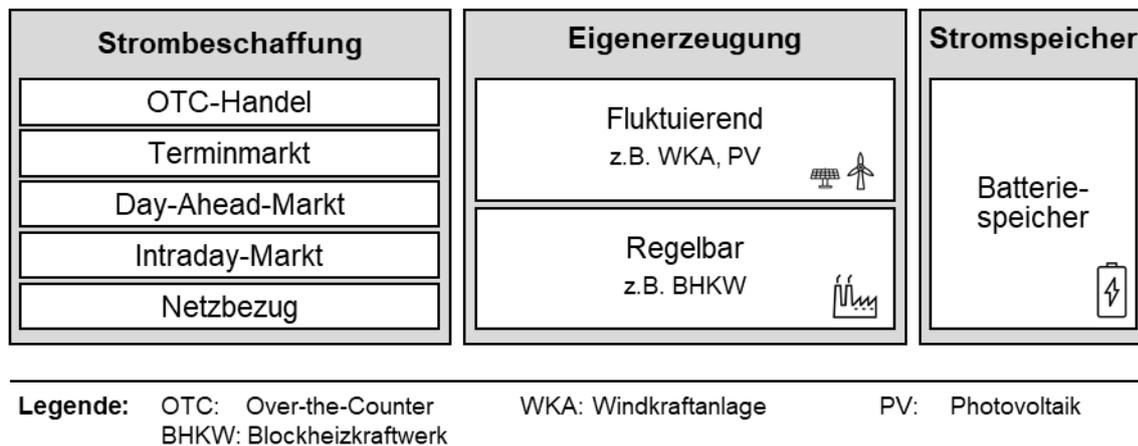


Abbildung 19: Relevante Energiebezugsoptionen

Eine weitere Möglichkeit für Unternehmen, durch flexiblen Energieverbrauch einen monetären Vorteil zu generieren, bietet die Teilnahme am Regelenergiemarkt (vgl. Abschnitt 2.2.4). Das teilnehmende Unternehmen wird im Zuge dessen bereits für das Vorhalten von Regelenergie finanziell entschädigt. Im Falle eines Abrufs der Regelenergie muss die Leistung jedoch zwangsläufig erbracht werden. Andernfalls drohen hohe Strafen (50HERTZ 2019). Eine Abwägung mit anderen Zielgrößen im Falle eines Abrufs ist somit überflüssig. Der für Industrieunternehmen relevante Energiemarkt wird jeweils für den Zeitraum von einer Woche ausgeschrieben. Der Ausschreibungszeitpunkt ist dabei in der Regel am Dienstag der Vorwoche (ACRICOLA ET AL. 2015). Die Entscheidung über die Teilnahme am Regelenergiemarkt liegt damit außerhalb des Zeithorizonts der Produktionssteuerung, die lediglich innerhalb eines Tages agiert (vgl. Abschnitt 2.3.1). Aus diesen Gründen wird die Teilnahme am Regelenergiemarkt im Kontext der vorliegenden energieorientierten Produktionssteuerung nicht weiter betrachtet.

5.1.2 Parameter zur Beschreibung der Energiebezugsoptionen

Die im weiteren Verlauf wesentlichen Parameter zur Charakterisierung von Energiebezugsoptionen werden wie folgt definiert:

- **Vorlaufzeit:** Zeitspanne, die zwischen der Entscheidung über den Abruf einer Energiebezugsoption und der tatsächlichen Stromlieferung verstreicht
- **Abrufbare Energiemenge:** die zu einem Zeitpunkt maximal und minimal abrufbare Energiemenge
- **Energiekosten:** die innerhalb einer Periode für eine bestimmte Energiemenge anfallenden Kosten

Die Energiekosten sowie die maximal und minimal abrufbare Energiemenge können innerhalb jedes Zeitintervalls variieren. Demgegenüber bleibt festzuhalten, dass die Vorlaufzeit, die für den Abruf einer Energiebezugsoption einzuplanen ist, in der Regel als konstant betrachtet werden kann.

5.1.3 Kategorisierung

Für die Entscheidungsfindung der Produktionssteuerung ist die Information essenziell, welche Energiebezugsoptionen zu dem jeweiligen Betrachtungszeitpunkt zur Verfügung stehen. Aus diesem Grund werden die unterschiedlichen Energiebezugsoptionen nachfolgend hinsichtlich ihrer jeweils benötigten *Vorlaufzeit* in die drei Kategorien *ad-hoc Energiebezug*, *kurzfristiger* sowie *mittel- und langfristiger Stromhandel* eingeteilt (siehe Abbildung 20).

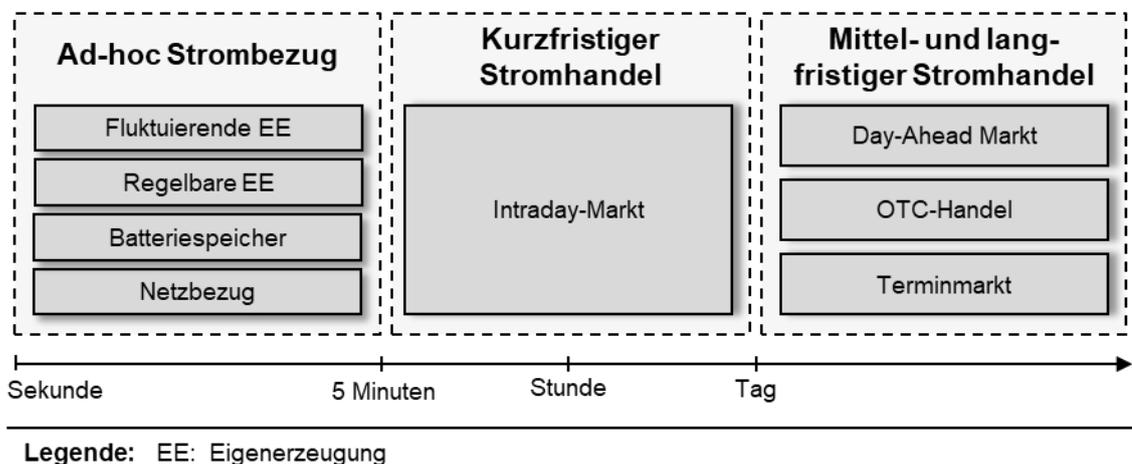


Abbildung 20: Kategorisierung der Energiebezugsoptionen gemäß der Vorlaufzeit

Die *ad-hoc Energiebezugsoptionen* können zum jeweils aktuellen Zeitpunkt ohne einen nennenswerten Zeitversatz angepasst werden. Die erforderliche Reaktionszeit liegt bei PV-Anlagen innerhalb von Sekunden (MACKENSEN ET AL. 2017). Ähnliche Reaktionszeiten erzielen als zweiter verbreiteter Vertreter von *fluktuierenden EE* Windkraftanlagen, die für die Primär- und Sekundärregelung ein-

gesetzt werden können, was zwangsläufig eine Reaktionszeit unterhalb von 30 Sekunden erfordert (GESINO 2011). *Regelbare EE* wie BHKW weisen im Vergleich eine höhere Trägheit auf. Deren Regelbarkeit hängt von der spezifischen technischen Umsetzung ab. Ein Kaltstart bis zur maximalen Erzeugungsleistung dauert hierbei in der Regel deutlich unter fünf Minuten, während Laständerungen zur Laufzeit wesentlich schneller umgesetzt werden können (PILGRAM & KARGER 2014). *Batteriespeicher* kommen aufgrund ihrer geringen Reaktionszeit im Bereich der unterbrechungsfreien Stromversorgung zum Einsatz (ANDRESEN ET AL. 2017). Auch vom Netz kann jederzeit kurzfristig mehr oder weniger Strom als vereinbart bezogen werden. Allerdings ist dies, wie in Abschnitt 2.2.4 dargelegt, mit hohen Kosten verbunden. Als Grundvoraussetzung für die Regelbarkeit von Erzeugungsanlagen und Batterien sind eine entsprechende informationstechnische Vernetzung und die Installation von geeigneter Hardware zur Datenerfassung und Steuerung erforderlich. Im Zuge der vorliegenden Arbeit wird das Vorhandensein dieser Infrastrukturkomponenten als gegeben betrachtet.

Um *kurzfristig Strom* zu handeln, eignet sich der *Intraday-Markt*. Dort kann mit mindestens fünf Minuten Vorlaufzeit Strom für festgeschriebene Viertelstundenintervalle gekauft und verkauft werden (vgl. Abschnitt 2.2.4).

Dem *mittel- und langfristigen Stromhandel* werden der *Day-Ahead-Markt*, der *OTC-Handel* und die *Terminmärkte* zugeordnet. Da deren Vorlaufzeit mehr als einen Tag beträgt, werden diese Strombezugsoptionen im Zuge der Produktionssteuerung, die innerhalb eines Tages agiert, nicht weiter betrachtet.

5.2 Aggregation und Modellierung

Nachfolgend liegt der Fokus auf der Modellierung der unter den gegebenen technischen und organisatorischen Randbedingungen abrufbaren Energiemenge und der in Abhängigkeit davon resultierenden Energiekosten jeder der im Kontext der Produktionssteuerung relevanten Energiebezugsoptionen.

5.2.1 Fluktuierende Eigenerzeugung

Die Erzeugungsleistung der fluktuierenden Eigenerzeugung (FEE) hängt maßgeblich von den aktuellen Wetterbedingungen ab. Die Größe von PV- und Windkraftanlagen wird durch die *absolute Nennleistung* beziffert. Sie gibt die maximal erzielbare Leistung einer Anlage unter optimalen, äußeren Bedingungen an. Im Zuge

der weiteren Modellierung wird zusätzlich der Begriff der *relativen Nennleistung* eingeführt. Er beschreibt die zum aktuellen Zeitpunkt aufgrund der vorliegenden Wetterbedingungen maximal mögliche Erzeugungsleistung. Die relative Nennleistung kann bei optimalen Witterungsbedingungen den Wert der absoluten Nennleistung annehmen, liegt im Normalfall jedoch darunter. Der *regelbare Bereich* von fluktuierenden Erzeugungsanlagen ist somit nach oben hin durch die relative Nennleistung begrenzt. Dagegen ist eine schnelle und flexible Leistungsreduzierung jederzeit möglich (vgl. Abschnitt 5.1.3). Zusammenfassend ergibt sich daher die Möglichkeit zur Regelung von fluktuierenden Anlagen unterhalb der relativen Nennleistung (vgl. Abbildung 21). Durch die Nutzung von Wettervorhersagen und daraus abgeleiteten Erzeugungsprognosen kann zudem die Erzeugungsleistung fluktuierender Anlagen vorhergesagt werden (vgl. Abschnitt 2.2.2).

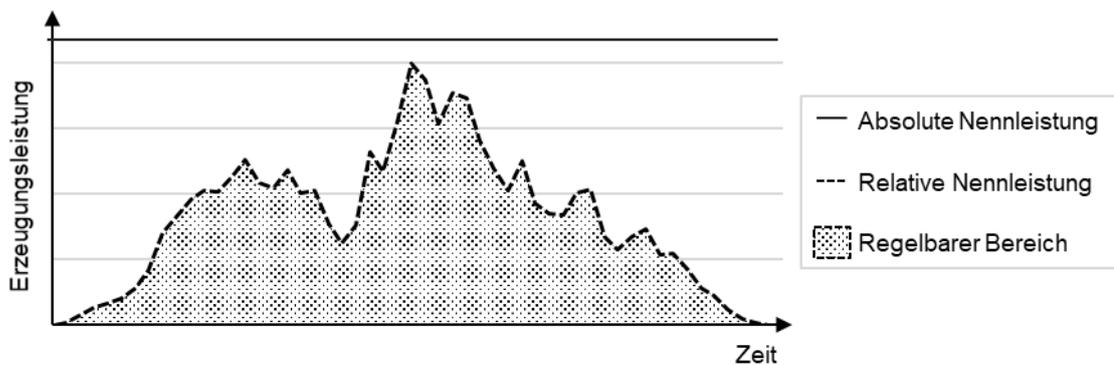


Abbildung 21: Regelbarer Leistungsbereich von fluktuierenden Eigenerzeugungsanlagen

Der Großteil der Stromgestehungskosten für Elektrizität aus erneuerbaren Energien ist auf die erforderlichen Investitionen für die Anlage und deren Anbindung zurückzuführen. Die laufenden Kosten setzen sich hauptsächlich aus Ausgaben für Pacht, Versicherung, Wartung oder Instandhaltung zusammen, deren jährliche Gesamthöhe bei 1,5 bis vier Prozent der Investitionssumme liegt. Im operativen Betrieb können diese Kosten jedoch nicht beeinflusst werden und sind somit nicht entscheidungsrelevant. Da kein Energieträger oder Ähnliches verbraucht wird, hat etwa eine temporäre Abregelung lediglich einen geringfügigen Einfluss auf die Betriebskosten. Aus diesem Grund wird der Strom aus fluktuierenden Erzeugungsanlagen näherungsweise als kostenlos modelliert. (KALTSCHMITT ET AL. 2013)

5.2.2 Regelbare Eigenerzeugung

Im Gegensatz zu Wind- oder PV-Anlagen kann bei regelbaren Eigenerzeugungsanlagen (REE), basierend auf fossilen oder erneuerbaren Energieträgern, jederzeit die maximale Nennleistung abgerufen werden. Gleichzeitig ist auch eine flexible Reduzierung der Leistung möglich. Der Teillastbetrieb ist jedoch auf eine bestimmte Schwelle begrenzt, die von dem jeweiligen Motortypen und der eingesetzten Technik abhängt (vgl. 2.2.2). Bei Gas-Diesel-Aggregaten beispielsweise liegt der niedrigste regelbare Wert bereits bei circa 70 Prozent der Nennleistung (SCHAUMANN & SCHMITZ 2010,). Gas-Otto BHKW können bis zu einer Teillast von 50 Prozent betrieben werden (HOLDACK 2015). Es ist somit eine flexible Anpassung der Erzeugungsleistung bis zu einer minimalen Teillast möglich. Darüber hinaus kann das Kraftwerk lediglich ganz abgeschaltet werden (vgl. Abbildung 22 links). Die Anpassung der Erzeugungsleistung von BHKWs ist nur mit einer sehr geringen Trägheit verbunden. Aus diesem Grund wird, wie bei der Modellierung des Verhaltens von BHKW weit verbreitet, auf das Abbilden von zeitlichen Hoch- und Runterfahrprozessen verzichtet (DVOŘÁK & HAVEL 2012).

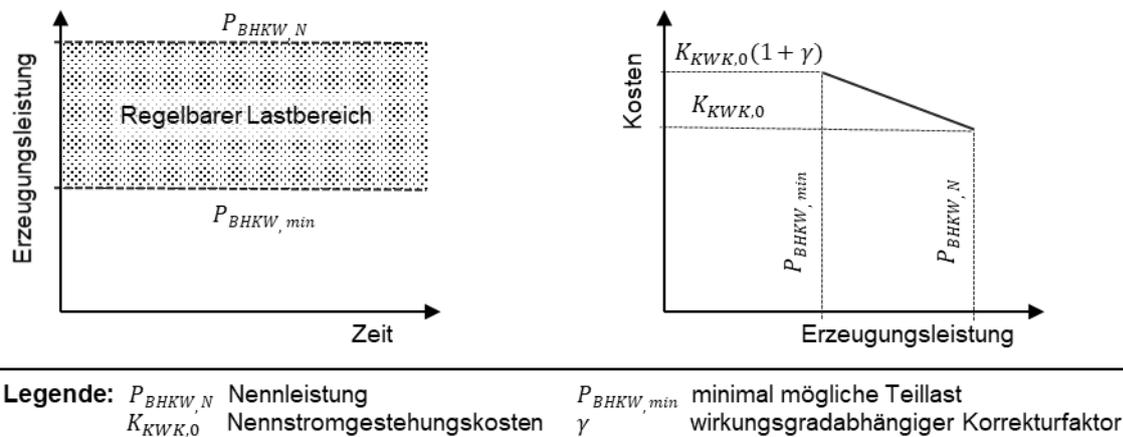


Abbildung 22: Leistungsbereich (links) und Stromgestehungskosten (rechts) von regelbaren Eigenerzeugungsanlagen

Ein großer Teil der Stromgestehungskosten von BHKW ist auf die Investitionskosten zurückzuführen. Diese sind stark von den jährlichen Volllaststunden abhängig, da bei einem hohen Nutzungsgrad die Investitionskosten auf eine größere erzeugte Energiemenge umgelegt werden können. Der zweite wesentliche Faktor für die Zusammensetzung der Stromgestehungskosten ist zudem der Brennstoffverbrauch. Da sich der Wirkungsgrad im Teillastbetrieb im Vergleich zur Volllast verschlechtert, steigt der spezifische Brennstoffverbrauch in Teillast an. Der elektrische Wirkungsgradverlust ist dabei abhängig vom jeweiligen Motorentyp. Bei

Gas-Otto-Motoren fällt dieser im minimalen Teillastbetrieb beispielsweise um etwa vier Prozent im Vergleich zur Volllast, was eine Erhöhung des spezifischen Brennstoffverbrauchs um etwa zehn Prozent nach sich zieht (SCHAUMANN & SCHMITZ 2010). Auf Seiten der Nutzenrechnung von KWK-Anlagen spielen der Erlös für die erzeugte Wärme und der für den erzeugten Strom eine Rolle. Der Fokus dieser Arbeit liegt ausschließlich auf der Betrachtung der elektrischen Energie, wodurch die Wechselwirkungen der Stromgestehungskosten mit dem Wärmebedarf nicht abgebildet werden. (WÜNSCH ET AL. 2014, HOLDACK 2015)

Demnach lässt sich festhalten, dass die Stromgestehungskosten bei BHKW von mehreren Einflussfaktoren abhängig sind: etwa von der Anzahl jährlicher Volllaststunden, dem Wärmebedarf und -erzeugung sowie dem aktuellen Betriebspunkt. Die kurzfristige Steuerung kann die jährlichen Volllaststunden aufgrund des begrenzten Betrachtungshorizonts kaum beeinflussen. Vielmehr wird auf eine Schätzung der spezifischen Stromgestehungskosten zurückgegriffen, die basierend auf den bisherigen, statischen Verfahren im Kontext von BHKW ermittelt werden. Auf Basis von Erfahrungswerten können dabei die jährlichen Volllaststunden abgeschätzt und daraus, zusammen mit dem Brennstoffverbrauch, die Nennstromgestehungskosten per Annuitätenmethode abgeleitet werden (vgl. VDI-RICHTLINIE 2067). Um dem in Teillast sinkenden Wirkungsgrad Rechnung zu tragen, wird zur spezifischen Kostenberechnung näherungsweise ein linearer Korrekturfaktor δ eingeführt (vgl. Abbildung 22 rechts):

$$K_{BHKW}(t) = k_{BHKW,0} \cdot \left(1 + \gamma \cdot \frac{1 - \frac{P_{BHKW}(t)}{P_{BHKW,N}}}{1 - \frac{P_{BHKW,min}}{P_{BHKW,N}}} \right) \cdot P_{BHKW}(t) \cdot \Delta t \quad (4)$$

$K_{BHKW}(t)$	Stromkosten des BHKW zum Zeitpunkt t
$k_{BHKW,0}$	Nennstromgestehungskosten des BHKW
γ	wirkungsgradabhängiger Korrekturfaktor
$P_{BHKW}(t)$	Leistung der BHKW-Anlage zum Zeitpunkt t
$P_{BHKW,N}$	Nennleistung der BHKW-Anlagen
$P_{BHKW,min}$	minimal mögliche Teillast der BHKW-Anlage
Δt	Länge des Betrachtungsintervalls

5.2.3 Stromspeicher

Um Speicher im Gesamtsystem abbilden zu können, sind Informationen zu der maximalen und minimalen Speicherleistung, der Kapazität, dem Wirkungsgrad sowie der Speicheralterung und anderen technischen Randbedingungen erforderlich. Die Modellierung der technischen Parameter hängt maßgeblich von der jeweiligen Speichertechnologie ab, wodurch eine allgemeingültige Formulierung nicht zielführend erscheint. Im Weiteren wird deshalb ein Batteriespeicher auf Basis von Li-Ionen modelliert, da diese Technologie aufgrund der aktuellen Entwicklungen von hoher Relevanz für die stationäre Anwendung in der Industrie ist (vgl. Abschnitt 2.2.3). Davon unabhängig können in das vorliegende System auch andere Speichertechnologien integriert werden. Einzige Voraussetzung hierfür ist ein geeignetes Modell, das die relevanten Parameter abbilden kann.

Der maximale Lade- beziehungsweise Entladestrom einer Batterie wird durch die C-Rate festgelegt (vgl. Abschnitt 2.2.3). Die Lade- und Entladeleistung P_{batt} kann somit nur innerhalb der dadurch vorgegebenen Grenzen liegen. Aufgrund des zeitlichen Diskretisierens in der vorliegenden Arbeit wird näherungsweise angenommen, dass die Lade-/Entladeleistung in jedem Zeitschritt konstant ist. Der aktuelle Ladezustand (engl. State of Charge, SOC) für jeden Zeitschritt der Länge Δt lässt sich auf Basis des vereinfachten Modells von MENDIL ET AL. (2018) wie folgt ermitteln:

$$SOC(t + \Delta t) = SOC(t) + \eta \int_t^{t+\Delta t} \frac{P_{batt}(u)}{C_N V(u)} du \quad (5)$$

$SOC(t)$ Ladezustand zum Zeitpunkt t

η Wirkungsgrad beim Laden/Entladen der Batterie

$P_{batt}(t)$ Lade-/Entladeleistung der Batterie zum Zeitpunkt t

$V(t)$ Batteriespannung zum Zeitpunkt t

C_N Nennkapazität der Batterie

Die Lebensdauer eines Batteriespeichers hängt maßgeblich von den Zyklen ab, denen der Speicher ausgesetzt ist. Im weiteren Verlauf wird das Degradationsmodell nach XU ET AL. (2018) herangezogen, um die Auswirkungen der Speicherbetriebsstrategie auf die Alterung modellieren zu können. Das verwendete Modell ist dabei auf die Nutzung von Li-Ionen-Batterien zugeschnitten. Falls andere Zelltypen verwendet werden, ist eine entsprechende Anpassung erforderlich. Durch den semi-empirischen Charakter ist lediglich eine begrenzte Anzahl an experimentellen Daten erforderlich. Die Anwendung ist im operativen Betrieb für beliebige

Zyklen und Zyklen-Wechsel möglich. Gleichzeitig können die einzelnen Degradationsmechanismen mit ausreichender Genauigkeit dargestellt werden. Prinzipiell können auch andere Alterungsmodelle in das Gesamtsystem integriert werden, sofern sie eine zyklusspezifische Berechnung ermöglichen. Der Alterungszustand einer Batterie wird im vorliegenden Kontext durch den Parameter L beschrieben, der zwischen 0 und 1 liegt, wobei für eine neue Batterie $L = 0$ gilt. Die Alterung L wird wie folgt berechnet:

$$L = 1 - \alpha_{sei} e^{-\beta_{sei} \sum_i^N f_{d,i}} - (1 - \alpha_{sei}) e^{-\sum_i^N f_{d,i}} \quad (6)$$

- L Alterungszustand der Batterie
- $\alpha_{sei}, \beta_{sei}$ Koeffizienten für das Ausbilden der Passivierungsschicht der Anode
- N Zyklenzahl
- $f_{d,i}$ Linearisierte Degradation pro Zyklus i

Dabei werden zur Ermittlung der Zyklen spezifischen Degradation $f_{d,i}$ die einzelnen Stressfaktoren Temperatur, Ladezustand (SOC), Entladetief (DOD) und die Zeit zunächst getrennt modelliert und anschließend zu einem gemeinsamen Modell zur Bewertung der resultierenden Batteriealterung zusammengefasst. Somit ergibt sich eine Degradation $f_{d,i}$ für einen Zyklus i mit:

$$f_{d,i}(t_i, \delta_i, \sigma_i, T) = [S_t(t_i) + S_\delta(\delta_i)] S_\sigma(\sigma_i) S_T(T) \quad (7)$$

- $f_{d,i}$ Degradation für den Zyklus i
- $S_t(t_i)$ Zeitliches Stressmodell
- $S_\delta(\delta_i)$ DOD-Stressmodell
- $S_\sigma(\sigma_i)$ SOC-Stressmodell
- $S_T(T)$ Temperatur-Stressmodell

Die SOC bedingte Degradation ist auf Mikrorisse durch interne mechanische Belastungen zurückzuführen und kann nach MILLNER (2010) berechnet werden:

$$S_\sigma(\sigma) = e^{k_\sigma(\sigma - \sigma_{ref})} \quad (8)$$

- $S_\sigma(\sigma)$ SOC-Stress Koeffizient
- k_σ SOC-Stress Koeffizient
- σ_{ref} Referenz SOC-Wert, in der Regel zwischen 0,4 und 0,5
- σ Durchschnittlicher SOC im aktuellen Zyklus

Integration des kurzfristigen Energiebezugs in die Produktionssteuerung

Ein weiterer wichtiger Faktor ist die Degradation in Abhängigkeit des DoD, die in einem nichtlinearen Zusammenhang stehen (ORTEGA-VAZQUEZ 2014). Da bisher keine ausreichenden experimentellen Studien vorhanden sind, wird an dieser Stelle ein empirisches Modell verwendet (XU ET AL. 2018):

$$S_{\delta}(\delta) = \frac{1}{k_{\delta 1} \delta^{k_{\delta 2}} + k_{\delta 3}} \quad (9)$$

$S_{\delta}(\delta)$	DOD-Stressmodell
$k_{\delta 1}, k_{\delta 2}, k_{\delta 3}$	Empirische Koeffizienten für die DoD bedingte Alterung
δ	DoD des aktuellen Zyklus

Für die Degradation über die Zeit wird näherungsweise ein linearer Verlauf angenommen:

$$S_t(t) = k_t t \quad (10)$$

$S_t(t_i)$	Zeitliches Stressmodell
k_t	Empirischer Koeffizient für die kalendarische Alterung
t	Zeitdauer des aktuellen Zyklus

Die temperaturbedingte Alterung wird nach LIAW ET AL. (2003) von der Arrhenius-Gleichung abgeleitet. Auf diese Weise können die Angaben zu Alterungsversuchen des Herstellers, die bei einer Referenztemperatur T_{ref} durchgeführt wurden, auf eine beliebige Zelltemperatur übertragen werden:

$$S_T(T) = e^{k_T(T-T_{ref}) \frac{T_{ref}}{TP}} \quad (11)$$

$S_T(T)$	Temperaturstressmodell
TP	durchschnittliche Zelltemperatur
T_{ref}	Referenztemperatur
k_T	Empirischer Koeffizient für die temperaturbedingte Alterung

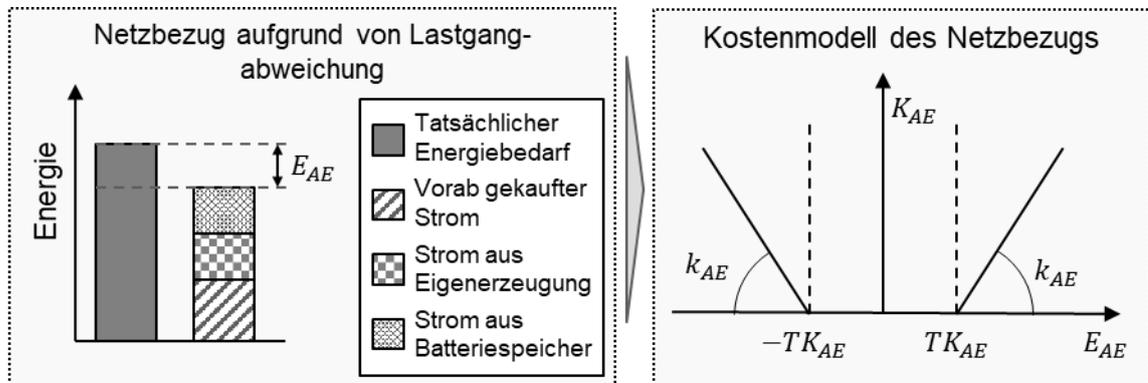
Die zur Modellierung erforderlichen Parameter α_{sei} , β_{sei} , $k_{\delta 1}$, $k_{\delta 2}$, $k_{\delta 3}$, k_{σ} , σ_{ref} , k_t und k_T können entsprechend dem in XU ET AL. (2018) vorgestellten Ansatz aus den Herstellerangaben einer Batterie ermittelt werden. Durch die beschriebene Modellierung ist eine Berechnung der Degradation des Batteriespeichers für jeden Zyklus ermöglicht. Auf diese Weise können die Investitionskosten einer Batterie in Abhängigkeit der tatsächlich verursachten Alterung auf jeden Zyklus umgelegt werden:

$$K_{Bat,i} = \Delta L_i \cdot INV_{Bat} \quad (12)$$

- $K_{Bat,i}$ Resultierende Alterungskosten der Batterie pro Zyklus i
- ΔL_i Veränderung der Alterung im Zyklus i
- INV_{Bat} Investitionskosten der Batterie

5.2.4 Netzbezug

Falls für ein Zeitintervall nicht genügend oder zu viel Energie aus Eigenerzeugung und bereits beschafftem Strom zur Verfügung steht, kann diese Differenz kurzfristig und ohne Zeitverzögerung vom öffentlichen Netz bezogen beziehungsweise eingespeist werden. Der Bezugs- und Einspeisemenge dieser Ausgleichsenergie sind theoretisch keine Grenzen gesetzt. Da der Energieversorger unter Umständen diese Lastgangabweichung ausgleichen muss, um die Netzstabilität zu gewährleisten, können hohe Kosten anfallen, falls die Abweichung eine bestimmte Toleranzgrenze TK_{LA} über- oder unterschreitet (vgl. Abschnitt 2.2.4). Auch wenn die exakte Höhe dieser Kosten erst im Nachgang beziffert wird, so können diese dennoch bereits vorab in Form des Strafkostensatzes grob abgeschätzt und modelliert werden (vgl. Abbildung 23).



Legende: E_{AE} vom Netz bezogene Ausgleichsenergiemenge
 K_{AE} Kosten für vom Netz bezogene Ausgleichsenergie
 TK_{AE} Toleranzkorridor für Lastgangabweichung

Abbildung 23: Ursachen (links) und Kosten (rechts) von kurzfristigem Netzbezug innerhalb eines Zeitintervalls

Integration des kurzfristigen Energiebezugs in die Produktionssteuerung

Die Stromkosten für Netzbezug setzen sich demnach folgendermaßen zusammen:

$$K_{AE}(t) = \begin{cases} E_{AE}(t) \cdot k_{AE} & E_{AE}(t) > TK_{AE} \\ 0 & TK_{AE} > E_{AE}(t) > -TK_{AE} \\ -E_{AE}(t) \cdot k_{AE} & E_{AE}(t) < -TK_{AE} \end{cases} \quad (13)$$

$K_{AE}(t)$ Kosten für vom Netz bezogene Ausgleichsenergie in der Periode t

$E_{AE}(t)$ vom Netz bezogene Ausgleichsenergiemenge in der Periode t

k_{AE} Strafkostensatz für Abweichung

TK_{AE} Toleranzkorridor für Lastgangabweichung

Daneben können im Falle einer Vollstromversorgung durch Lastspitzen hohe Zusatzkosten entstehen. Da der Kunde eines Energieversorgers einen vorab vereinbarten Leistungspreis k_{LS} für jedes Kilowatt der höchsten Lastspitze innerhalb eines Abrechnungszeitraums bezahlen muss, lässt sich dieser Kostenfaktor genau beziffern. Als Ausgangspunkt dient dazu die im aktuellen Abrechnungszeitraum bisherige Lastspitze LS_{max} , da ausschließlich Lastspitzen, die darüber hinausgehen, weitere Kosten verursachen. Demnach lassen sich die Kosten für Lastspitzen $K_{LS}(t)$ folgendermaßen berechnen:

$$K_{LS}(t) = \begin{cases} k_{LS} \cdot (P_{ES}(t) - LS_{max}) & , P_{ES}(t) - LS_{max} > 0 \\ 0 & , P_{ES}(t) - LS_{max} \leq 0 \end{cases} \quad (14)$$

$K_{LS}(t)$ Kosten für Lastspitze im 15-Minuten-Intervall t

k_{LS} Mit Energieversorger vereinbarter Leistungspreis für Lastspitzen

$P_{ES}(t)$ Durchschnittlich vom externen Stromnetz bezogene Leistung im Intervall t

LS_{max} aktuelle, kostenwirksame Lastspitze

5.2.5 Stromhandel am Intraday-Markt

Am Intraday-Markt kann mit einer Vorlaufzeit von mindestens fünf Minuten Strom für einzelne 15-Minuten-Intervalle gekauft werden (vgl. Abschnitt 2.2.4). Zum Zeitpunkt des Kaufs werden neben dem Strompreis zusätzlich pauschale und anteilige Ordergebühren fällig. Gesetzliche Abgaben, wie etwa die EEG-Umlage oder Steuern, werden erst zum Zeitpunkt des eigentlichen Verbrauchs des Stroms kostenwirksam (DIHK 2019). Da es sich um tatsächlich anfallende Kosten handelt, die den Stromhandel maßgeblich verteuern, werden diese für den Marktagenten als entscheidungsrelevant erachtet und somit direkt bei der Kaufabwicklung beaufschlagt.

Beim kurzfristigen Verkauf von Energie fallen die identischen Gebühren wie beim Kauf an. Zu beachten ist, dass durch den Verkauf von Strom nicht nur ein gewisser Preis Erlöst werden kann. Gleichzeitig sind für die verkaufte Strommenge auch keine Steuern und Umlagen zu entrichten, da diese nicht tatsächlich verbraucht wird. Diese Randbedingung muss bei der Kostenbewertung des Stromhandels berücksichtigt werden. Angenommen, für ein Zeitintervall wurde vorab am Day-Ahead-Markt eine bestimmte Energiemenge $E_I(t)$ beschafft. Die dafür angefallenen Kosten beinhalten bereits die für den Verbrauch der Energie anfallenden Steuern und Umlagen k_{SU} . Wird diese Energiemenge kurz vor dem geplanten Verbrauch verkauft, so werden die bereits einkalkulierten Kosten für Steuern und Umlagen zusätzlich gespart. Wurde die verkaufte Strommenge nicht von extern beschafft, sondern kommt etwa aus Eigenerzeugungsanlagen, so wurden diese Kosten vorab nicht entrichtet und können somit auch nicht bilanziert werden.

Insgesamt ergibt sich mit dieser Fallunterscheidung folgende Berechnung der Stromkosten für den Handel am Intraday-Markt:

$$K_I(t) = \begin{cases} \frac{K_{P,I}}{z} + E_I(t) \cdot (k_{I,t} + k_{rel,I} + k_{SU}) & , E_I(t) > 0 \\ \frac{K_{P,I}}{z} - E_I(t) \cdot (k_{I,t} + k_{rel,I} - k_{SU}) & , -E_V(t) < E_I(t) < 0 \\ \frac{K_{P,I}}{z} - E_I(t) \cdot (k_{I,t} + k_{rel,I}) + E_V(t) \cdot k_{SU} & , E_I(t) < -E_V(t) \end{cases} \quad (15)$$

- $K_I(t)$ Kosten für Strom aus Intraday-Handel-Zeitintervall t
- $K_{P,I}$ Gebührenpauschale pro Order für Intraday-Handel
- z Anzahl der für die aktuelle Order betroffenen Zeitintervalle
- $E_I(t)$ Vom Intraday-Markt bezogene Strommenge in Zeitintervall t
- $E_V(t)$ Vorab beschaffte Energiemenge für das Zeitintervall t
- $k_{I,t}$ Preis am Intraday-Markt im Zeitintervall t
- $k_{rel,I}$ Anteilige Gebühr
- k_{SU} Kostensatz für Steuern und Umlagen

5.2.6 Zusammenfassende Betrachtung des Energiebezugs

Aus der bisherigen Betrachtung zu den Energiebezugsoptionen eines Unternehmens lassen sich folgende Punkte festhalten:

- **Fluktuierende Eigenerzeuger:** Strom aus fluktuierenden Eigenerzeugungsanlagen schlägt in der kurzfristigen Betrachtung als kostenfrei zu Buche. Eine Abregelung der Erzeugungsleistung von fluktuierenden EE sollte daher nur in Ausnahmefällen erfolgen.
- **Regelbare Eigenerzeuger:** Die Leistung regelbarer Erzeuger kann kurzfristig flexibel angepasst werden. Ein Teillastbetreiber ist jedoch nur bis zu einer unteren Grenze möglich. Die Stromgestehungskosten nehmen im Teillastbetrieb zu.
- **Batteriespeicher:** Die Batterie bietet ein sehr hohes und kurzfristiges Flexibilisierungspotenzial. Sie eignet sich, um temporäre Mengen- und Preisunterschiede auszugleichen.
- **Netzbezug:** Liegt die Differenz zwischen der gekauften und der tatsächlich aus dem Netz bezogenen Strommenge außerhalb eines bestimmten Toleranzkorridors, ist mit hohen Zusatzkosten zu rechnen. Eine derartige Lastabweichung sollte, wenn möglich, vermieden werden. Dieselbe Empfehlung lässt sich im Falle einer Vollstromversorgung hinsichtlich der Vermeidung von Lastspitzen aussprechen.
- **Intraday-Handel:** Durch den Handel am Intraday-Markt können kurzfristige Lastabweichungen ausgeglichen und Preisschwankungen ausgenutzt werden. Dabei sind neben dem Strompreis auch die Kosten für Gebühren sowie Steuern und Umlagen zu beachten.
- **Mittel- und langfristigen Stromhandel:** Strom aus dem mittel- und langfristigen Stromhandel wird vorab und somit außerhalb des Entscheidungshorizonts einer Produktionssteuerung gekauft. Die bereits bezahlten Preise sind somit im Kontext der Produktionssteuerung nicht mehr entscheidungsrelevant.

5.3 Ableitung des Steuerungsschemas

Durch die Kategorisierung der Strombezugsoptionen anhand der Vorlaufzeit und der verschiedenen Ansätze zur Modellierung (vgl. Abschnitt 5.2) wurden wesentliche Unterschiede zwischen Teilen der Energiebezugsoptionen deutlich. Um dennoch eine integrierte Steuerung aller Energiebezugsoptionen zu ermöglichen, werden diese zur Entwicklung des Steuerungskonzeptes im weiteren Verlauf dieses Kapitels in die zwei Teilsysteme *Ad-hoc Strombezug* und *kurzfristiger Stromhandel* unterteilt (vgl. Abbildung 20).

5.3.1 Ad-hoc Strombezug

Im Zuge des Ad-hoc Strombezugs steht die kostenoptimale Steuerung der Eigen-erzeugungsanlagen, der Batterie und des Netzbezugs im Vordergrund. Dabei steht eine bestimmte Menge an vorab gekauftem Strom zur Verfügung. Um das komplexe Teilsystem steuern zu können, wird eine Reduzierung des Aktionsraums angestrebt. Insgesamt lässt sich das in Abbildung 24 dargestellte Steuerungsschema ableiten, das aus neun einzelnen Schritten besteht. Die entwickelte Steuerungsroutine wird dabei alle drei Minuten durchlaufen, um auf diese Weise innerhalb der abrechnungsrelevanten 15 Minuten den Stromverbrauch korrigieren zu können. Hiermit wird jedes 15-minütige Hauptintervall in fünf gleiche Teilintervalle zerlegt. Eine wichtige Prämisse ist zudem, dass zum Startzeitpunkt der Routine bereits alle Produktionsressourcen ihre Aktionen für das anstehende Teilintervall ausgewählt haben (vgl. Abschnitt 7.1). Somit steht die für das Teilintervall benötigte Energiemenge bereits weitestgehend fest.

Im ersten Schritt der Steuerungsroutine wird zunächst die *für das aktuelle Hauptintervall bereits beschaffte Energiemenge (Operation Nummer 1)* abgerufen, die aus den Daten des Stromhandels hervorgeht. Im nächsten Schritt kann, darauf aufbauend, die *im aktuellen Teilintervall vorhandene Strommenge (2)* aus vorab gekauftem Strom berechnet werden. Dazu wird die im Hauptintervall bisher bereits verbrauchte Menge des beschafften Stroms herangezogen und die verbliebene Energie gleichmäßig auf die ausstehenden Teilintervalle des Hauptintervalls aufgeteilt. Da die ausgewählten Aktionen der Produktionsressourcen bereits bekannt sind, lässt sich gleichzeitig auch der *Energiebedarf für das aktuelle Teilintervall ableiten (3)*. Um abschließend die insgesamt verfügbare Strommenge zu erhalten, wird die für das anstehende Teilintervall prognostizierte, *verfügbare Erzeugungsleistung der FEE addiert (4)*.

Die bisherigen Schritte (1)-(4) verlaufen rein regelbasiert auf Basis von einfachen mathematischen Rechenoperationen. Zur Ermittlung der *Lade- beziehungsweise Entladeleistung des Speichers (5)* wird nun auf einen RL-Agenten zurückgegriffen. Auf den Zustands- und Aktionsraum der Entscheidung wird in Abschnitt 7.3.3 näher eingegangen. Nachdem die Lade- oder Entladeleistung der Batterie ermittelt ist, wird im nächsten Schritt die *Erzeugungsleistung der REE angepasst (6)*.

Integration des kurzfristigen Energiebezugs in die Produktionssteuerung

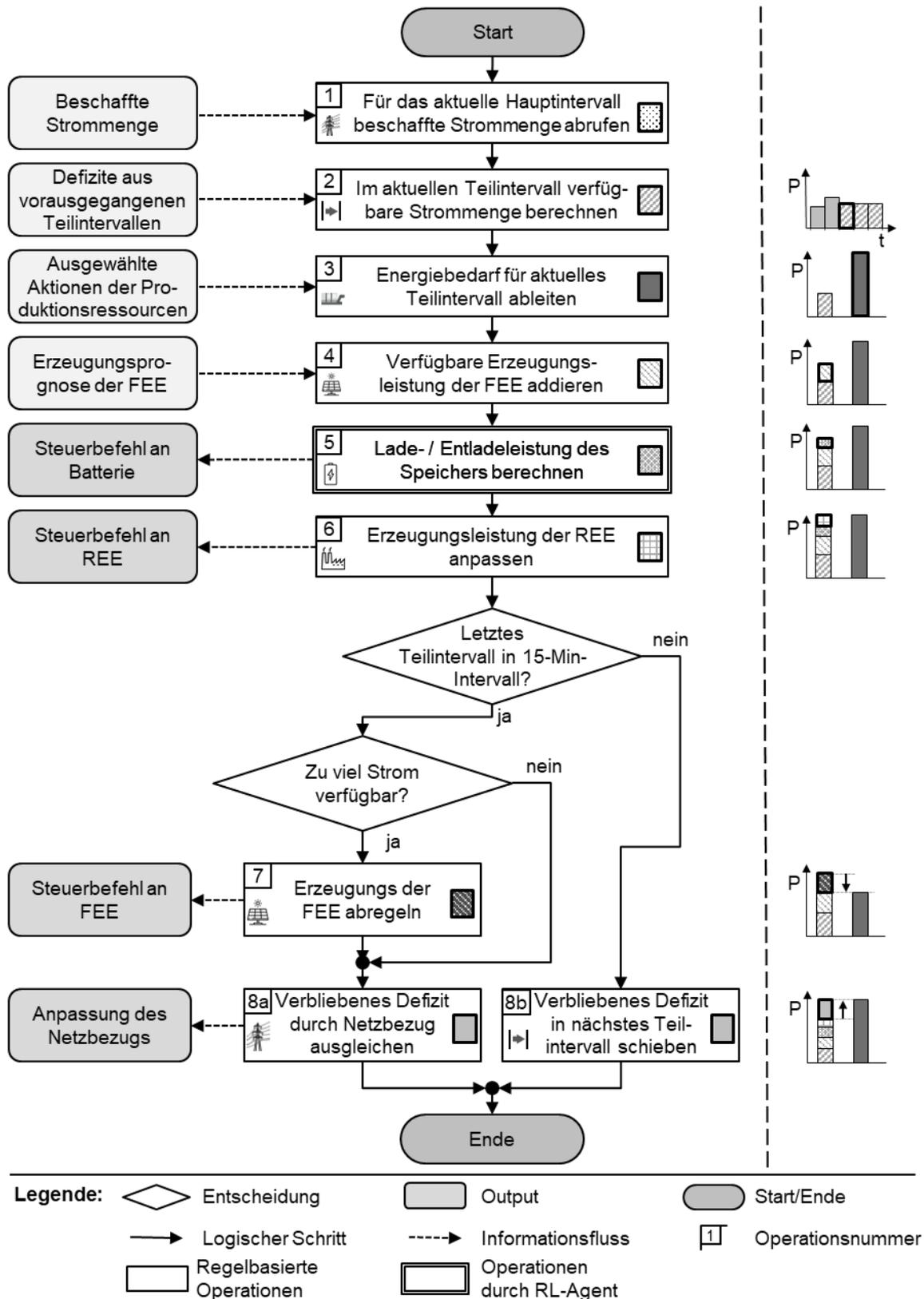


Abbildung 24: Steuerungsschema des Ad-hoc Strombezugs

Die REE sollen idealerweise das bestehende Defizit zwischen dem Energiebedarf und der verfügbaren Energiemenge inklusive der Energie aus der Batterie schließen. Ist das bestehende Defizit zu groß, so dass es nicht vollständig durch die Erzeugungsleistung der REE geschlossen werden kann, wird dennoch die maximal mögliche Erzeugungsleistung angestrebt. Liegt das Defizit umgekehrt unterhalb des minimal möglichen Teillastbetriebs des REE, so werden diese vollständig abgeschaltet.

Falls weiterhin ein Defizit besteht, wird nun in zwei Fälle unterschieden. Da erst nach dem jeweiligen Ende eines Hauptintervalls die Abrechnung erfolgt, zieht ausschließlich ein Defizit Kosten nach sich, das sich aus dem Durchschnitt aller fünf Teilintervalle eines Hauptintervalls ergibt. Wird beispielsweise die verfügbare Strommenge im dritten Teilintervall nicht vollständig verbraucht, so steht der überschüssige Strom für das vierte und fünfte Teilintervall zur Verfügung. Handelt es sich um das letzte Teilintervall innerhalb eines Hauptintervalls, wird dagegen zunächst im Falle einer zu hohen Stromverfügbarkeit das Defizit durch die *Abregelung der Erzeugungsleistung aus FEE (7)* ausgeglichen. Ist auch bei vollständiger Abschaltung der FEE zu viel Strom verfügbar oder ist umgekehrt zu wenig Strom vorhanden, wird dieses *verbliebene Defizit schließlich aus dem öffentlichen Netz ausgeglichen (8a)*. Das ist mit einer kostspieligen Lastgangabweichung verbunden (vgl. Abschnitt 5.2.4). Folgen noch weitere Teilintervalle innerhalb des aktuellen 15-Minuten-Intervalls, so kann das nach der Anpassung der FEE *verbliebene Defizit in das nächste Teilintervall geschoben werden (8b)*, indem rein rechnerisch die verbleibende Menge des beschafften Stroms um die Höhe des Defizits verändert wird.

Das vorgestellte schematische Vorgehen schränkt im Kontext eines Lösungsansatzes mit MARL wesentlich die möglichen Systemfreiheitsgrade ein. Der Aktionsraum des Speicheragenten zur Steuerung aller Elemente des Ad-hoc Strombezugs wird auf das Laden und Entladen der Batterie reduziert. Auf diese Weise kann die Komplexität des Systems deutlich verringert werden. Die übrigen Steuergrößen werden auf Basis von einfachen und statischen Regeln festgelegt.

Der RL-Agent zur Steuerung der Batterie lernt im Zuge des Trainings die Mechanismen und Auswirkungen der Regeln anhand von zahllosen Simulationsläufen. Er kann durch dieses Erfahrungswissen seine Strategie auch im Hinblick auf die implementierten Regeln optimieren. Die Grundvoraussetzung dafür ist allerdings, dass die für die Anwendung der Regeln relevanten Einflussgrößen im Zustandsraum des Agenten übergeben werden. Auf diese Weise können alle Schritte des

Steuerungsschemas implizit in der Entscheidung des Batterie-Agenten berücksichtigt werden. Zudem bietet das abgeleitete Steuerungsschema ausreichend Flexibilität, um Abweichungen zwischen dem tatsächlichen und dem auf Basis der ausgewählten Aktionen der Ressourcenagenten geschätzten Energieverbrauch auszugleichen. Als Puffer dienen hier insbesondere die regelbasierten Steuerungen der Eigenerzeugungsanlagen und des Netzbezugs. Bei Letzterem zieht eine Abweichung innerhalb eines bestimmten Toleranzbandes keine Kosten nach sich und kann somit in Kauf genommen werden (vgl. Abschnitt 5.2.4).

5.3.2 Kurzfristiger Stromhandel

Im Gegensatz zum Ad-hoc Strombezug besteht der kurzfristige Stromhandel aus nur einer Operation, und zwar dem Handel von Strom am Intraday-Markt. Dieser ist mit großer Preisunsicherheit verbunden (DIHK 2019). Da der Strom kontinuierlich ab 15:00 Uhr des Vortags bis fünf Minuten vor eigentlicher Lieferung gehandelt wird und sich der Preis während dieser Zeit dynamisch ändert, kann beispielsweise Strom für ein 15-Minuten-Intervall um 9:00 Uhr am Vormittag zu anderen Börsenpreisen als am Mittag gekauft werden (BERTSCH ET AL. 2017). Die Modellierung dieser marktseitigen Unsicherheit ist aufgrund der vielfältigen Einflussfaktoren sowie der Abhängigkeiten von anderen Marktteilnehmern komplex und wird im Kontext der vorliegenden Arbeit nicht modelliert. Vielmehr soll der Intraday-Handel dazu benutzt werden, kurzfristige Abweichungen zwischen Energieangebot und -verbrauch auszugleichen und Preisextreme auszunutzen.

Aus diesem Grund ist der Stromhandel im Kontext der vorliegenden Arbeit auf das jeweils nächste Zeitintervall begrenzt. Ein Kauf oder Verkauf von Strom über diesen begrenzten Zeithorizont hinaus wird nicht berücksichtigt (vgl. Abbildung 25). Obwohl theoretisch lediglich fünf Minuten Vorlaufzeit zum Stromhandel am Intraday-Markt erforderlich sind, werden im weiteren Verlauf dieser Arbeit derartige Aktionen neun Minuten vor Anbruch des entsprechenden 15-Minuten-Intervalls ausgewählt. Auf diese Weise bleibt ausreichend Zeit für die technische und organisatorische Abwicklung des Kaufs. Die eigentliche Entscheidung über den kurzfristigen Kauf oder auch Verkauf von Strom wird mittels eines lernenden Agenten getroffen, dessen Funktionsweise in Abschnitt 7.3.4 aufgezeigt wird.

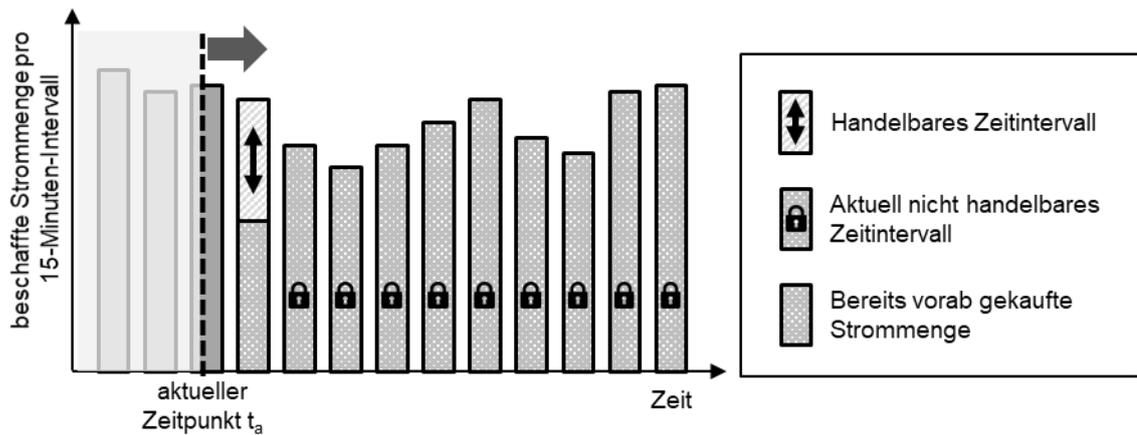


Abbildung 25: Handelbare Zeitintervalle des kurzfristigen Stromhandels

5.4 Fazit

Das Ergebnis des vorliegenden Kapitels ist ein Steuerungsschema zur Integration von kurzfristigem Stromhandel, Batteriespeicher, Eigenerzeugung und Netzbezug. Um dieses Ziel zu erreichen, wurde im ersten Schritt die Betrachtung auf die im Kontext der Produktionssteuerung relevanten Energiebezugsoptionen eingegrenzt, wobei die Vorlaufzeit als entscheidendes Merkmal identifiziert wurde. Im nächsten Schritt erfolgte die Modellierung aller Energiebezugsoptionen in Form von Kosten sowie der technischen und organisatorischen Randbedingungen. Zur Integration in die Produktionssteuerung wurden die Bezugsoptionen in zwei Teilsysteme eingeteilt: einerseits der kurzfristige Stromhandel, der auf Basis eines RL-Agenten stets mit neun Minuten Vorlauf Strom für ein anstehendes 15-Minuten-Intervall kauft oder verkauft. Andererseits steuert im Rahmen des Ad-hoc-Bezugs ein RL-Agent die Lade-/Entladeleistung des Batteriespeichers, während die übrigen Eigenerzeugungsanlagen mit Hilfe einer regelbasierten Logik in den sequenziellen Ablauf integriert werden.

Mit der im vorliegenden Kapitel erarbeiteten Modellierung der Energiebezugsoptionen und der Ableitung des Steuerungskonzeptes wird die Vorarbeit geschaffen, den Energiebezug in die Produktionssteuerung zu integrieren. Dieser verbleibende nächste Schritt erfolgt im nachfolgenden Kapitel 7.

6 Gesamtkostenbasiertes Zielsystem

Im vorliegenden Kapitel wird ein gesamtkostenbasiertes Zielsystem für die Produktionssteuerung entwickelt, in dem alle von der Produktionssteuerung beeinflussbaren Kosten enthalten sind und auf diese Weise die logistischen Zielgrößen repräsentieren. Um dieses Ziel zu erreichen, wird zunächst in Abschnitt 6.1 der Betrachtungsraum eingegrenzt. Anschließend werden die aus der Literatur bekannten Kostenfaktoren analysiert (Abschnitt 6.2). Die durch diese Schritte gefilterten Kosten werden in Abschnitt 6.3 mit Hilfe eines Wirkgefüges in ihre Kostenbestandteile zerlegt und die aufgedeckten monetären Auswirkungen in Abschnitt 6.4 anhand von mathematischen Formeln modelliert. Auf diese Weise wird eine Berechnung der Kosten zur Laufzeit der Produktionssteuerung ermöglicht. Alle erforderlichen Informationen werden abschließend in einem Datenmodell (Abschnitt 6.5) zusammengefasst.

6.1 Vorauswahl der Kostenfaktoren

Die Herstellkosten produzierender Unternehmen können nach PLINKE ET AL. (2015) im Sinne der Zuschlagskalkulation in Material- und Fertigungskosten gegliedert werden. Die Materialkosten setzen sich zusammen aus Materialeinzelkosten, die direkt einem gefertigten Produkt zugeordnet werden können, und Materialgemeinkosten, die beispielsweise Lagerkosten oder Anfahrverluste beinhalten. Die wesentlichen Bestandteile der Fertigungskosten sind die Fertigungseinzelkosten wie Fertigungslöhne und die Fertigungsgemeinkosten, welche sich unter anderem aus kalkulatorischen Kosten für Gebäude, Abschreibungen von Investitionen und Zinsen zusammensetzen. Darüber hinaus können unter Umständen zusätzlich Sondereinzelkosten der Fertigung etwa für Patente oder Spezialwerkzeuge anfallen.

Hinsichtlich einer gesamtkostenbasierten Zielfunktion für die Produktionssteuerung sind lediglich diejenigen Kosten von Relevanz, die durch Entscheidungen der Produktionssteuerung beeinflusst werden können. Bei Betrachtung der in Anlehnung an BULLINGER ET AL. (2009) der PPS vorgelagerten Schritte Produktentwicklung sowie Prozess-Engineering und Fabrikplanung wird deutlich, dass ein wesentlicher Teil der Herstellkosten bereits zu einem früheren Zeitpunkt feststeht (vgl. Abbildung 26). So wird der Materialbedarf für jedes später zu fertigende Produkt bereits weitestgehend im Zuge der Produktentwicklung festgelegt. Dadurch ist der wesentliche Bedarf an Roh- und Hilfsstoffen bereits durch den Arbeitsplan und damit auch die Materialeinzelkosten bestimmt. Eine Beeinflussung durch

Gesamtkostenbasiertes Zielsystem

Maßnahmen der Produktionssteuerung ist nicht mehr möglich (DANNHAUSER 2000). Lediglich der Materialausschuss beim Rüsten oder Lagern kann sich kurzfristig steuerungsbedingt noch ändern. Das Prozess-Engineering und die Fabrikplanung bestimmen die Gestaltung des Produktionssystems und die erforderlichen Investitionen in Gebäuden oder Produktionsressourcen. Die entstehenden Kosten aufgrund von Abschreibungen und Zinsen können im nachgelagerten Fabrikbetrieb ebenfalls nicht mehr beeinflusst werden. In der Kostenartenrechnung werden diese Kosten folglich auch Fixkosten genannt, da sie innerhalb eines definierten Betrachtungszeitraums nicht verändert werden können (HORSCH 2015). So bezeichnet LÖDDING (2008) Finanzierungskosten von Produktionsanlagen im Kontext der Fertigungssteuerung als „sunk costs“, weil sie zu diesem Zeitpunkt nicht mehr beeinflussbar sind und dadurch auch keine Entscheidungsrelevanz aufweisen. Lediglich der Lebensdauerverbrauch von Verschleißteilen wird durch die PPS bedingt, da in diesem Schritt die Auslastung einer Produktionsressource festgelegt wird.

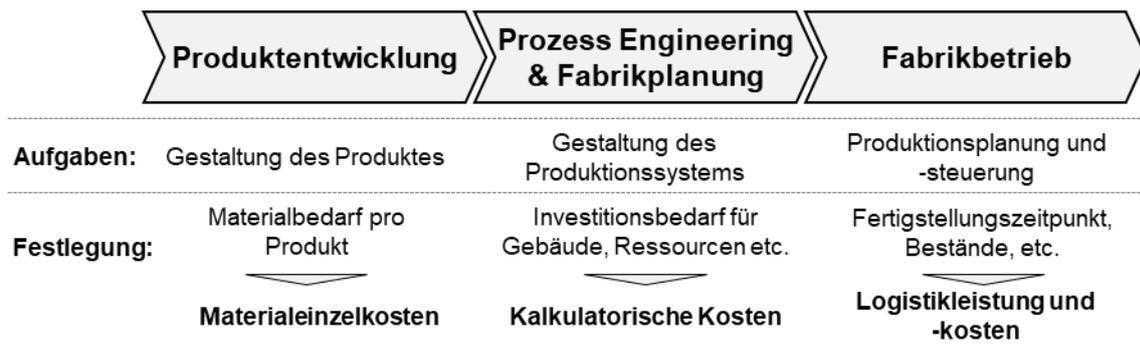


Abbildung 26: Festlegung der Herstellkosten im Produktlebenszyklus

Dennoch existieren neben den Materialgemeinkosten und kalkulatorischen Kosten noch zahlreiche weitere Kostenfaktoren, die in der Regel unter den Begriff *Logistikkosten* zusammengefasst werden und deren Beeinflussbarkeit durch die Produktionssteuerung zu prüfen ist (LÖDDING 2008). Zudem beeinflusst die PPS maßgeblich die *Logistikleistung*, deren Nichterfüllung zusätzliche Kosten nach sich ziehen kann. Die vorgelagerte Produktionsplanung legt bereits ein Produktionsprogramm fest. Im Zuge dessen werden beispielsweise die Schichtzeiten festgelegt. Es kann unter idealen Bedingungen in dieser Form abgearbeitet werden (vgl. Abschnitt 2.3). In diesem Schritt werden kostenrelevante Entscheidungen, wie etwa die Losgrößenbildung oder die Maschinenbelegung, getroffen, die von der Produktionssteuerung nicht mehr beeinflusst werden können. Zur Bewertung eines alternativen Plans der Produktionssteuerung ist es daher ausreichend, die Änderungen der Kosten im Vergleich zum *Ist-Plan* der Produktionsplanung zu betrachten. Wie in

der beispielhaften Darstellung in Abbildung 27 angedeutet, fallen ausschließlich zusätzliche Aufwände, wie in diesem Fall für Mehrarbeit und Rüsten, ins Gewicht. Bei der nachfolgenden Analyse geht es somit um eine relative Kostenbetrachtung, im Zuge derer die Kostenauswirkung mehrerer Lösungsalternativen verglichen werden.

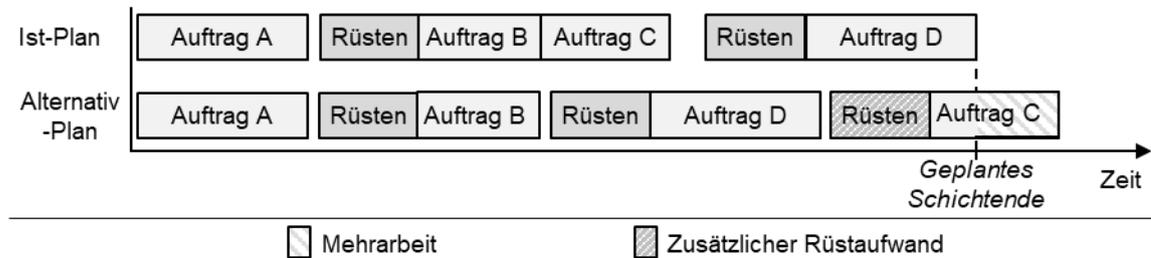


Abbildung 27: Beispielhafte Produktionspläne für eine Produktionsressource

Die Produktionssteuerung greift in der Regel nicht in die Bearbeitungsparameter auf Anlagenebene ein. Daher erfolgt die Kostenbetrachtung in diesem Kapitel unter der Annahme, dass die Produktionsqualität nicht durch Maßnahmen der Produktionssteuerung beeinflusst werden kann.

6.2 Analyse der durch die Produktionssteuerung beeinflussbaren Kostenverursacher

Als Ausgangspunkt der weiteren Analyse werden die insgesamt neun in der Literatur identifizierten Kostenfaktoren (vgl. Abschnitt 3.3) herangezogen. Um diese im Hinblick auf die Entscheidungsrelevanz für die Produktionssteuerung zu bewerten, werden zunächst die entsprechenden *Kostenverursacher* abgeleitet. Sie beschreiben das Ereignis, das die durch die Kostenfaktoren repräsentierten, monetären Mehraufwände auslöst. Beispielsweise werden die Terminüberschreitungskosten durch das Ereignis einer verspäteten Fertigstellung verursacht. Im zweiten Schritt wird eine qualitative Überprüfung durchgeführt, ob die Produktionssteuerung die identifizierten Kostenverursacher beeinflussen kann. Damit eine detailliertere Betrachtung möglich wird, erfolgt eine Unterteilung in die drei Funktionen der Produktionssteuerung nach LÖDDING (2008): *Auftragsfreigabe*, *Reihenfolgebildung* und *Kapazitätssteuerung*. Da die Belegungsplanung, wie bereits in 2.3.1 beschrieben, im Kontext dieser Arbeit nicht als Teil der Steuerung betrachtet wird, erfolgt keine Berücksichtigung dieses Schrittes im Zuge der weiteren Analyse. Die

Ergebnisse sind anhand einer Beeinflussungsmatrix (vgl. Abbildung 28) zusammengefasst. Nachfolgend wird die theoretische Herleitung dieser Erkenntnisse diskutiert.

Die Ursache für zusätzliche *Terminüberschreitungskosten* ist eine zu späte Fertigstellung eines Produktionsauftrags. Die Auftragsfreigabe bestimmt genauso wie die Reihenfolgebildung direkt den Bearbeitungszeitpunkt eines Auftrags. Demnach können beide Funktionen zu einem verspäteten Fertigstellungstermin führen. Da die Kapazitätssteuerung nach LÖDDING (2008) beispielsweise in Form einer Kapazitätserhöhung durch Überstunden ebenfalls in der Lage ist, den Bearbeitungszeitpunkt direkt zu verändern, liegt auch hier eine Beeinflussung vor. Umgekehrt können alle drei Funktionen der Produktionssteuerung durch ihre Möglichkeit zur Verschiebung des Bearbeitungszeitpunkts auch eine zu frühe Fertigstellung von Aufträgen hervorrufen, was wiederum zu Terminunterschreitungskosten führen kann. Folglich können sowohl *Terminüber-* als auch *Terminunterschreitungskosten* von allen drei Funktionen der Produktionssteuerung beeinflusst werden.

Nichtnutzungskosten entstehen durch Maschinenstillstände. Während dieser führt die jeweilige Maschine keine wertschöpfende Bearbeitung durch (NIEHUES 2017). Da die Auftragsfreigabe maßgeblich die kurzfristige Auslastung der jeweiligen Produktionsressource bestimmt, können auf diese Weise Maschinenstillstände aufgrund mangelnder Aufträge verursacht werden. Dies zieht eine reduzierte Auslastung nach sich. Dieser Einfluss ist etwa auch im Modell zur Fertigungssteuerung von NYHUIS (2008) abgebildet. Demgegenüber ist die Reihenfolgeplanung zwar in der Lage, durch ein Vertauschen von Aufträgen den Bearbeitungszeitpunkt zu verändern und Einfluss auf die erforderlichen Rüstaufwände zu nehmen (WIENDAHL 2010). Da Rüstvorgänge jedoch keine wertschöpfende Tätigkeit sind, kann insgesamt die Dauer von nichtwertschöpfenden Maschinenzuständen innerhalb einer Periode durch eine Reihenfolgeänderung nicht beeinflusst werden. Die Kapazitätssteuerung hingegen kann durch eine Verlängerung der Produktionszeiten theoretisch temporäre Stillstände begünstigen. Im Falle einer Kapazitätserhöhung bei gleichbleibendem Produktionsprogramm erhöht sich somit zwangsläufig die Dauer von nicht wertschöpfenden Betriebsphasen (LÖDDING 2008).

Die *Maschinenkosten* werden meist in Form eines Maschinenstundensatzes beziffert, der sich aus den kalkulatorischen Ausgaben für die Abschreibung, Zinsen, Räumlichkeiten und Instandhaltung zusammensetzt (VDI-RICHTLINIE). Es wird deutlich, dass kurzfristige Entscheidungen nicht in der Lage sind, diese Kosten zu

beeinflussen (vgl. Abschnitt 6.1). Eine Ausnahme ist die Instandhaltung. Das Erfordernis von Instandhaltungsmaßnahmen hängt, unter der Prämisse einer zustandsabhängigen Instandhaltungsstrategie, von dem Lebensdauerverbrauch der Verschleißteile und den Betriebsstunden einer Produktionsanlage beziehungsweise dem resultierenden Output ab (DANNHAUSER 2000). Dieser wird durch die Produktionsplanung bei der Erstellung des Belegungsplans definiert. Die Auftragsfreigabe, Reihenfolgebildung und Kapazitätssteuerung bestimmen zwar den Zeitpunkt der Bearbeitung, die Anzahl der von jeder Maschine zu bearbeitenden Aufträge bleibt davon jedoch unberührt. Über eine Produktionsperiode von beispielsweise einer Schicht betrachtet kann die Produktionssteuerung demnach den Bearbeitungszeitpunkt einzelner Aufträge ändern. In Summe bleiben die von jeder Maschine bearbeiteten Aufträge und damit der Lebensdauerverbrauch unverändert.

Rüstkosten werden von Rüstvorgängen verursacht, die zwischen der Bearbeitung von unterschiedlichen Aufträgen auf einer Anlage erforderlich sind. Die Anzahl der Rüstvorgänge in einer Produktionsperiode hängt daher maßgeblich von der Auftragsreihenfolge und somit dem Schritt der Reihenfolgebildung ab (ADAM 1998). Die Auftragsfreigabe und Kapazitätssteuerung beeinflussen die Auftragsreihenfolge dagegen nicht.

Werkzeugen wird in der Regel, ähnlich wie Maschinen, eine definierte Standzeit zugeordnet, nach deren Ablauf eine Reparatur oder Überholung erforderlich ist (SCHENK 2010). Jedem Bearbeitungsschritt kann somit ein gewisser Lebensdauerverbrauch in Form von Verschleiß zugeordnet werden, woraus sich die jeweiligen *Werkzeugkosten* ergeben (MILBERG 1992). Die für eine Bearbeitung zu verwendenden Werkzeuge und Parameter sind bereits im Arbeitsplan festgelegt (KURBEL 2013). Da die Auftragsfreigabe, Reihenfolgebildung oder Kapazitätssteuerung keine Änderungen der Prozessparameter oder der Werkzeugwahl vornehmen, ist eine Beeinflussung des Werkzeugverschleißes und damit der Werkzeugkosten nicht gegeben.

Der Aufbau von Beständen verursacht eine Erhöhung von *Bestandskosten*. Kurzfristig können Bestandserhöhungen auftreten, insbesondere durch Materialpuffer zwischen aufeinanderfolgenden Produktionsressourcen, die eine Erhöhung der Durchlaufzeit nach sich ziehen (DANGELMAIER 2009). Die Höhe von Pufferbeständen wird wesentlich vom Auftragszugang durch die Auftragsfreigabe beeinflusst (WIENDAHL 2010). Außerdem kann die Reihenfolgebildung oder Kapazitätssteuerung einen kurzfristigen Aufbau von Materialpuffern nach sich ziehen,

Gesamtkostenbasiertes Zielsystem

wenn beispielsweise zusätzliche Rüstvorgänge oder Stillstandzeiten verursacht werden.

Kosten für Intralogistik entstehen durch Transporte von Halbzeugen, Bauteilen oder Werkzeugen. Im Zuge der Maschinenbelegung wird bereits in der Produktionsplanung bestimmt, an welchem Ort die jeweiligen Materialien benötigt werden. Bleibt im Zuge der Produktionssteuerung die Maschinenfolge unverändert, kann lediglich der Zeitpunkt der erforderlichen Bereitstellung beeinflusst werden. Die Transportwege und -aufwände bleiben unberührt. Demnach können die Kosten für Intralogistik im Zuge einer Produktionssteuerung, die keine Anpassung der Maschinenbelegung vornimmt, vernachlässigt werden. (NIEHUES 2017)

Kostenfaktor	Kostenverursacher	Funktionen der Produktionssteuerung		
		Auftragsfreigabe	Reihenfolgebildung	Kapazitätssteuerung
Terminüberschreitungskosten	Zu späte Fertigstellung	●	●	●
Terminunterschreitungskosten	Zu frühe Fertigstellung	●	●	●
Nichtnutzungskosten	Zusätzlicher Maschinenstillstand	●	○	●
Maschinenkosten	Zusätzlicher Maschinenverschleiß	○	○	○
Rüstkosten	Zusätzliche Rüstvorgänge	○	●	○
Werkzeugkosten	Zusätzlicher Werkzeugverschleiß	○	○	○
Bestandskosten	Bestandserhöhung	●	●	●
Kosten für Intralogistik	Zusätzlicher Transport	○	○	○
Lohnkosten	Mehrarbeit	○	○	●

Legende: ● Beeinflussung ○ Keine Beeinflussung

Abbildung 28: Herleitung und Beeinflussung der Kostenverursacher durch die Funktionen der Produktionssteuerung

Mehrarbeit in Form von Zusatzschichten oder Überstunden führt zu einer Erhöhung der *Lohnkosten*. Im Zuge der Produktionsplanung wird der benötigte Arbeitsaufwand abgeschätzt und ein Schichtplan erstellt (HACKSTEIN 1989). Die Produktionssteuerung hat im Rahmen der Kapazitätssteuerung die Aufgabe, bei kurzfristigem Bedarf die bereits existierende Personalplanung entsprechend zu verändern (LÖDDING 2008). Aus Kostensicht ist für die Produktionssteuerung daher lediglich die relative Änderung der Personalstunden zum ursprünglichen Plan relevant. Die

Reihenfolgebildung und Auftragsfreigabe haben keine Auswirkungen auf den Personaleinsatz und somit auf die Personalkosten.

Die Ergebnisse der in diesem Unterkapitel dargelegten Analyse sind in Abbildung 28 zusammengefasst. Von den aus in der Literatur abgeleiteten neun Kostenfaktoren werden gemäß der durchgeführten Analyse die Maschinen-, Werkzeugkosten und Kosten für Intralogistik nicht durch die Produktionssteuerung beeinflusst. Sie sind daher für die weitere Betrachtung nicht von Relevanz.

6.3 Ableiten der monetären Auswirkungen

Im vorangegangenen Abschnitt wurden die relevanten Kostenfaktoren ausgewählt und die jeweiligen Kostenverursacher identifiziert. Um die einzelnen Bestandteile genauer untersuchen zu können, werden in diesem Abschnitt die monetären Auswirkungen der jeweiligen Kostenverursacher analysiert. Dazu werden in einem Wirkgefüge zunächst die organisatorischen und technischen Auswirkungen analysiert. Daraus lassen sich die monetären Folgen ableiten. Das in Abbildung 29 dargestellte Wirkgefüge resultiert aus dieser Herangehensweise und wird nachfolgend hergeleitet. Aus der systematischen Analyse der identifizierten Kostenverursacher gehen insgesamt sieben verschiedene Kostenbestandteile hervor. Dabei werden die *Verzugskosten und Konventionalstrafen*, *Kosten durch Vertrauensverluste*, *Erhöhte Versandkosten* und *zusätzliche Lagerkosten* aufgrund der Abhängigkeit zum Fertigstellungstermin zu den *terminbezogenen Kosten* zusammengefasst.

Die *verspätete Fertigstellung* eines Auftrags kann zu einer nicht fristgerechten Auslieferung an den Kunden führen. Abhängig von den Vertragsvereinbarungen können dem Lieferanten in diesem Fall Konventionalstrafen auferlegt werden. Zudem leidet das Vertrauensverhältnis zwischen Kunden und Lieferanten, was möglicherweise langfristig zu Auftragseinbußen führt (LÖDDING 2008). Eine verspätete Fertigstellung kann unter Umständen ganz oder teilweise durch einen Expressversand ausgeglichen werden, der jedoch in der Regel mit höheren Versandkosten verbunden ist. Umgekehrt kann eine *zu frühe Fertigstellung* eines Auftrags dazu führen, dass das Produkt bis zur Auslieferung zwischengelagert werden muss und dadurch zusätzliche Lagerkosten verursacht.

Maschinenstillstände ziehen einen Kapazitätsverbrauch nach sich. Währenddessen wird keine Wertschöpfung kreiert. Aus reiner Kostensicht fallen demnach durch einen Stillstand zunächst keine direkten Ausgaben an. Viele Stillstände verschlechtern jedoch maßgeblich die Auslastung. Die dabei vergeudete Zeit kann

bewirken, dass innerhalb der Produktionsperiode nicht mehr alle geplanten Aufträge fristgerecht abgearbeitet werden können. Die wesentliche potenzielle Auswirkung von Maschinenstillständen ist somit eine verspätete Fertigstellung von Aufträgen, was wiederum in Konventionalstrafen, Vertrauensverlust und erhöhten Versandkosten resultieren kann.

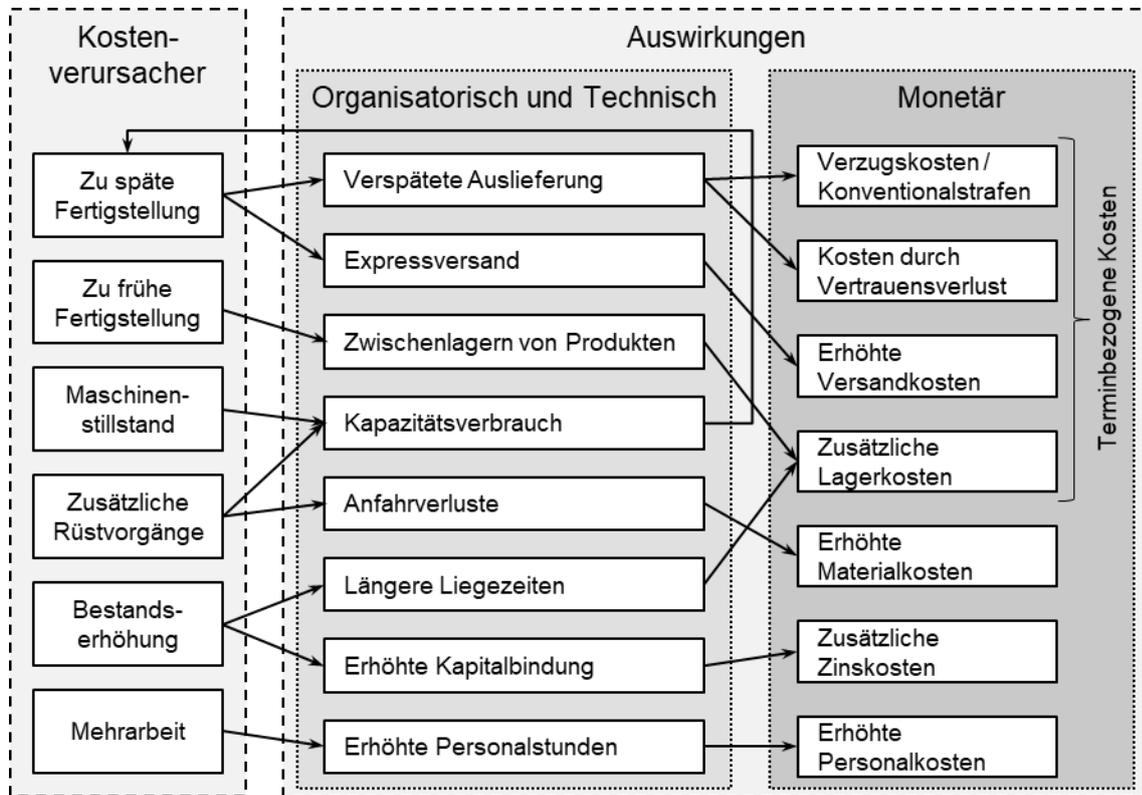


Abbildung 29: Wirkgefüge der Kostenverursacher und deren Auswirkungen

Zusätzliche Rüstaufwände ziehen mehrere negative Auswirkungen nach sich. Zum einen stehen während des Rüstvorgangs das dafür benötigte Personal und die zu rüstende Produktionsressource nicht zur Bearbeitung von Aufträgen zur Verfügung. Dieser zusätzliche Kapazitätsverbrauch kann analog zu Maschinenstillständen in einer verspäteten Fertigstellung von Aufträgen resultieren. Außerdem können bei Rüstvorgängen Anfahrverluste und damit ein erhöhter Materialverbrauch entstehen, beispielsweise, wenn nach dem Austausch eines Werkzeugs eine bestimmte Hochfahrzeit erforderlich ist, während der Ausschuss produziert werden kann (KLETTI & SCHUMACHER 2014).

Eine temporäre *Erhöhung von Beständen* zieht einerseits eine längere Liegezeit von Produkten nach sich (WIENDAHL 2010). Falls dazu auf externe Pufferkapazitäten zurückgegriffen werden muss, fallen zusätzliche Lagerkosten an. Da bestehende Materialpuffer in der Regel lediglich kalkulatorische Kosten für Investition,

Raumbedarf etc. verursachen, werden durch deren Nutzung keine entscheidungsrelevanten Kosten für die Produktionssteuerung wirksam (vgl. Abschnitt 6.1). Andererseits hat eine Bestandserhöhung zusätzlich eine kurzfristige Erhöhung des gebundenen Kapitals zur Folge, was wiederum Zinskosten verursacht.

Die Erweiterung der Kapazität durch Mehrarbeit führt zu zusätzlichen Arbeitsstunden der Mitarbeiter. Als Folge fallen erhöhte Personalkosten an, wobei neben den reinen Überstunden je nach Tarif- und Schichtmodell auch Zulagen zu entrichten sind, beispielsweise für Wochenendarbeit.

6.4 Modellierung der Kostenbestandteile

Im folgenden Unterkapitel werden die in den beiden vorausgegangenen Abschnitten identifizierten monetären Auswirkungen mathematisch modelliert. Auf diese Weise können die einzelnen Kostenverursacher im Falle des Auftretens quantifiziert werden.

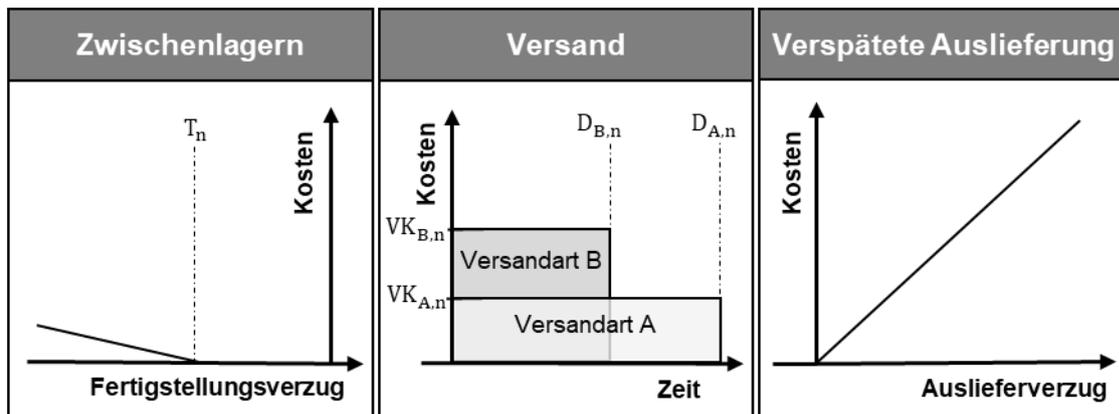
6.4.1 Terminbezogene Kosten

Da sich eine *zu frühe* und eine *zu späte Fertigstellung* gleichermaßen auf den Fertigstellungstermin beziehen, werden diese beiden Aspekte in einem *terminbezogenen Kostenmodell* zusammengefasst. Gemäß dem Wirkgefüge in Abbildung 29 setzen sich die terminbezogenen Kosten aus den Bestandteilen Verzugskosten beziehungsweise Konventionalstrafen, Kosten durch Vertrauensverlust, erhöhte Versandkosten und zusätzliche Lagerkosten zusammen.

Der Fertigstellungsverzug eines Auftrages lässt sich durch die Differenz aus geplantem und tatsächlichem Fertigstellungstermin berechnen. Ist der Verzug negativ, so liegt eine zu frühe Fertigstellung vor und es kann erforderlich sein, das fertige Produkt bis zur planmäßigen Auslieferung an den Kunden einzulagern (SCHMIDT ET AL. 2014). Dieser Mechanismus wird mit Hilfe einer Toleranzperiode T_n modelliert, ab deren Unterschreitung Lagerkosten anfallen. Die Lagerkosten setzen sich zusammen aus Kosten für die Lagerkapazität wie Raum und Abschreibungen, aus dem beim Ein- und Auslagern erforderlichen Handling in Form von Personalkosten und den Verwaltungskosten (TILLENBURG 2008). Ein Großteil dieser Kosten ist nicht steuerungsrelevant (vgl. Abschnitt 6.1). Falls jedoch bei der Lagerhaltung auf externe Partner zurückgegriffen wird, werden alle resultierenden Kosten

auf die entsprechende Rechnung umgelegt. Der Kostensatz ist in der Regel abhängig von der Lagerzeit und dem erforderlichen Platzbedarf (WEBER 2012). Aus diesem Grund werden die Lagerkosten näherungsweise als lineare Funktion modelliert, die mit zunehmender frühzeitiger Fertigstellung ansteigt (vgl. Abbildung 30).

Der Kunde ist in vielen Fällen auf die zuverlässige und pünktliche Anlieferung des Lieferanten angewiesen, etwa um keine Auslastungsverluste oder gar Stillstände der eigenen Produktion zu riskieren (LÖDDING 2008). Droht ein Auftrag, zu spät beim Kunden anzukommen, kann die Verzögerung möglicherweise durch die Wahl einer anderen Versandart ausgeglichen werden. Jede Versandart wird wesentlich von der Dauer und den Kosten charakterisiert. Expressversand weist etwa eine geringe Dauer bei vergleichsweise hohen Kosten auf (LÖDDING 2008). In Abbildung 30 werden beispielhaft zwei unterschiedliche Versandoptionen modelliert. Die Verfügbarkeit von Versandoptionen ist vorab zu prüfen.

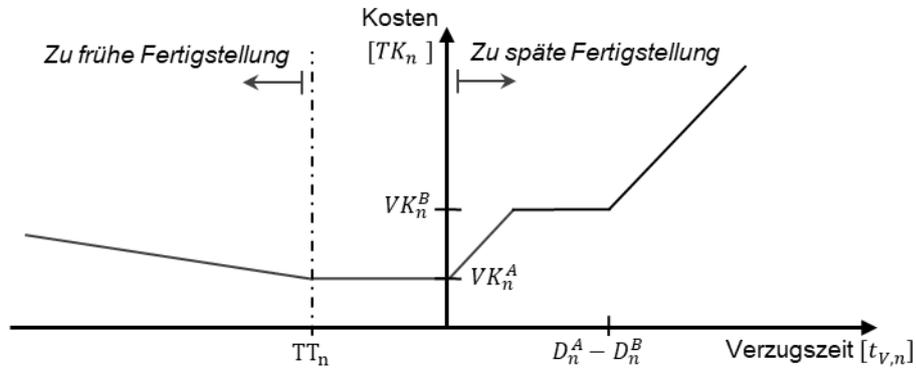


Legende: T_n Termintoleranz für Auftrag n
 $VK_{A,n}$ Kosten für Versandart A und Auftrag n
 $D_{A,n}$ Dauer von Versandart A und Auftrag n
 $VK_{B,n}$ Kosten für Versandart B und Auftrag n
 $D_{B,n}$ Dauer von Versandart B und Auftrag n

Abbildung 30: Bestandteile der terminbezogenen Kosten (ROESCH ET AL. 2018)

Im Falle einer verspäteten Lieferung können dem Lieferanten vertraglich vereinbarte Konventionalstrafen auferlegt werden, die mit zunehmender Verzögerung ansteigen (LÖDDING 2008). Die Höhe der Strafen kann demnach direkt dem gültigen Liefervertrag entnommen werden. Ein anderer wichtiger Aspekt bei der verspäteten Auslieferung ist der daraus resultierende Vertrauensverlust auf Seiten des Kunden. Die zukünftigen Auswirkungen, beispielsweise in Form entgangener Aufträge, lassen sich nicht direkt monetär beziffern und können lediglich abgeschätzt werden. Im schlimmsten Fall wird der Kunde in Zukunft die Aufträge bei einem zuverlässigeren Wettbewerber platzieren (LÖDDING 2008). Insbesondere, da sich Vertrauensbildung oft als langwieriger und aufwendiger Prozess erweist,

ist dieser Faktor nicht zu vernachlässigen (SCHUH 2011). Der Vertrauensverlust steigt schätzungsweise analog zu den Konventionalstrafen mit zunehmender Verspätung an. Aus diesem Grund werden die Verzugskosten näherungsweise als lineare Funktion modelliert, die mit wachsender Verzugszeit ansteigt.



Legende:	TT_n Termintoleranz für Auftrag n	$VK_{B,n}$ Kosten für Versandart B und Auftrag n
	$VK_{A,n}$ Kosten für Versandart A und Auftrag n	$D_{B,n}$ Dauer Versandart B und Auftrag n
	$D_{A,n}$ Dauer Versandart A und Auftrag n	$t_{v,n}$ Verzugszeit von Auftrag n

Abbildung 31: Berechnung der terminbezogenen Kosten (ROESCH ET AL. 2018)

Zusammengesetzt aus den drei Kostenbestandteilen ergibt sich nachfolgende Kostenfunktion zur Berechnung der durch die Fertigstellung eines Auftrags n zum Zeitpunkt t angefallenen Kosten:

$$TK_n(t) = \begin{cases} VK_n^A + lk_n(TT_n - t + t_{F,n}) & , t - t_{F,n} \leq TT_n \\ VK_n^A & , TT_n < t - t_{F,n} \leq 0 \\ VK_n^A + vk_n(t - t_{F,n}) & , 0 < t - t_{F,n} \leq \frac{(VK_n^B - VK_n^A)}{vk_n} \\ VK_n^B & , \frac{(VK_n^B - VK_n^A)}{vk_n} < t - t_{F,n} \leq \\ & \leq t - t_{F,n} + D_n^A - D_n^B \\ VK_n^B + vk_n(t - t_{F,n} - D_n^A + D_n^B) & , D_n^A - D_n^B < t - t_{F,n} \end{cases} \quad (16)$$

- $TK_n(t)$ Terminbezogene Kosten von Auftrag n bei Fertigstellungstermin t
- $t_{F,n}$ Soll-Fertigstellungstermin von Auftrag n
- TT_n Termintoleranz von Auftrag n
- lk_n spezifischer Lagerkostensatz von Auftrag n
- VK_n^A Kosten für Versandart A und Auftrag n
- VK_n^B Kosten für Versandart B und Auftrag n
- D_n^A Dauer von Versandart A und Auftrag n
- D_n^B Dauer von Versandart B und Auftrag n
- vk_n spezifischer Verzugskostensatz von Auftrag n

Im Zuge der Modellierung der terminbezogenen Kosten wurden ausschließlich Lagerkosten berücksichtigt, die für fertiggestellte Produkte anfallen. Darüber hinaus können nach Abbildung 29 zusätzliche Lagerkosten durch erhöhte Material- und Halbzeug-Bestände in der Produktion verursacht werden. Deren Höhe hängt dabei stark davon ab, ob interne Pufferkapazitäten zur Verfügung stehen oder kurzfristige zusätzliche Kapazitäten von extern geschaffen werden müssen. Ist Ersteres der Fall, so hat dies keine Auswirkungen auf die nicht steuerungsrelevanten Abschreibungen und Raumkosten (vgl. Abschnitt 6.1), die einen wesentlichen Kostenbestandteil der Lagerkosten bilden. Die Schaffung von externen Lagerkapazitäten ist mit einem vergleichsweise deutlich größeren, organisatorischen Aufwand verbunden. Aufgrund des Entscheidungshorizonts der Produktionssteuerung von etwa einem Tag verursacht eine Verschiebung des Fertigungszeitpunkts innerhalb dieser zeitlichen Grenzen lediglich eine sehr temporäre Veränderung der Bestände. Aus diesem Grund erscheint es unwahrscheinlich, dass so kurzfristige externe Lagerkapazitäten genutzt werden. Vielmehr werden die im Produktionssystem vorhandenen und gleichzeitig auch endlichen Pufferkapazitäten ausgenutzt. Die Pufferkapazitäten geben somit eine harte Randbedingung für die Produktionssteuerung vor.

6.4.2 Erhöhte Materialkosten

Rüstvorgänge verbrauchen für die Dauer des Rüstens in erster Linie Ressourcen- und Personalkapazität, was sich wiederum, wie in Abschnitt 6.3 beschrieben, negativ auf die Terminerfüllung auswirken kann. Als einziger direkter Kostenbestandteil verbleiben daher die gestiegenen Materialkosten durch Anfahrverluste. Diese können sich auf einer Maschine je nach zu rüstendem Auftrag deutlich unterscheiden. Auch wenn die Höhe der Anfahrverluste schwanken kann, wird in dieser Arbeit eine näherungsweise Abschätzbarkeit des Materialverbrauchs von Rüstvorgängen auf Basis von Erfahrungswerten angenommen. Mittels der jeweiligen Materialpreise können somit die Kosten KR_r von einem Rüstvorgang r berechnet werden:

$$KR_r = mk_m \cdot AV_{m,r} \quad , r \in \{R_n\} \quad (17)$$

KR_n Zusätzliche Rüstkosten für die Fertigung von Auftrag n

mk_m Kostensatz von Material m

$AV_{m,r}$ Anfahrverluste an Material m bei Rüstvorgang r

R_n Für die Fertigung von Auftrag n erforderliche Rüstvorgänge

Die entsprechenden Werte können etwa einer Rüstmatrix entnommen werden, die jeden Rüstvorgang charakterisiert. Für den Fall mehrerer verbrauchter Materialien bei einem Rüstvorgang ist eine Erweiterung der Formel erforderlich.

6.4.3 Zusätzlich Zinskosten

Der Aufbau von Beständen innerhalb der Produktion zieht eine Erhöhung des gebundenen Kapitals nach sich, wofür kalkulatorisch eine entsprechende Verzinsung fällig wird. Kommt es durch Maßnahmen der Produktionssteuerung für eine bestimmte Periode zu einer Bestandserhöhung, so können sich aus dem für diesen Zeitraum fälligen Zinssatz bestimmte Kosten ergeben.

In der Produktion ist der Auftragszugang durch das vorab definierte Produktionsprogramm der Produktionsplanung festgelegt. Gleiches gilt für den Abgang, der in Form der Fertigstellungstermine geregelt wird. Eine allgemeine Veränderung der Bestände über einen längeren Zeitraum ist dementsprechend als eine Stellgröße der Produktionsplanung zu sehen und kann nicht von der Produktionssteuerung beeinflusst werden (NIEHUES 2017). Zudem beziehen sich die Kapitalbindungskosten auf den Zeitraum zwischen der Beschaffung von Material bis zum Erlöseingang, der aus dem daraus gefertigten Produkt generiert wird (APPELFELLER & BUCHHOLZ 2011). Dementsprechend verursacht der Bestandsaufbau erst dann kalkulatorische Kosten, wenn es zu einer Verzögerung des Erlöseingangs kommt. Im Kontext der Produktionssteuerung kann ein derartiger Aufschub durch eine verspätete Auslieferung an den Kunden verursacht werden. In diesem Fall werden neben diesen erhöhten Kapitalbindungskosten auch die zu Beginn dieses Unterkapitels modellierten Verzugskosten fällig. Aufgrund der den Verzugskosten zugrundeliegenden Konventionalstrafen und des Vertrauensverlusts kann davon ausgegangen werden, dass diese die vergleichsweise geringen Zinskosten deutlich übersteigen. Aus diesen Gründen erscheinen die Zinskosten im Kontext der Produktionssteuerung bei einer gleichzeitigen Betrachtung der Verzugskosten vernachlässigbar.

6.4.4 Erhöhte Personalkosten

Durch kurzfristige Kapazitätserweiterungen in Form von Überstunden und Sonderschichten fallen zusätzliche Personalkosten an. Deren Höhe hängt dabei einerseits von der Menge an zusätzlich erbrachten Arbeitsstunden ab. Andererseits können unter Umständen auch die Lohnkosten durch tariflich vereinbarte Sonderzah-

lungen oder andere Zuschüsse ansteigen (DANNHAUSER 2000). Diese sind beispielsweise im Falle von Nachtschichten oder Wochenendarbeit vom Arbeitgeber einzukalkulieren und bei der Entscheidung über eine Kapazitätserhöhung zu berücksichtigen. Neben einer Kapazitätserweiterung ist im Kontext der Produktionssteuerung auch eine kurzfristige Reduzierung der Kapazitäten denkbar, etwa bei temporär geringer Auslastung oder sehr hohen Energiepreisen gegen Ende oder Anfang einer Schicht. Generell sind die Möglichkeiten und Grenzen einer kurzfristigen Kapazitätserweiterung oder -reduzierung vorab mit den Mitarbeitern und anderen relevanten Stellen wie dem Betriebsrat zu klären. Die dabei gezogenen Randbedingungen dienen als Eingangsgröße für die Produktionssteuerung.

Zusammenfassend können die aus einer kurzfristigen Kapazitätsänderung direkt resultierenden Gesamtkosten aus dem jeweilig fälligen Personalstundensatz und der Menge an Arbeitsstunden berechnet werden. Im Falle einer Kapazitätsreduzierung fällt die Menge der Arbeitsstunden PS negativ aus:

$$PK_n = pk * PS_n \quad (18)$$

- PK_n Zusätzliche Personalkosten für die Fertigung von Auftrag n
- pk Personalkostensatz
- PS_n Zusätzlich geleistete Personalstunden für die Fertigung von Auftrag n

6.4.5 Funktion des gesamtkostenbasierten Zielsystems

Die produktionslogistischen Auswirkungen einer Planalternative auf die Gesamtkosten berechnen sich aus der Summe der drei einzelnen Kostenbestandteile:

$$GK_n(t) = \sum TK_n(t) + \sum KR_n + PK_n \quad (19)$$

- $GK_n(t)$ Gesamtkosten für die Fertigung von Auftrag n
- $TK_n(t)$ Terminbezogene Kosten von Auftrag n mit Fertigstellungszeitpunkt t
- KR_n Zusätzliche Rüstkosten für die Fertigung von Auftrag n
- PK_n Zusätzliche Personalkosten für die Fertigung von Auftrag n

Neben den hier genannten Kostenfaktoren von Seiten der Produktion werden im Rahmen der vorliegenden Arbeit des Weiteren die Energiekosten betrachtet, die im Zuge der Integration des Energiebezugs im Abschnitt 5.2.6 hergeleitet wurden.

6.5 Datenmodell

Zur Berechnung der in Abschnitt 5.4. abgeleiteten Kostenbestandteile ist eine Datengrundlage in Form von bestimmten Kennzahlen erforderlich. Da die benötigten Informationen aus heterogenen Quellen stammen, werden diese in Form des nachfolgenden Datenmodells strukturiert. Das Modell beschränkt sich ausschließlich auf die Beschreibung der Daten zur Berechnung der Kostenbestandteile, ohne die Energiekosten zu beinhalten. Für ein geeignetes Datenmodell der energieorientierten Produktionssteuerung wird auf die Vorarbeit von SCHULTZ (2018) verwiesen.

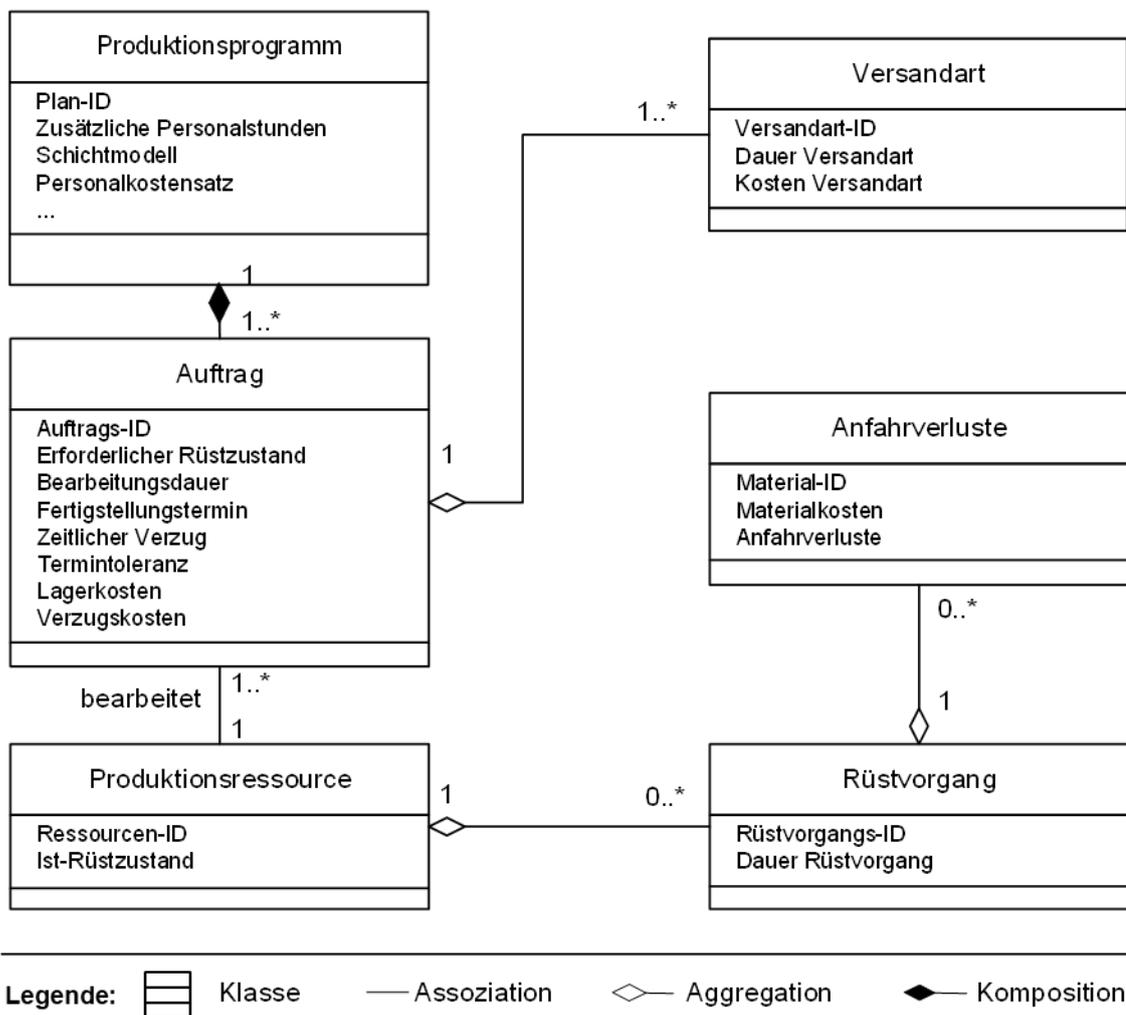


Abbildung 32: Datenmodell für das gesamtkostenbasierte Zielsystem

Für ein definiertes Produktionssystem, bestehend aus einer oder mehreren Produktionsressourcen, werden übergeordnet die zu erbringenden Arbeitszeiten und Schichten vorab durch die Produktionsplanung festgelegt. Auf dieser Ebene können zudem die Informationen über die Höhe von Mehrarbeit und die dafür fälligen

Personalstundensätze abgegriffen und hinterlegt werden. Ein Produktionsprogramm besteht zudem aus einer bestimmten Anzahl von Aufträgen, die über eine Auftragsnummer eindeutig identifiziert werden. Zudem werden Aufträge durch den erforderlichen Rüstzustand, die Bearbeitungsdauer, den spätesten Fertigstellungstermin, die Termintoleranz, die Lager- und Verzugskosten charakterisiert. Für jeden Auftrag ist zudem mindestens eine Versandart hinterlegt, die sich durch die Dauer und Kosten des Versands untereinander differenzieren. Jeder Auftrag ist zudem einer Produktionsressource mit einem aktuellen Ist-Rüstzustand zugeordnet. Für jeden auf der jeweiligen Ressource möglichen Rüstvorgang sind dessen Dauer, die Höhe der Anfahrverluste und die zugrundeliegenden Materialkosten aufgeführt. Das Datenmodell ist in Abbildung 32 zusammengefasst.

6.6 Fazit

In diesem Kapitel wurde ein gesamtkostenbasiertes Zielsystem für die Produktionssteuerung abgeleitet, das die Beantwortung der zweiten Forschungsfragen der vorliegenden Arbeit (Abschnitt 1.3) beinhaltet. Zu Beginn wurden dazu eine Vorauswahl und Eingrenzung der zu analysierenden Kostenfaktoren getroffen (Abschnitt 6.1). Als Ausgangspunkt für die weitergehende Analyse dienten die in der Literatur betrachteten Kostenfaktoren, deren Beeinflussbarkeit anhand der drei Aufgaben der Produktionssteuerung bewertet wurde (Abschnitt 6.2). Durch den Aufbau eines Wirkgefüges konnten Wechselwirkungen zwischen einzelnen Kostenverursachern systematisch aufgedeckt und die monetären Auswirkungen identifiziert werden (Abschnitt 6.3).

Auf dieser Basis erfolgten eine detaillierte Betrachtung und Modellierung der Kostenbestandteile, die aus den vorangegangenen Schritten resultierten (Abschnitt 6.4). Die resultierenden Lager-, Versand-, Verzugs-, Rüst- und Personalkosten werden abschließend in einem Datenmodell beschrieben (Abschnitt 6.5). Werden diese anhand der aufgezeigten Modellierungsansätze berechnet, so sind eine direkte Quantifizierung und Überführung der logistischen Zielgrößen in Kosten möglich. Die Relevanz der Kostenbestandteile ist ein allgemeingültiges Ergebnis. Durch die parametrierbare Modellierung der Kosten ist zudem eine Übertragbarkeit der Berechnung gewährleistet. Gegebenenfalls müssen einzelne Formeln für spezifische Anforderungen eines Unternehmens erweitert werden. Im Hinblick auf die weitere Struktur der vorliegenden Arbeit stellt die kostenbasierte Zielfunktion einen Input für das in Kapitel 7 zu entwickelnde Steuerungssystem dar.

7 Integrierte Produktionssteuerung mittels Reinforcement Learning

Die im vorliegenden Kapitel zu entwickelnde integrierte Steuerung ist der letzte Lösungsbaustein dieser Arbeit. Um die in Kapitel 3 abgeleiteten Anforderungen zu erfüllen, wird zu diesem Zweck, wie in Abschnitt 4.1 hergeleitet, ein Ansatz mit MARL verfolgt, dessen Ablauf zunächst in Abschnitt 7.1 aufgezeigt wird. Die Umgebung bildet die Grundlage für jegliche Optimierung mittels Reinforcement Learning. Die Modelle der beiden Elemente Strommarkt und ad-hoc Strombezug wurden bereits im vorherigen Kapitel dargelegt. Aus diesem Grund wird in Abschnitt 7.2 die Modellierung des Produktionssystems mit den zugehörigen Produktionsressourcen und Aufträgen beschrieben. Auf dieser Basis erfolgen anschließend die Modellierung der RL-Agenten (Abschnitt 7.3), die Formulierung der Belohnungsfunktion aus den beiden Hauptbestandteilen Produktions- und Energiekosten sowie insbesondere deren Vergabe (Abschnitt 7.4). Die Auswahl und Implementierung eines geeigneten Lernverfahrens sind in Abschnitt 7.5 beschrieben. Abschließend werden die entwickelten Bausteine und Abläufe in einer Trainingsumgebung zusammengeführt (Abschnitt 7.6).

7.1 Aufbau und Ablauf

Die energieorientierte Produktionssteuerung ist das zentrale Element zur gezielten Anpassung des Energiebedarfs in der Produktion und der Integration des Energiebezugs. Aufgrund der vorliegenden Anforderungen und des Lösungsansatzes mit MARL ergibt sich die in Abbildung 18 dargestellte Gesamtübersicht der Steuerung. Die Elemente gliedern sich in drei wesentliche Teilsysteme: die Steuerung mit den einzelnen RL-Agenten; die Umgebung, bestehend aus Strommarkt, ad-hoc Strombezug und Produktionssystem sowie die Belohnungsfunktion.

Während des Trainings interagieren die RL-Agenten durch die Wahl von Aktionen mit der Umgebung in Form einer Simulation. Die Auswirkungen dieser Aktionen werden mit Hilfe der Belohnungsfunktion bewertet und der generierte Erfahrungsschatz durch das Training in einem neuronalem Netz konserviert (vgl. Abschnitt 2.5.3). Zur Modellierung des Produktionsablaufs wird der zeitliche Verlauf des Gesamtsystems in konstante Zeitschritte mit einer Länge von drei Minuten diskretisiert, die damit auch die Taktung der Ressourcenagenten bestimmt. Prinzipiell kann die Länge der Zeitschritte auch anders festgelegt werden. Eine sinnvolle

Wahl hängt maßgeblich von dem Produktionssystem und der durchschnittlichen Länge der zu bearbeitenden Aufträge ab. Da der Energiemarkt in 15-Minuten-Intervallen (vgl. Abschnitt 2.2.4) abgewickelt wird, sollte im Zuge einer energieorientierten Steuerung die zeitliche Diskretisierung so gewählt werden, dass ein Vielfaches davon 15 Minuten ergibt. Die Festlegung auf einen Takt von drei Minuten bietet zudem die Chance, dass die Ressourcen- und der Speicheragent innerhalb der abrechnungsrelevanten 15-minütigen Hauptintervalle den Energiebedarf anpassen können (vgl. Abschnitt 5.3.1).

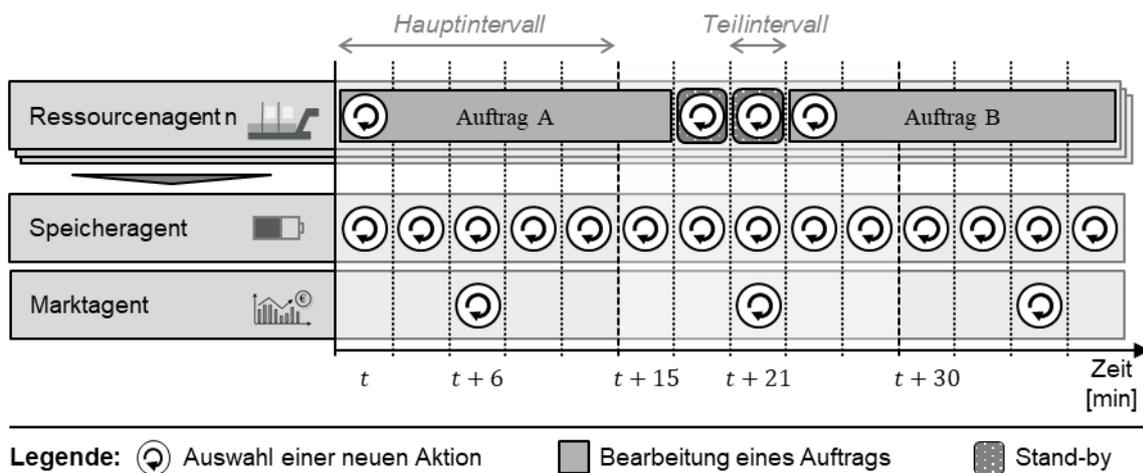


Abbildung 33: Simulationsablauf und Aktionsauswahl der Agenten
(in Anlehnung an ROESCH ET AL. 2020)

Der resultierende zeitliche Ablauf für die Auswahl von Aktionen der einzelnen Agenten ist in Abbildung 33 dargestellt. Während die *Ressourcenagenten* immer nach Abschluss einer Aktion, also am Ende einer Bearbeitung oder eines Stand-by-Schrittes eine neue Aktion auswählen, wird der *Speicheragent* alle drei Minuten aktiv und bestimmt den Lade- beziehungsweise Entladestrom der Batterie für den nächsten Zeitschritt. Im Gegensatz dazu handelt der *Marktagent* lediglich alle 15 Minuten, da Strom nur für Intervalle dieser Länge gehandelt werden kann (vgl. Abschnitt 5.3.2). Der Marktagent agiert dabei mit einer Vorlaufzeit von neun Minuten vor dem eigentlichen Beginn des 15-Minuten-Intervalls, in das die Auswirkungen des Handels fallen. Während eines Zeitschritts sind dabei zunächst alle Ressourcenagenten parallel an der Reihe, ihre nächsten Aktionen auszuwählen – vorausgesetzt, eine Aktion wurde soeben abgeschlossen. Ist ein Agent beispielsweise noch mit der Abarbeitung eines Produktionsauftrags beschäftigt, so führt dieser im aktuellen Zeitschritt keine Aktion aus. Dieser Sachverhalt und dessen

Auswirkungen spielen für die Gestaltung des Zustandsraums der Agenten im späteren Teil der Arbeit eine wichtige Rolle und werden in Abschnitt 7.3.1 genauer beleuchtet.

Durch die vorgelagerte Aktionsauswahl der Ressourcenagenten sind die im Anschluss agierenden Markt- und Speicheragenten in der Lage, bei ihrer Aktionswahl bereits die anstehenden Aktionen der Ressourcen zu berücksichtigen. Dies bietet insbesondere für den Batterieagenten eine hilfreiche Informationsgrundlage, da somit der für das Teilintervall zu erwartende Energieverbrauch bereits abgeschätzt werden kann (vgl. Abschnitt 5.3.2). Insgesamt handelt es sich demnach um ein System mit asynchron operierenden Agenten.

7.2 Modellierung des Produktionssystems

Im Folgenden wird die Modellierung des Produktionssystems vorgestellt, die die Basis für das Trainieren der Ressourcenagenten bildet. Zunächst wird dabei auf die einzelnen Produktionsressourcen eingegangen, bevor die zu bearbeitenden Produktionsaufträge abstrahiert werden und die Einbindung stochastischer Ereignisse skizziert wird.

7.2.1 Produktionsressourcen

Ein Produktionssystem besteht aus einer definierten Anzahl p an Produktionsressourcen PR , die vorgegebene Aufträge unter bestimmten Randbedingungen abarbeiten (vgl. Abschnitt 7.2.2). Die innerhalb einer Schicht von einer Produktionsressource PR_p zur Bearbeitung ausstehenden Produktionsaufträge befinden sich in der Warteschlange. Die Zuteilung, welche Produktionsressource den jeweiligen Auftrag zu fertigen hat, wurde im Zuge der Belegungsplanung bereits vorab durchgeführt (vgl. Abschnitt 2.3.1) und ist daher im Zuge der vorliegenden Betrachtung schon festgelegt.

Die Modellierungsparameter der einzelnen Produktionsressourcen sind in Abbildung 34 dargestellt. Während der Energieverbrauch im Stand-by-Zustand einen statischen Wert besitzt, verändern sich die Parameter zur Angabe der *aktuellen Aktion*, der *Dauer bis zum Ende der aktuellen Aktion* und des *aktuellen Rüstzustands* mit Fortlaufen der Simulation. Die Bearbeitung eines Auftrags setzt jeweils einen definierten Rüstzustand der Produktionsressource voraus. Soll ein Auftrag bearbeitet werden, der einen anderen Rüstzustand als den aktuellen erfordert, muss

die Produktionsressource vor Beginn der Bearbeitung entsprechend gerüstet werden. Die möglichen Rüstvorgänge können dabei anhand der *Rüstvorgangs-ID* eindeutig zugeordnet werden. Zudem ist jeweils die *Dauer des Rüstvorgangs* hinterlegt. Außerdem ist jede Produktionsressource lediglich in der Lage, eine definierte Auswahl an Auftragstypen zu bearbeiten.

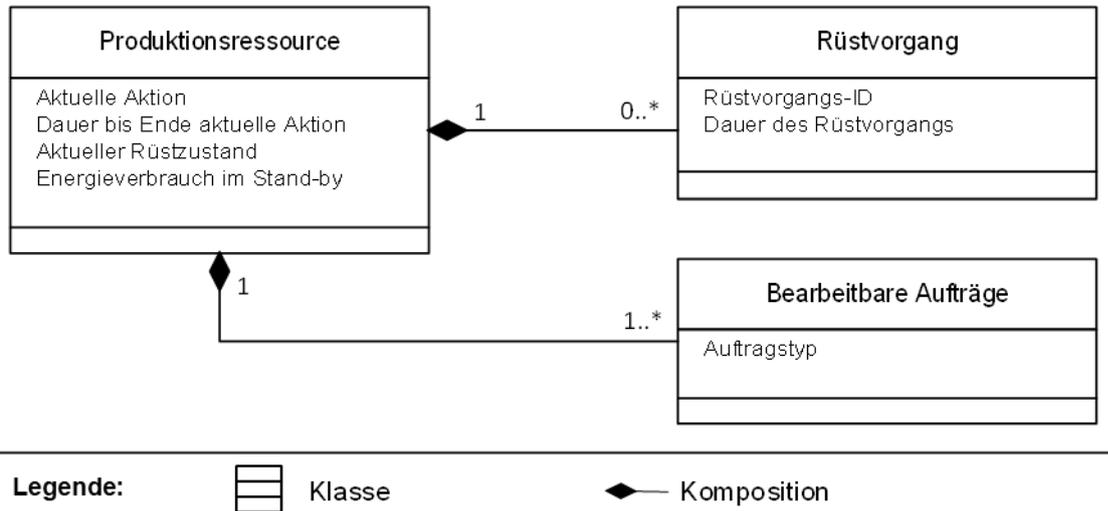


Abbildung 34: Modellierung der Produktionsressourcen in UML-Notation

Hinsichtlich der zeitlichen Modellierung wird eine Unterteilung in einzelne Schichten mit einer definierten Länge T_{max} von beispielsweise acht Stunden vorgenommen. Prinzipiell können Aufträge auch nach Beendigung einer Schicht abgearbeitet werden. Dies ist aber mit einem pauschalen Kostenaufschlag pro zusätzlichem Zeitintervall vom Ende der Schicht bis zur Fertigstellung verbunden. Auf diese Weise werden die Kosten von Überstunden (vgl. Abschnitt 6.4) abgebildet.

7.2.2 Produktionsaufträge

Die Aufgabe der Produktionsressourcen ist es, die von der übergeordneten Produktionsplanung vorgegebenen Aufträge abzuarbeiten. Die einzelnen Aufträge werden in Auftragstypen eingeteilt, die durch die in Tabelle 1 dargestellten Merkmale charakterisiert werden. Neben den konstanten Merkmalen, die für jeden Auftrag eines Typs identisch sind, ist der Fertigstellungstermin ein für jeden Auftrag individueller Parameter. Eine wichtige Grundlage für die zu modellierenden Parameter der Aufträge bildet zudem das in Kapitel 6 entwickelte Datenmodell des gesamtkostenbasierten Zielsystems (Abbildung 32).

Tabelle 1: Parameter der einzelnen Aufträge

Kategorie	Parameter	Einheit	Erklärung
Konstant für jeden Auftrag eines Typs	Bearbeitungs-dauer	min	Für die Fertigstellung des Auftrages benötigte Zeit
	Leistungsbedarf	kW	Wert des durchschnittlichen Leistungsbedarfs über die Bearbeitungsdauer
	Produktionsressource		Produktionsressource, die zur Bearbeitung des Auftrags in der Lage ist
	Erforderlicher Rüstzustand		Zur Bearbeitung des Auftrags auf der Produktionsressource erforderlicher Rüstzustand
	Verzugskosten	€/min	Relative Kosten, die bei Überschreiten des Fertigstellungstermins anfallen
	Lagerkosten	€/min	Relative Kosten, die bei Unterschreiten des Fertigstellungstermins anfallen
	Versandvarianten und Kosten	min; €	Falls unterschiedliche Versandarten vorhanden sind, deren Dauer und Kosten
Auftrags-individuell	Fertigstellungs-termin	min	Zeitpunkt, zu dem der Auftrag fertiggestellt sein muss in Minuten ab Schichtbeginn

7.2.3 Stochastische Ereignisse

Im Produktionsumfeld treten stochastische Ereignisse auf, die es von der Produktionssteuerung auszugleichen gilt (vgl. Abschnitt 1.2). Damit das Agentensystem lernen kann, in derartigen Situationen richtig zu reagieren, ist eine entsprechende Modellierung erforderlich. Unvorhergesehene und damit stochastische Ereignisse können in Anlehnung an NIEHUES (2017) im Produktionsumfeld folgende Auswirkungen nach sich ziehen:

- Zeitliche Abweichung zwischen Ist- und Plan-Verlauf bei der Bearbeitung von Aufträgen (zum Beispiel längere Bearbeitungszeiten aufgrund technischer Probleme im Prozess)
- Temporäre Reduzierung des Kapazitätsangebots (zum Beispiel durch Maschinenausfälle)
- Temporäre Reduzierung oder Erhöhung der Kapazitätsnachfrage (zum Beispiel durch neue ankommende Eilaufträge)

Im Modell des Produktionssystems können diese stochastischen Ereignisse durch eine zufällig auftretende Abweichung der Bearbeitungszeiten, Maschinenausfälle

und Eilaufträge modelliert werden. Zu diesem Zweck kann eine Normalverteilung herangezogen werden, deren Standardabweichung und Mittelwert an die tatsächliche Auftrittswahrscheinlichkeit im realen Produktionssystem angelehnt sind.

7.3 RL-Agenten

Die RL-Agenten sind das zentrale Element der vorliegenden Produktionssteuerung. Im Kontext von RL werden Agenten maßgeblich durch ihren Aktions- und Zustandsraum charakterisiert. Im Hinblick auf das angestrebte kooperierende Verhalten aller Agenten spielt neben der Belohnungsvergabe die Gestaltung des Zustandsraums eine wichtige Rolle. Wenn die Agenten den aktuellen Beobachtungsraum der jeweils anderen Agenten kennen, kann dies eine wichtige Wissensbasis für aufeinander abgestimmte Entscheidungen sein. In den nachfolgenden Unterkapiteln wird in diesem Zusammenhang zunächst auf Aspekte der *allgemeinen Gestaltung des Zustandsraums* eingegangen, die gleichermaßen alle Agenten betreffen. Darauf aufbauend werden die *Aktions- und Zustandsräume der Ressourcen-, Speicher- und des Handelsagenten* vorgestellt.

7.3.1 Allgemeine Gestaltung des Zustandsraums

Der Zustandsraum entspricht der Repräsentation der Umgebungswahrnehmung des Agenten. Zur Einordnung sei angemerkt, dass in der Literatur unter dem Begriff Zustandsraum in Bezug auf Automaten eine Beschreibung von diskreten Zuständen zur Modellierung von Maschinen verstanden wird (MALKIS & BROY 2018). Im Kontext der vorliegenden Arbeit soll der Zustandsraum im Sinne des englischen Begriffs „state space“ verstanden werden, der im Bereich RL die definierte Menge an Informationen beschreibt, die ein RL-Agent bei der Beobachtung seiner Umgebung wahrnimmt (SUTTON & BARTO 2018). Im Bereich von MARL hat der Zustandsraum einen signifikanten Einfluss auf die Ergebnisse, gleichzeitig existiert eine Vielzahl an Modellierungsmöglichkeiten (GABEL 2009). Im Zuge der Erstellung der vorliegenden Arbeit wurden im Hinblick auf die Gestaltung der Zustandsräume der einzelnen Agenten zahlreiche Lösungsansätze auf ihre Eignung hin experimentell untersucht. Dabei wurden unter anderem Feature-Learning-Methoden (LIU ET AL. 2015), basierend auf Autoencoder (LANGE ET AL. 2012), und klassische Verfahren der Merkmalsselektion angewendet (CHANDRASHEKAR & SAHIN 2014). Die nachfolgende Beschreibung spiegelt die Erkenntnisse aus diesen Untersuchungen wider.

Jeder RL-Agent verfügt über einen individuellen Zustandsraum. Darin sind alle Informationen enthalten, die für die Auswahl einer Aktion von Relevanz sein können. Einige Aspekte betreffen jedoch alle Agenten gleichermaßen. Diese umfassen insbesondere die Herangehensweise bei der *Modellierung der zu bearbeitenden Aufträge*, die Gestaltung eines *globalen Beobachtungsraums*, die Berücksichtigung des *bereits geplanten Energieverbrauchs* sowie die Einbeziehung der *Prognose der Energieverfügbarkeit und der Strompreise*. Diese vier Aspekte werden nachfolgend allgemein für alle Agententypen beleuchtet.

Modellierung der zu bearbeitenden Aufträge

Bei den Entscheidungen der RL-Agenten spielen die zu bearbeitenden Aufträge und deren Modellierung eine wichtige Rolle. Zur Beschreibung sind prinzipiell die zwei in Abbildung 35 dargestellten Herangehensweisen möglich.

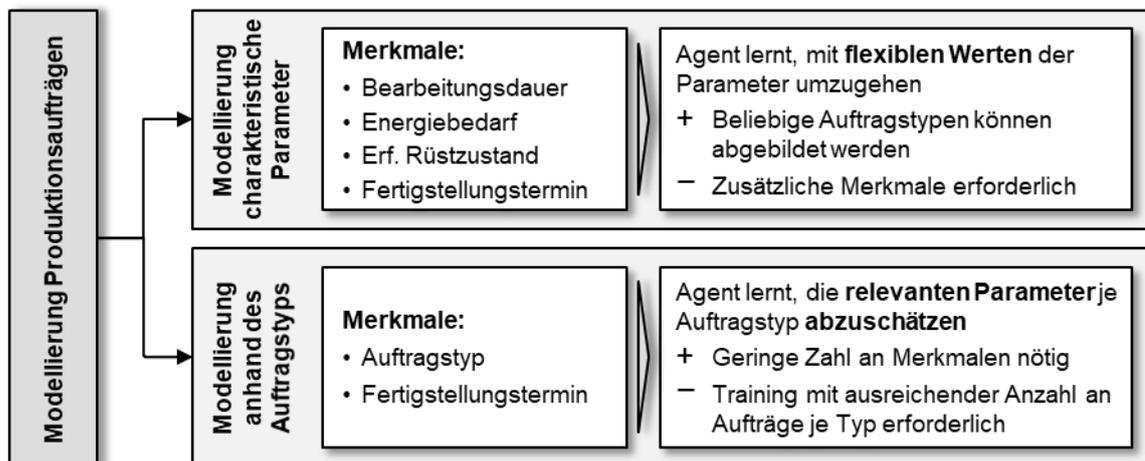


Abbildung 35: Ansätze zur Modellierung von Produktionsaufträge für RL-Agenten

Eine Modellierungsmöglichkeit ist die Anwendung *charakteristischer Parameter*. Dabei werden einzelne Aufträge anhand auftragsindividueller Merkmale beschrieben. Die Parameter werden direkt als diskrete oder kontinuierliche Werte übergeben, was zu einer großen Zahl an Merkmalen für jeden Auftrag führt. Eine alternative Variante bietet die alleinige *Modellierung anhand des Auftragsstyps*, wobei die Agenten keine direkten Informationen zu den charakteristischen Eigenschaften der einzelnen Aufträge erhalten. Wurde eine ausreichende Anzahl von Aufträgen jedes Auftragsstyps während des Trainings bearbeitet, so sind die Agenten in der Lage, durch das erlangte Erfahrungswissen die charakteristischen Eigenschaften abzuschätzen, wie etwa die Bearbeitungsdauer der Auftragsstypen.

Welche Herangehensweise zur Beschreibung der Aufträge auszuwählen ist, hängt maßgeblich von den Gegebenheiten des realen Produktionssystems ab. Wird eine sehr beschränkte Zahl von sich wiederholenden Auftragstypen bearbeitet, so ist eine Modellierung mittels der Auftragstypen vorzuziehen. Denn insgesamt kann bei diesem Ansatz der Zustandsraum der Agenten mit einer geringen Zahl an Merkmalen beschrieben werden, was sich positiv auf die Rechenzeit auswirkt. Die Anzahl der Eingangsmerkmale steigt jedoch kontinuierlich mit jedem zusätzlich möglichen Auftragstyp. Ist die Auftragsvarianz groß oder handelt es sich gar um eine kundenindividuelle Fertigung und jeder Auftragstyp wird einmalig gefertigt, so ist eine Beschreibung auf Basis der charakteristischen Auftragsmerkmale zwingend erforderlich. Dies ist jedoch erst ab einer Anzahl von etwa zehn möglichen Auftragstypen sinnvoll, da zur Beschreibung der Aufträge auf diese Art prinzipiell mehr Merkmale erforderlich sind. Diese steigen jedoch nicht mit zunehmenden Auftragstypen an.

In dem in Kapitel 8 betrachteten Anwendungsfall liegt eine Werkstattfertigung mit sich wiederholenden Auftragstypen vor. Aus diesem Grund wird im weiteren Verlauf ausschließlich auf eine Auftragsbeschreibung auf Basis des Auftragstyps zurückgegriffen. Durch die Anwendung von charakteristischen Auftragsparametern ist eine generische Übertragung des Lösungsansatzes auf Produktionssysteme mit hoher Auftragsvarianz möglich.

Globaler Beobachtungsraum

Aus den eingangs des Abschnitts erwähnten Gründen der Kooperation wird den Ressourcenagenten ein globaler Beobachtungsraum gewährt. Dies ist insbesondere im Kontext des Energieverbrauchs relevant. Hat ein Agent beispielsweise demnächst einen sehr dringenden, energieintensiven Auftrag zu fertigen, so können die anderen Ressourcenagenten etwa mit diesem Wissen ihren Energieverbrauch vorausschauend reduzieren oder der Marktagent kann zusätzlichen Strom kaufen und den erhöhten Bedarf auf diese Weise ausgleichen.

Aus dieser Überlegung heraus wird neben dem eigenen Zustandsraum des Agenten zusätzlich der Zustandsraum aller anderen Agenten übergeben und somit ein globaler Beobachtungsraum gebildet, wie in Abbildung 38 angedeutet ist. Ein Nachteil dieses Ansatzes ist, dass die Zahl der Input-Merkmale für die neuronalen Netze deutlich steigt. Dies macht eine komplexere Netzarchitektur erforderlich und zieht einen erhöhten Rechenaufwand beim Training nach sich. Aus diesem Grund und auf Basis der Ergebnisse der durchgeführten Experimente wird der Zustand der

jeweils anderen Agenten abstrahiert, indem etwa im Kontext der Produktionsressourcen lediglich die Anzahl der bearbeitenden Aufträge und Informationen zur aktuell ausgewählten Aktion übergeben werden. Auf eine detaillierte Aufteilung nach Auftragsstyp wird verzichtet, um auf diese Weise die Anzahl der Eingangsmerkmale zu reduzieren.

Bereits geplanter Energieverbrauch

Im Rahmen der vorliegenden agentenbasierten Produktionssteuerung wird kein deterministischer Produktionsplan berechnet. Vielmehr entscheidet jeder Agent reaktiv nach Beendigung eines Auftrags, welche Aktion als Nächstes auszuführen ist. Aus diesem Grund ist keine deterministische Vorhersage des Energieverbrauchs möglich. Dennoch lässt sich zumindest ein Teil des Strombedarfs im nächsten Zeitintervall abschätzen. Die dazu erforderliche Herangehensweise ist nachfolgend erläutert.

Für die Bearbeitung einzelner Aufträge werden in der Regel mehrere diskrete Zeitschritte benötigt. Daher wählt in einem Zeitschritt lediglich ein Teil der Ressourcenagenten eine neue Aktion aus, während der Rest noch mit der Abarbeitung eines Auftrags beschäftigt ist. Zur Verdeutlichung ist eine derartige Situation beispielhaft in Abbildung 36 dargestellt. Im aktuellen Teilintervall ist ausschließlich der Ressourcenagent 1 an der Reihe, eine neue Aktion auszuwählen, weil die Abarbeitung von Auftrag A nun beendet ist. Die anderen Ressourcenagenten 2 bis 4 arbeiten noch an der Fertigstellung eines Auftrags und wählen erst zu einem späteren Zeitpunkt wieder neue Aktionen aus.

Haben einzelne Agenten noch für mindestens einen Zeitschritt mit der Bearbeitung eines Auftrags zu tun, so lässt sich gemäß der Modellierungsart der Produktionsaufträge auch der dafür erforderliche Energieverbrauch für den nächsten Zeitschritt deterministisch vorhersagen. Im Sinne des kooperierenden Verhaltens kann diese Informationen also verwendet werden, um dem Teil der Agenten, die im jeweiligen Zeitschritt eine neue Aktion auswählen, Informationen über die bereits sicher verplante Energiemenge zur Verfügung zu stellen. Konkret wird die Zahl der Agenten übergeben, die im aktuellen Zeitschritt keine neue Aktion auswählen, sowie die Summe des bereits eingeplanten Leistungsbedarfs, der sich aus den einzelnen Bedarfen der in Bearbeitung befindlichen Aufträge zusammensetzt. In dem Beispiel aus Abbildung 36 wählen aktuell drei Agenten keine Aktionen aus: Ressourcenagent 2, 3 und 4. Der bereits eingeplante Leistungsbedarf setzt sich in diesem Fall aus dem Energiebedarf von Auftrag B, C und D zusammen. Bei den im Zuge dieser Arbeit durchgeführten Simulationsstudien hat sich gezeigt, dass diese

Berücksichtigung des bereits eingeplanten Energieverbrauchs zur signifikanten Verbesserung der Ergebnisse beiträgt.

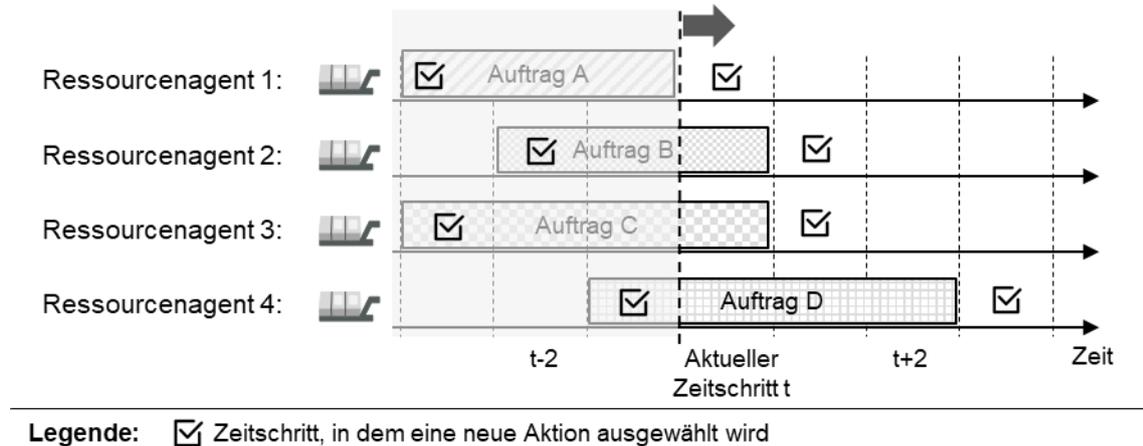


Abbildung 36: Beispielhaftes Szenario der Auswahl von Aktionen durch die Ressourcenagenten

Prognose der Energieverfügbarkeit und Strompreise

Im Hinblick auf den Energieverbrauch sollen die Agenten vorausschauend agieren (vgl. Abschnitt 3.4). Zu diesem Zweck ist erforderlich, dass auch die zukünftige Entwicklung der Energieverfügbarkeit aus bereits gekaufter Energiemenge und Eigenerzeugung in die Entscheidungen aller Agenten einfließt. Anhand einer Preisprognose kann zudem der zu erwartende Strompreis am Intraday-Markt abgeschätzt werden. Aus diesem Grund werden im Zustandsraum nicht nur die Merkmale des aktuellen Hauptintervalls, sondern auch die zukünftigen Werte dieser drei Zeitreihen abgebildet.

Um die Anzahl der zu übergebenden Merkmale zu reduzieren, werden dabei, wie in Abbildung 37 beispielhaft für die Erzeugungsprognose dargestellt, 15-Minuten-Werte ab einem Zeithorizont von einer Stunde stufenweise zusammengefasst. Insgesamt ergeben sich auf diese Weise jeweils acht Merkmale, deren Aussagekraft bis vier Stunden in die Zukunft reicht. Alle drei Zeitreihen sind mit Unsicherheiten behaftet. Die Erzeugungsleistung von EE kann nicht exakt vorhergesagt werden. Die bereits gekaufte Energiemenge kann sich kurzfristig aufgrund des Markagenten ändern. Auch die Vorhersehbarkeit der Strompreise ist begrenzt. Die Agenten müssen demnach im Zuge des Trainings lernen, trotz derartiger Abweichungen ihre Strategie zu optimieren.

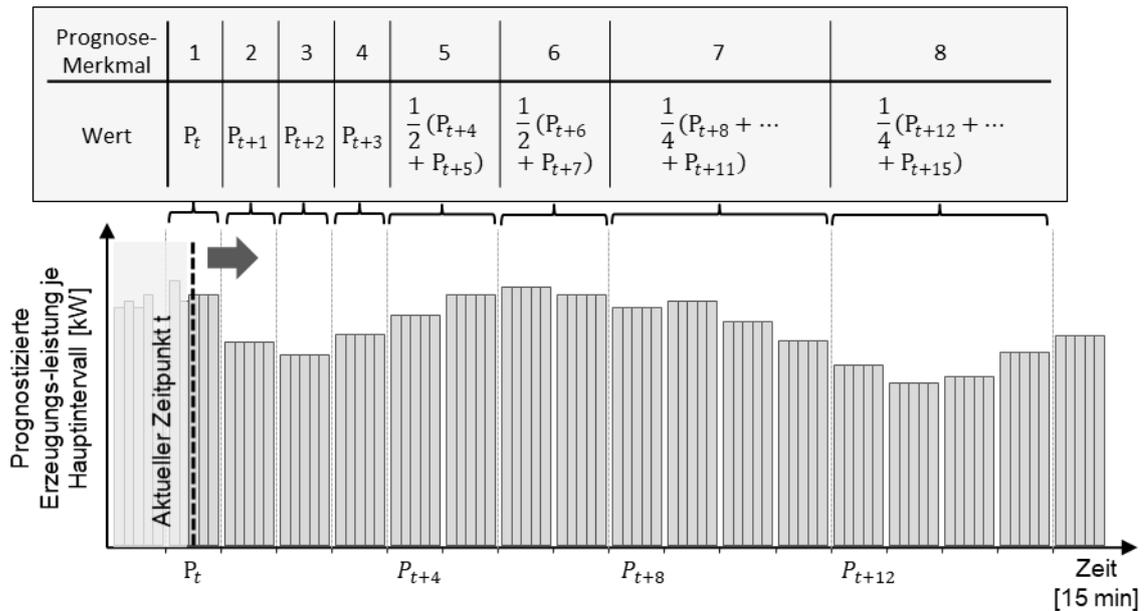


Abbildung 37: Merkmale zur Beschreibung der drei Zeitreihen der Energieverfügbarkeit am Beispiel der prognostizierten Erzeugungsleistung

7.3.2 Aktions- und Zustandsraum der Ressourcenagenten

Zustandsraum

Die zentrale Aufgabe der Ressourcenagenten ist es, die in der Warteschlange anstehenden Aufträge abzuarbeiten. Aus diesem Grund spielen diese eine wichtige zentrale Rolle bei der Gestaltung des Zustandsraums, wobei auf die in Abschnitt 7.3.1 beschriebene Modellierung anhand des Auftragsstyps zurückgegriffen wird. Jeder Ressourcenagent erhält mit den Informationen zu der *Anzahl der zu fertigenden Aufträge je Auftragsstyp* sowie den jeweils *nächsten drei Fertigstellungsterminen* ein detailliertes Bild über den eigenen Zustand. Im Kontext der Informationen zu den einzelnen Aufträgen ist zudem der *aktuelle Rüstzustand* der Produktionsressource relevant. Weil die Anzahl der erforderlichen Rüstvorgänge maßgeblich von der Bearbeitungsreihenfolge abhängt, kann der Agent diesen Aspekt ebenfalls bei der Entscheidungsfindung berücksichtigen. Indem der aktuelle Rüstzustand als Input in den Zustandsraum übergeben wird, kann der Agent im Laufe des Trainings lernen, mit dem Rüstzustand verbundene Folgen abzuschätzen und diese in seinen Entscheidungen zu berücksichtigen. Die zugrundeliegende Logik eignet sich der Agent dabei selbstständig auf Basis der im Training gesammelten Erfahrungen.

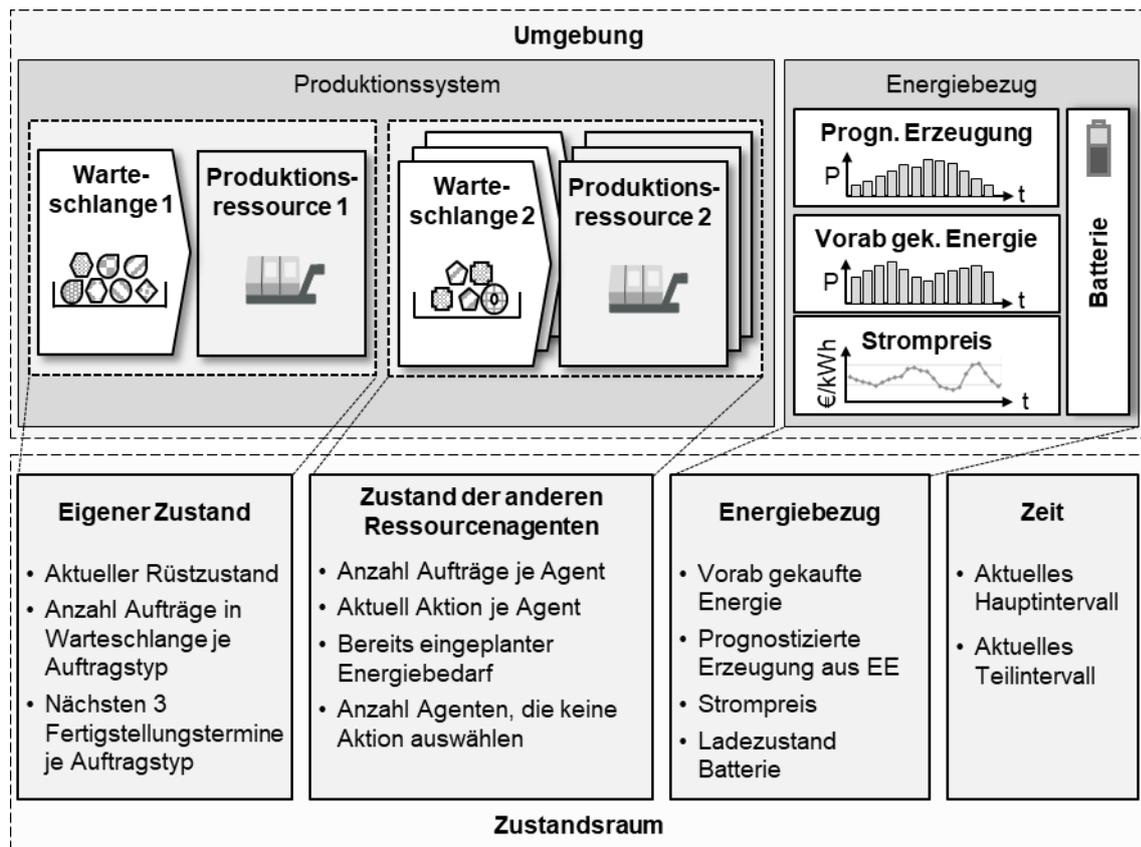


Abbildung 38: Zustandsraum eines Ressourcenagenten am Beispiel von Ressourcenagent 1

Daneben erhält jeder einzelne Ressourcenagent im Sinne des in Abschnitt 7.3.1 beschriebenen globalen Beobachtungsraums auch Einblicke in den aktuellen Zustand der übrigen Ressourcenagenten, indem jeweils die *aktuelle Aktion jedes Agenten* und die *Anzahl der ausstehenden Aufträge* übergeben werden. Zudem kann der *bereits geplante Energieverbrauch* zusammen mit der *Anzahl der Agenten, die keine Aktion auswählen*, da sie im aktuellen Zeitschritt noch mit der Bearbeitung eines Auftrags beschäftigt sind, abgeschätzt werden. Von Seiten des Energiebezugs bilden die *bereits vorab gekaufte Energiemenge*, die *prognostizierte Erzeugung der EE* und die *Strompreisprognose* weitere Eingangsmerkmale, deren Zeitreihen jeweils in der in Abschnitt 7.3.1 beschriebenen Form mit über die Zeit abnehmender Granularität einfließen. Neben diesen Informationen zur zukünftigen Energieverfügbarkeit ist zudem der *aktuelle Ladezustand der Batterie* eine wichtige Information. Dieser Wert beeinflusst maßgeblich die Möglichkeit des Gesamtsystems, kurzfristige Abweichungen zwischen Stromangebot und -nachfrage durch die Batterie auszugleichen.

Insgesamt ergibt sich damit für die Ressourcenagenten ein globaler Zustandsraum, in dem alle relevanten Informationen zusammengeführt sind. Dieser setzt sich im Wesentlichen aus den Informationen zum eigenen Zustand, zum Zustand der anderen Ressourcenagenten sowie aus dem Energiebezug und Zeitverlauf zusammen. Die dafür erforderlichen Merkmale sind in Abbildung 38 abstrahiert zusammengefasst. Bei einer Implementierung in einem System mit fünf Ressourcenagenten sind für diese Modellierungsweise insgesamt mehr als 80 Eingangsmerkmale erforderlich.

Aktionsraum

Der Aktionsraum spannt die Menge an möglichen diskreten Aktionen auf, die den Ressourcenagenten zur Auswahl stehen, um in der Umgebung zu agieren und ihre Ziele zu erreichen (WANNAGAT & VOGEL-HEUSER 2008). Nach erfolgreichem Abschluss einer Aktion können die Produktionsressourcen aus folgenden Aktionen auswählen:

- **Stand-by:** Diese Aktion weist eine feste Dauer von einem Zeitschritt und einem definierten Energieverbrauch (vgl. Abschnitt 7.2.1) auf. Ein Agent kann jederzeit Stand-by auswählen.
- **Bearbeitung eines Auftragsstyps:** Eine Produktionsressource kann die Bearbeitung eines Auftragsstyps wählen. Dabei wird jeweils der nächste fertigzustellende Auftrag des ausgewählten Auftragsstyps bearbeitet. Befindet sich kein Auftrag eines Typs in der Warteschlange und wird dieser dennoch vom Agenten zur Bearbeitung ausgewählt, führt dies zu einer „Bad Action“, die für den Agenten mit hohen Strafkosten verbunden ist. Ein Ressourcenagent kann somit zu jedem Zeitpunkt aus allen Aktionen wählen. In der jeweiligen Situation unzulässige Aktionen werden über die Belohnungsfunktionen bestraft. Der Agent lernt dadurch im Training, unzulässige Aktionen zu vermeiden.

Die Aktion *Rüsten* kann von der Produktionsressource nicht bewusst ausgewählt werden. Vielmehr wird Rüsten automatisch vor Bearbeitungsbeginn durchgeführt, wenn der aktuelle Rüstzustand von dem für die Bearbeitung des ausgewählten Auftrags erforderlichen Rüstzustand abweicht. Die Fertigstellungsdauer eines Auftrags setzt sich somit aus der Summe der erforderlichen Rüstzeit und der Bearbeitungsdauer zusammen.

7.3.3 Aktions- und Zustandsraum des Speicheragenten

Zustandsraum

Der Speicheragent ist dem ad-hoc Strombezug zuzuordnen und legt die Lade- beziehungsweise Entladeleistung des Speichers fest. Dementsprechend sind im Zustandsraum des Speicheragenten alle Informationen zu akkumulieren, die einen Einfluss auf die Steuerungsstrategie des Speichers nehmen.

Darüber hinaus ist neben der verfügbaren Energiemenge der aktuelle und zukünftige Energiebedarf des Produktionssystems von großer Relevanz. Da die Ressourcenagenten ihre Aktionen reaktiv nach jedem abgeschlossenen Auftrag neu auswählen, existiert kein deterministischer Produktionsplan. Damit ist auch keine exakte Vorhersage des zukünftigen Energieverbrauchs möglich. Um dem Speicheragenten dennoch eine möglichst detailreiche Berücksichtigung des Energiebedarfs zu ermöglichen, werden zwei Ansätze verfolgt. Einerseits kann der Energiebedarf zumindest im aktuellen Zeitschritt exakt bestimmt werden – vorausgesetzt, alle Ressourcenagenten haben bereits für den jeweiligen Zeitschritt ihre Aktionen ausgewählt. Aus diesem Grund wird der Ablauf des Systems so gestaltet (vgl. Abschnitt 5.3.1), dass zu Beginn eines neuen Zeitschritts zunächst die Ressourcenagenten parallel an der Reihe sind, eine neue Aktion auszuwählen und erst im Anschluss der Batterieagent aktiv wird. Somit sind zum Zeitpunkt der Aktionsauswahl des Batterieagenten bereits die Aktion der Ressourcenagenten und damit auch der daraus resultierende *Gesamtenergieverbrauch im aktuellen Zeitschritt* bekannt. Andererseits wird der Speicheragent neben dieser exakten Verbrauchsbestimmung für den aktuellen Zeitschritt in die Lage versetzt, durch geeignete Merkmale den demnächst anstehenden Energiebedarf der Ressourcenagenten abschätzen zu können. Zu diesem Zweck wird dem Speicheragenten für jeden Ressourcenagenten der *summierte Energiebedarf aller offenen Aufträge* und aller *innerhalb der nächsten Stunde zu fertigenden Aufträge* übergeben. Ein zusätzlicher Input sind die *Summen des Zeitbedarfs zur Fertigung aller* und lediglich der *innerhalb der nächsten Stunde zu bearbeitenden Aufträge*, da dadurch etwa der Spielraum der Ressourcenagenten für Stand-by-Aktionen abgeschätzt werden kann.

Von großer Relevanz ist dabei die verfügbare Energiemenge, die sich sowohl aus dem *bereits vorab gekauften Strom* sowie der *Erzeugungsprognose aus EE* und der *Strompreisprognose* zusammensetzt. Analog zu dem in Abbildung 37 dargestellten Verfahren im Kontext der Ressourcenagenten erfolgt auch im Zuge des Speicheragenten eine stufenweise Abbildung der Zeitreihen mit abnehmender Granularität. Von Seiten des Energiebezugs ist außerdem der aktuelle Ladezustand des

Speichers von essenzieller Bedeutung. Insgesamt ergibt sich daraus der in Abbildung 39 dargestellte Zustandsraum des Speicheragenten.

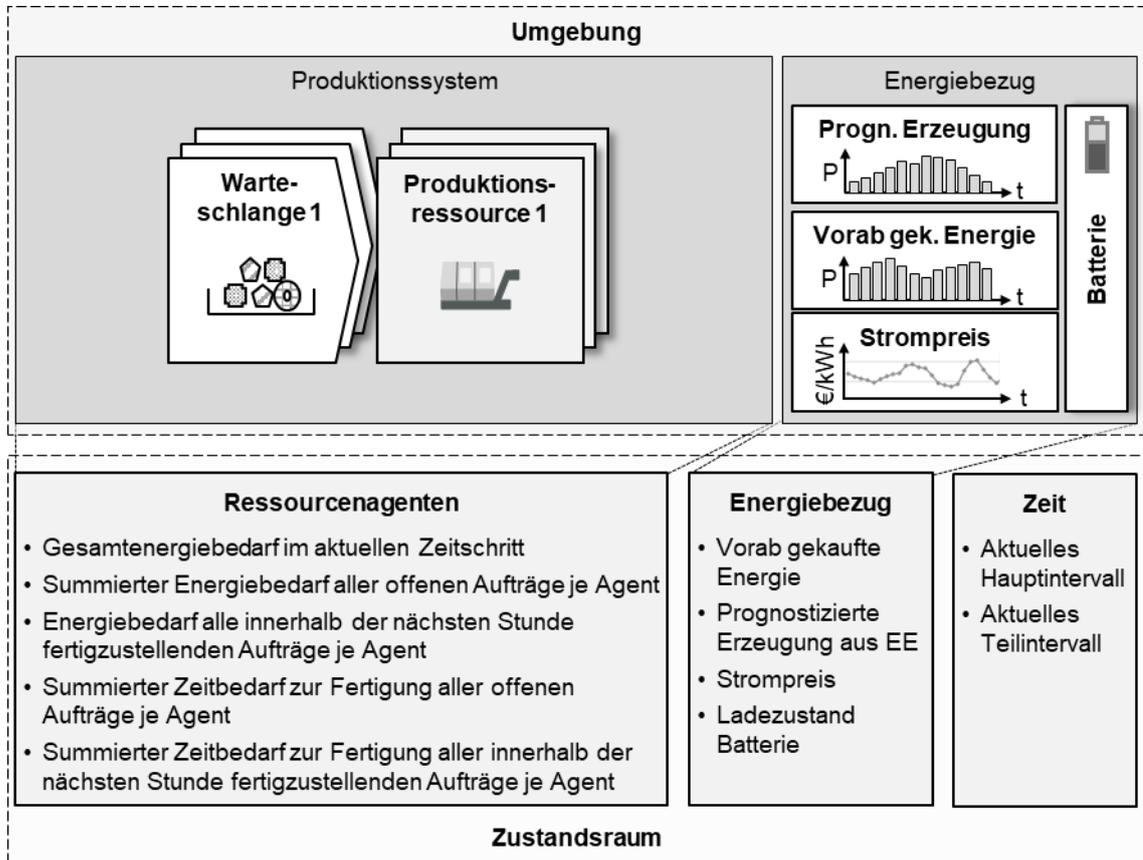


Abbildung 39: Zustandsraum des Speicheragenten

Aktionsraum

Die möglichen Aktionen des Speicheragenten werden in Form eines diskreten Aktionsraums modelliert, der sich aus insgesamt 51 Werten zusammensetzt. Die Wahl ist ein Kompromiss zwischen einer möglichst variablen Lade- und Entladeleistung einerseits und der zunehmenden Komplexität für das Training durch einen größer werdenden Aktionsraum andererseits. Die von der C-Rate (vgl. Abschnitt 2.2.3) bestimmte maximale und minimale Lade- beziehungsweise Entladeleistung ist der erste und letzte diskrete Wert des Aktionsraums. Die übrigen 49 Intervalle werden linear zwischen diesen beiden Leistungswerten verteilt. Der 26. Wert des Aktionsraums entspricht somit einem Ladestrom von 0 kW.

Falls der Batterieagent die Kapazität des Batteriespeichers überschreitet und beispielsweise versucht, mehr Strom zu entladen, als aktuell im Speicher vorhanden

ist, so wird eine „Bad Action“ ausgegeben. Die damit einhergehenden Strafkosten in der Belohnung bringen den Batterieagenten während des Trainings dazu, derartige Aktionen zu unterlassen.

7.3.4 Aktions- und Zustandsraum des Marktagenten

Zustandsraum

Der Marktagent regelt den kurzfristigen Stromhandel und hat die Aufgabe, gewinnbringend den Zeitpunkt und die Höhe eines Stroman- oder -verkaufs zu steuern. Ähnlich wie im Kontext des Speicheragenten sind daher der zukünftige Energieverbrauch und die Energieverfügbarkeit von Bedeutung. Im Vergleich dazu agiert der Marktagent aufgrund der erforderlichen Vorlaufzeit für den Stromhandel vergleichsweise weiter in der Zukunft, da bereits neun Minuten vor Anbruch eines neuen Hauptintervalls über den Kauf/Verkauf von Energie entschieden werden muss (vgl. Abschnitt 5.2.5).

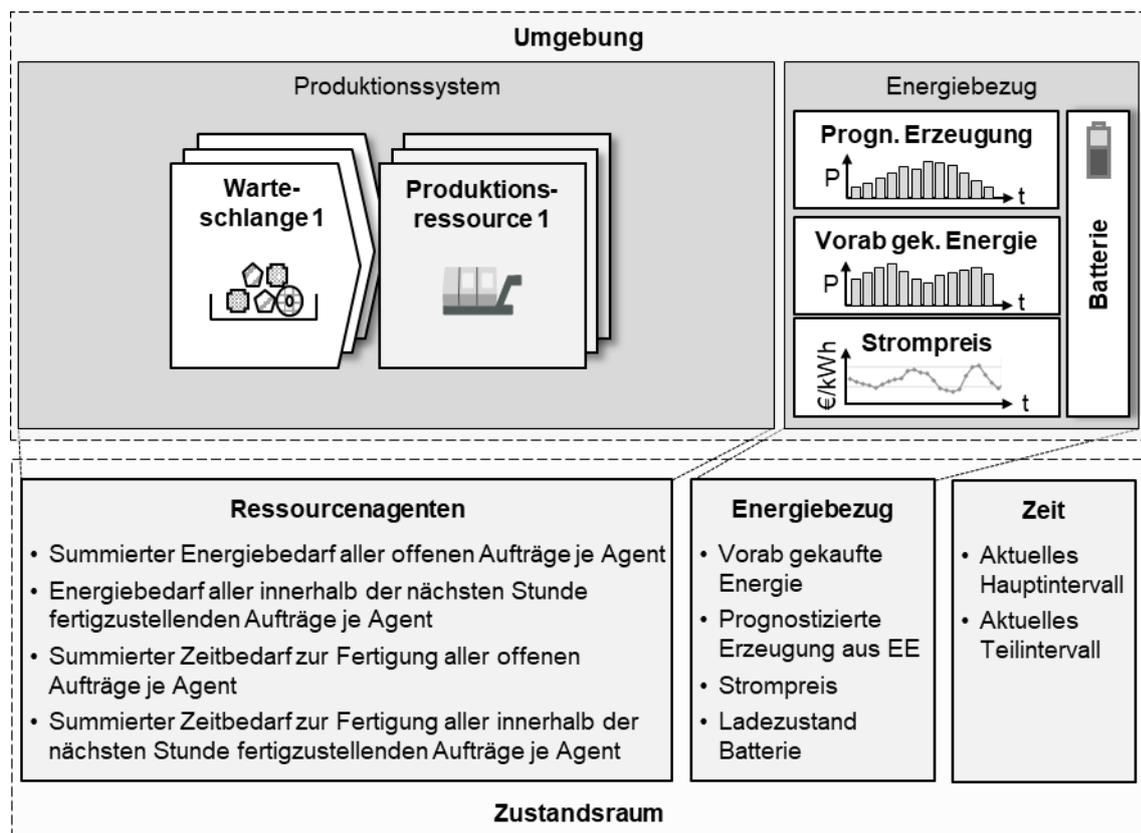


Abbildung 40: Zustandsraum des Marktagenten

Demzufolge ist der Aufbau des Zustandsraums sehr ähnlich zu dem des Speicheragenten, wobei lediglich die Elemente zur Beschreibung des kurzfristigen Energieverbrauchs fehlen. Die Energieverfügbarkeit aus vorab gekaufter Energiemenge und Erzeugungsprognose sowie die Strompreise werden wie im Zuge des Speicheragenten (vgl. Abschnitt 7.3.3) einbezogen. Hervorzuheben ist zudem der Ladezustand des Speichers, da dies ein Zusammenspiel zwischen Stromspeicher und kurzfristigem Stromhandel ermöglicht. Dadurch können Strompreisschwankungen ausgenutzt werden. Die Abschätzung des Energieverbrauchs erfolgt mit der Berechnung des Gesamtenergieverbrauchs aller Aufträge und der innerhalb der nächsten Stunde fertigzustellenden Aufträge ebenfalls analog zum Speicheragenten. Die resultierende Zusammensetzung des Zustandsraums des Marktagenten ist in Abbildung 40 dargestellt.

Aktionsraum

Der Aktionsraum des Marktagenten wird analog zu dem des Speicheragenten gestaltet. Auch hier wird ein Aktionsraum mit insgesamt 51 diskreten möglichen Schritten aufgespannt, die linear zwischen einer maximal zu kaufenden und einer maximal zu verkaufenden Energiemenge verteilt werden. Die Höhe der beiden Grenzwerte ergibt sich aus den maximal sinnvollen Handelsvolumina des Marktagenten im Kontext des Gesamtsystems. Zudem muss es sich nicht zwangsläufig um einen Aktionsraum handeln, der den Kauf und Verkauf von Strom symmetrisch repräsentiert.

Falls keine Rückspeisung von Strom aus Eigenerzeugung ins Netz möglich ist, darf der Marktagent lediglich den Strom verkaufen, der vorab am Strommarkt beschafft wurde. Diese Regelung wird erneut durch eine „Bad Action“ abgebildet. Die negative Belohnung daraus animiert den Marktagenten dazu, ein derartiges Handeln zu unterlassen. Bei der Gestaltung des Aktionsraums des Marktagenten sind darüber hinaus regulatorische Randbedingungen zu beachten. Sind etwa Mindestmengen für den Handel von Strom oder Vorgaben zur kleinsten handelbaren Einheit vorgegeben, so muss der Aktionsraum entsprechend diesen Anforderungen angepasst werden.

7.4 Belohnungsfunktion und -vergabe

Neben der Modellierung des Zustandsraums bilden die Belohnungsfunktion und deren Vergabe das zweite kritische Element bei der Anwendung von MARL (DEWEY 2014). Daher wird nachfolgend zunächst die Bedeutung einer *globalen und lokalen Belohnung* im Kontext des vorliegenden Systems verdeutlicht. Danach wird mit den *Energie- und Produktionskosten* auf die beiden zentralen Elemente der Belohnungsfunktion eingegangen. Anschließend erfolgt eine Zusammenfassung der *resultierenden Belohnungsvergabe*.

7.4.1 Globale und lokale Belohnung

Aus dem vorliegenden Anwendungsfall der energieorientierten Produktionssteuerung resultiert eine komplexe Belohnungsfunktion, die im Wesentlichen in die beiden Bestandteile Energiekosten und Produktionskosten untergliedert werden kann. Erstere gehen aus dem Kostenmodell der unterschiedlichen Energiebezugsoptionen in Kapitel 5 hervor. Letztere sind das Ergebnis des gesamtkostenbasierten Zielsystems in Kapitel 6.

Die Gestaltung der Belohnungsfunktion hat darüber hinaus maßgeblichen Einfluss auf das kooperierende Verhalten der einzelnen Agenten. Im Falle einer globalen Belohnung erhalten alle Agenten gemeinsam die identische Belohnung und lernen dadurch, mit Hilfe einer gemeinsamen Kooperation die Belohnung zu maximieren (vgl. Abschnitt 2.5.4). Demgegenüber steht die Möglichkeit einer lokalen Belohnung, die lediglich an einzelne Agenten zurückgegeben wird. Bei der Vergabe einer pauschalen, globalen Belohnung stehen die einzelnen Agenten vor der Herausforderung, zur Optimierung der Strategie ihren eigenen Einfluss auf die resultierende Belohnung herausfiltern zu müssen. Dieses Problem wird in der Literatur als „*Credit Assignment Problem*“ bezeichnet (NGYEN ET AL. 2018) und ist in Abbildung 41 beispielhaft im Kontext des vorliegenden Systems demonstriert.

Angenommen, die gemeinsame Belohnung aller Agenten fällt in einer Schicht aufgrund eines zu spät fertiggestellten Auftrags schlecht aus. Im Zuge des Lernprozesses stehen nun alle Agenten vor dem Zuordnungsproblem, welche der durchgeführten Aktionen die negative Belohnung verursacht hat. Am Beispiel des Marktagenten wird deutlich, dass dieser durch seine Aktionen die negative Belohnung nicht beeinflussen konnte. Dennoch erhält der Marktagent diesen Teil der Belohnung und muss seine Strategie auf dieser Basis verbessern, ohne seine Aktionen

direkt der negativen Belohnung zuordnen zu können. Dies behindert den Lernfortschritt. Aus diesem Grund wird angestrebt, lediglich die Faktoren als Bestandteil der Belohnung eines Agenten zu modellieren, die der Agent maßgeblich durch seine Aktionen beeinflussen kann. Alle anderen sogenannten exothermen Elemente der Belohnungsfunktion sind zu entfernen, da diese das Trainingsergebnis und die Lösungsqualität in Summe verschlechtern können (DIETTERICH ET AL. 2018). Im diskutierten Beispiel hat lediglich der Ressourcenagent die verspätete Fertigstellung zu verantworten. Daher wird nur dieser die entsprechende Belohnung erhalten.

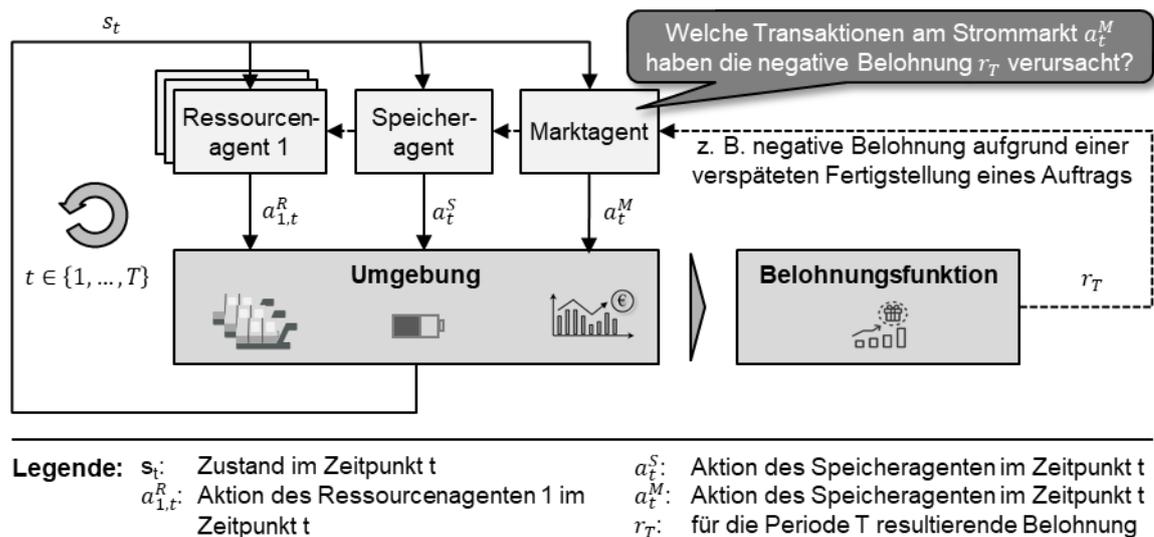


Abbildung 41: Beispielhafte Darstellung des Credit-Assignment-Problems

Im vorliegenden System beeinflussen alle Agenten die Energiekosten direkt. Die Produktionskosten werden hingegen lediglich von den jeweiligen individuellen Ressourcenagenten bestimmt. So wird beispielsweise die verzögerte Fertigstellung eines Auftrags, die zu hohen Strafkosten führt, von dem zuständigen Ressourcenagenten verursacht. Weder die anderen Ressourcenagenten noch Speicher- oder Marktagent können den Bearbeitungszeitpunkt eines Auftrags durch ihre Aktionen direkt beeinflussen. Aus diesem Grund wird die Belohnungsfunktion in der vorliegenden Arbeit in zwei Bestandteile zerlegt: die globalen Energiekosten und die lokalen Produktionskosten.

7.4.2 Energiekostenbestandteil der Belohnung

Die Energiekosten werden gleichermaßen von allen Agenten beeinflusst und daher entsprechend an alle übermittelt. Es handelt sich demnach um einen globalen Belohnungsteil. Die Ausgaben für Energie fallen jeweils zum Ende eines 15-minütigen Hauptintervalls oder beim Kauf beziehungsweise Verkauf von Strom an. Da sich etwa Produktionsaufträge über mehrere Hauptintervalle hinweg erstrecken können, ist bei Ressourcenagenten nicht immer eine direkte Zuordnung der resultierenden Energiekosten zu einer bestimmten Aktion möglich. Aus diesem Grund erhalten alle Agenten die kumulierten Energiekosten als Belohnung am Ende einer Produktionsschicht. Die Energiekosten setzen sich aus den in Abschnitt 5.2 hergeleiteten Kosten für Eigenerzeugung, kurzfristigen Netzbezug, Stromhandel und Batteriealterung zusammen. Darüber hinaus lernen RL-Agenten, die Belohnung zu maximieren. Das Ziel der vorliegenden Arbeit ist die Minimierung der Kosten. Ein RL-Agent lernt dagegen, eine Belohnung zu maximieren. Aus diesem Grund wird durch die Vergabe von negativen Kosten aus dem Minimierungs- ein Maximierungsproblem geschaffen:

$$R_s^E = - \left(\sum_{t=0}^T K_{BHKW}(t) + K_{AE}(t) + K_{LS}(t) + K_I(t) + K_{Bat,i} \right) \quad (20)$$

R_s^E	Energiekostenbestandteil der Belohnung für die Schicht s
T	Anzahl der Hauptintervalle in einer Schicht
$K_{BHKW}(t)$	Stromkosten des BHKW im 15-Minuten-Intervall t
$K_{AE}(t)$	Kosten für vom Netz bezogene Ausgleichsenergie im Intervall t
$K_{LS}(t)$	Kosten für Lastspitze im Intervall t
$K_I(t)$	Kosten für Strom aus Intraday-Handel im Intervall t
$K_{Bat,i}$	Resultierende Alterungskosten der Batterie pro Zyklus i

7.4.3 Produktionskostenbestandteil der Belohnung

Die Produktionskosten eines Auftrags werden ausschließlich durch denjenigen Agenten bestimmt, der die Bearbeitung des Auftrages übernimmt. Deswegen wird die entsprechende Belohnung in Form der Produktionskosten vergeben, wenn ein Ressourcenagent eine Aktion abgeschlossen hat. Es handelt sich um einen lokalen Belohnungsteil, der unmittelbar nach Abschluss einer Aktion eines Ressourcenagenten ausgeschüttet wird.

Zur ganzheitlichen Berechnung der von der Produktionssteuerung verursachten Produktionskosten wird die in Abschnitt 6.4 hergeleitete Formel (19) verwendet. Im Kontext der vorliegenden Arbeit ergeben sich mehrere Besonderheiten, die nachfolgend unter den Stichworten *Verzugskosten von Aufträgen in der Warteschlange* und *relative Belohnungsvergabe* erläutert werden. Darauf aufbauend erfolgt die mathematische Formulierung der *Belohnungsfunktion für die Produktionskosten*.

Verzugskosten von Aufträgen in der Warteschlange

Um das Lernverhalten der Agenten zu verbessern, wird die Berechnung der terminbezogenen Kosten modifiziert. Die resultierenden Verzugskosten eines Auftrags werden im vorliegenden Lösungsansatz nicht erst nach entsprechender Fertigstellung des jeweiligen Auftrags vergeben. Verzugskosten werden vielmehr fällig, wenn nach Abschluss einer Aktion der späteste erforderliche Bearbeitungsbeginn bereits überschritten wurde, da zu diesem Zeitpunkt schon der Verzug eines Auftrags absehbar ist.

Zur Verdeutlichung dieses Ansatzes ist in Abbildung 42 ein beispielhaftes Szenario einer Produktionsressource dargestellt, bei dem die Aufträge A und B innerhalb der vorgegebenen Fertigstellungstermine zu bearbeiten sind. Im aktuellen Zeitschritt t wurde Auftrag A fertiggestellt. Der Ressourcenagent bekommt demnach eine entsprechende Belohnung und wählt eine neue Aktion aus. In dem dargestellten Szenario wurde *Auftrag A* rechtzeitig fertiggestellt und verursacht keine Verzugskosten. Des Weiteren muss *Auftrag B* zum *Zeitpunkt $t+2$* fertiggestellt werden. Aufgrund der dafür erforderlichen Bearbeitungsdauer von vier Zeitschritten ist dies jedoch aktuell nicht mehr möglich. Der Bearbeitungsbeginn hätte spätestens zum *Zeitpunkt $t-1$* erfolgen müssen. Dementsprechend erhält der Agent in diesem Fall im *Zeitschritt t* eine negative Belohnung aufgrund der bereits aktuell absehbaren Verspätung von *Auftrag B* um mindestens einen Zeitschritt. Somit erhält der Agent bereits eine Bestrafung, wenn der Fertigstellungszeitpunkt noch nicht überschritten, aber die Verspätung eines Auftrags bereits absehbar und unausweichlich ist. Die Simulationsstudien im Zuge der vorliegenden Arbeiten haben gezeigt, dass durch diese frühzeitige Vergabe einer negativen Belohnung, falls sich eine Verspätung abzeichnet, die Qualität der Ergebnisse deutlich verbessert wird. Die Agenten lernen auf diese Weise schneller, eine Terminüberschreitung von Fertigungsaufträgen zu vermeiden.

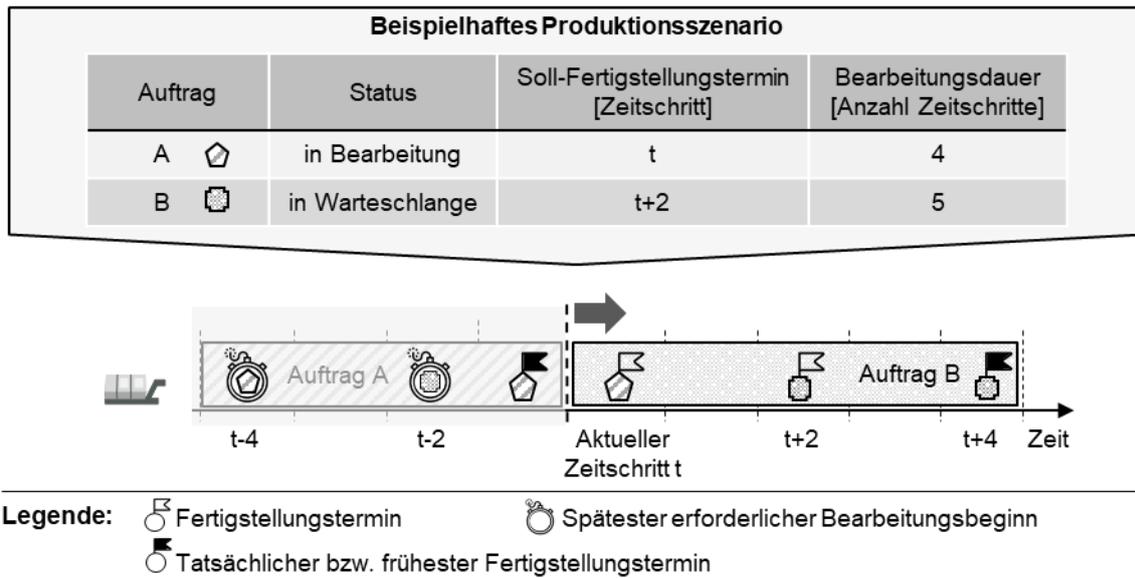


Abbildung 42: Beispielszenario für die Belohnungsvergabe bei Fertigstellungszug

Dementsprechend werden für alle Aufträge, die sich in der Warteschlange einer Produktionsressource befinden und deren Fertigstellungstermin im aktuellen Zeitschritt aufgrund der erforderlichen Bearbeitungsdauer absehbar überschritten wird, Verzugskosten berechnet. Zu diesem Zweck wird die Bearbeitungsdauer auf den aktuellen Zeitschritt aufgeschlagen, um so einen absehbaren Verzug zu identifizieren. Nachfolgend ist die mathematische Formulierung dieser Funktion dargestellt:

$$WK_{n_p}(t) = \begin{cases} TK_{n_p}(t + d_{n_p}) & , t + d_{n_p} \geq t_{F,n_p} \\ 0 & , t + d_{n_p} < t_{F,n_p} \end{cases} \quad (21)$$

$WK_{n_p}(t)$ Terminbezogene Kosten des Auftrags n_p in der Warteschlange

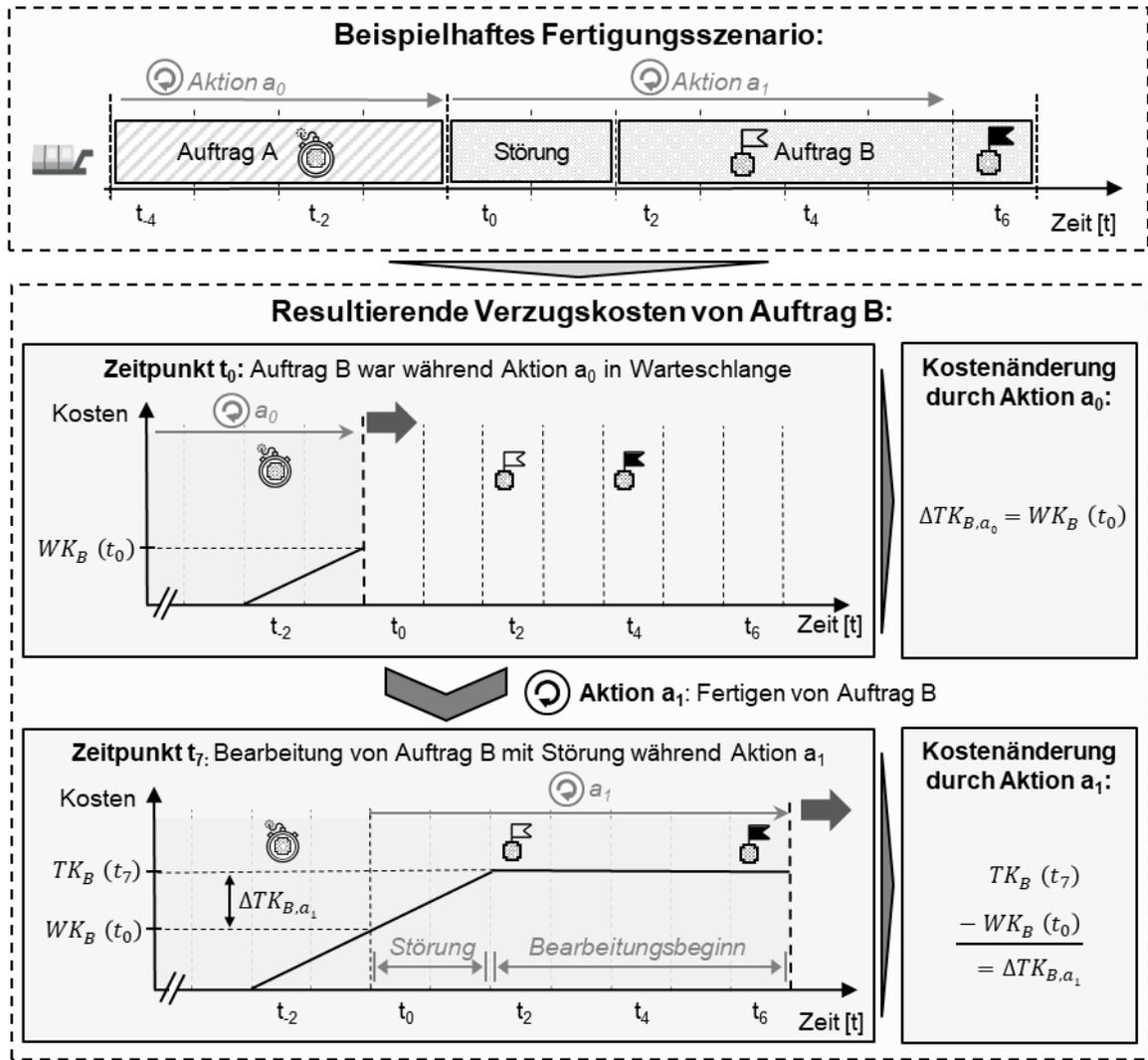
$TK_{n_p}(t)$ Terminbezogene Kosten von Auftrag n_p (siehe Formel (16))

d_{n_p} Erforderliche Bearbeitungsdauer von Auftrag n_p

t_{F,n_p} Fertigstellungstermin von Auftrag n_p

Relative Belohnungsvergabe

Durch die kontinuierliche Berechnung der Verzugskosten von Aufträgen, die sich in der Warteschlange befinden, ist es möglich, dass ein Ressourcenagent bei verspäteter Fertigstellung eines Auftrags bereits Teile der daraus resultierenden Belohnung erhalten hat.



- | | |
|---|---|
| <p>Legende:</p> <ul style="list-style-type: none"> Soll-Fertigstellungstermin von Auftrag B Spätester erforderlicher Bearbeitungsbeginn von Auftrag B Frühester bzw. tatsächlicher Fertigstellungstermin von Auftrag B | <ul style="list-style-type: none"> $WK_B(t)$: Terminbezogene Kosten von Auftrag n in der Warteschlange $TK_n(t)$: Terminbezogene Kosten bei Fertigstellung von Auftrag n $\Delta TK_{n,a}$: Veränderung der terminbezogenen Kosten von Auftrag n durch Aktion a |
|---|---|

Abbildung 43: Berechnung der relativen Belohnung bei Vergabe der terminbezogenen Kosten von Aufträgen in der Warteschlange

Zur Verdeutlichung ist eine derartige Konstellation beispielhaft in Abbildung 43 dargestellt. Dabei wird das in Abbildung 42 eingeführte Produktionsszenario erneut aufgegriffen. Der alleinige Betrachtungsfokus liegt nun auf der Kostenentwicklung von Auftrag B. Nach Beendigung der Aktion a_0 im Zeitschritt t_0 ist Auftrag A fertig. Der Soll-Fertigstellungstermin von Auftrag B kann zu diesem Zeitpunkt nicht mehr gehalten werden. Aufgrund der Fertigungsdauer von fünf Zeit-

schritten liegt die Verspätung bei mindestens zwei Zeitschritten. Bei Annahme einer linearen Kostenfunktion zieht dies die Kosten $WK_B(t_o)$ nach sich, die dem Agenten als Belohnung zugewiesen werden.

Der Agent wählt mit der Aktion a_1 die Bearbeitung von Auftrag B aus, kann diesen aufgrund einer unerwarteten Störung aber erst im Zeitschritt t_2 beginnen. Nach Ende der Aktion im Zeitschritt t_7 liegt der Gesamtverzug von Auftrag B bei vier Zeitschritten. Die entsprechenden terminbezogenen Kosten $TK_B(t_7)$ können berechnet werden. Im Zeitschritt t_0 hat der Agent jedoch bereits eine entsprechende Belohnung für den Verzug um zwei Zeitschritte erhalten. Seitdem hat sich die Fertigstellung störungsbedingt um zwei weitere Zeitschritte verschoben. Der für den Agenten im Zeitschritt t_7 ausstehende Kostenteil $\Delta TK_{B,a_1}$ berechnet sich demnach aus der Differenz zwischen den absoluten Verzugskosten $TK_B(t_7)$ im Zeitschritt t_7 und den Kosten $WK_B(t_o)$ im Zeitschritt t_o , als der Auftrag B noch in der Warteschlange der Produktionsressource war.

Das Beispiel verdeutlicht, dass im vorliegenden System die Agenten nach jeder Aktion lediglich den Teil der Belohnung bekommen, der seit letztmaligem Erhalt einer Belohnung hinzugekommen ist. Diese *relative Belohnungsvergabe* ist erforderlich, um durch die Berücksichtigung des Verzugs von Aufträgen in der Warteschlange keine Verzugskostenbestandteile mehrfach einfließen zu lassen.

Produktionskostenbestandteil der Belohnung

Neben den vorangegangenen Überlegungen sind zwei weitere Modifikationen bei der Formulierung der Belohnungsfunktion für die Produktionskosten zu beachten. Die Agenten müssen lernen, die Auswahl von nicht ausführbaren Aktionen, sogenannte „Bad Actions“ (vgl. Abschnitt 7.3.2), zu vermeiden. Wird daher eine „Bad Action“ ausgewählt, so erhält der Agent die Strafe r_{BA} . Analog zu den Energiekosten (vgl. Abschnitt 7.4.2) ist zudem durch Vergabe der negativen Produktionskosten die Umformulierung in ein Maximierungsproblem erforderlich.

Der Produktionskostenbestandteil der Belohnung setzt sich demnach aus der Gesamtkostenfunktion des abgeschlossenen Auftrags $GK_n(t)$ und den Verzugskosten der Aufträge in der Warteschlange $WK_n(t)$ zusammen, wovon aufgrund der relativen Belohnungsvergabe die bereits nach Ende der vorausgegangenen Aktion erhaltene Belohnung $WK_i(t - d_a)$ abgezogen wird. Insgesamt lässt sich der Produktionskostenbestandteil der Belohnung wie folgt berechnen:

$$R_{t,a}^p = \begin{cases} - \left(GK_{n_a}(t) + \sum_{i=0}^{n_{p,t}} WK_i(t) - WK_i(t - d_a) \right) & , a \notin \{A_{BA,t}\} \\ - r_{BA} & , a \in \{A_{BA,t}\} \end{cases} \quad (22)$$

- $R_{t,a}^p$ Produktionskostenbestandteil der Belohnung für Produktionsressource p im Zeitpunkt t für Aktion a
- $n_{p,t}$ Aufträge in der Warteschlange von Produktionsressource p zum Zeitpunkt t
- $GK_n(t)$ Gesamtkostenfunktion von Auftrag n (vgl. Formel (7))
- $WK_n(t)$ Terminbezogene Kosten des Auftrags n in der Warteschlange (vgl. Formel (21))
- d_a Tatsächliche Dauer der abgeschlossenen Aktion a
- r_{BA} Strafbelohnung einer „Bad Action“
- $\{A_{BA,t}\}$ Menge an Aktionen, die zum Zeitpunkt t nicht ausgeführt werden können

7.4.4 Resultierende Belohnungsvergabe

Aus dem Systemaufbau und der Belohnungsvergabe ergibt sich insgesamt der in Abbildung 44 schematisch dargestellte Ablauf. Das Durchlaufen der Steuerungslogik beginnt zu jedem Zeitschritt mit dem Handelsagenten, der lediglich alle 15 Minuten aktiv wird. Als Nächstes sind die Ressourcenagenten an der Reihe, im Falle einer abgeschlossenen Aktion die entsprechende sofortige Belohnung einzuholen und eine neue Aktion auszuwählen. Der Speicheragent wählt als einziger Agent in jedem Zeitschritt eine neue Aktion aus. Falls das Ende der aktuellen Produktionsschicht angebrochen ist, werden im letzten Schritt die Gesamtenergiekosten berechnet und in Form einer globalen Belohnung an alle RL-Agenten verteilt.

Neben dem aufgezeigten Ablauf erhalten sowohl der Speicher- als auch der Marktagent im Falle der Auswahl einer im aktuellen Zustand unzulässigen Aktion („Bad Action“) eine Strafe. Diese wird analog zu den Produktionsagenten in Form eines pauschalen Kostenfaktors vergeben. Aus Gründen der Übersichtlichkeit ist diese einfache Vergabemechanik in Abbildung 44 nicht dargestellt.

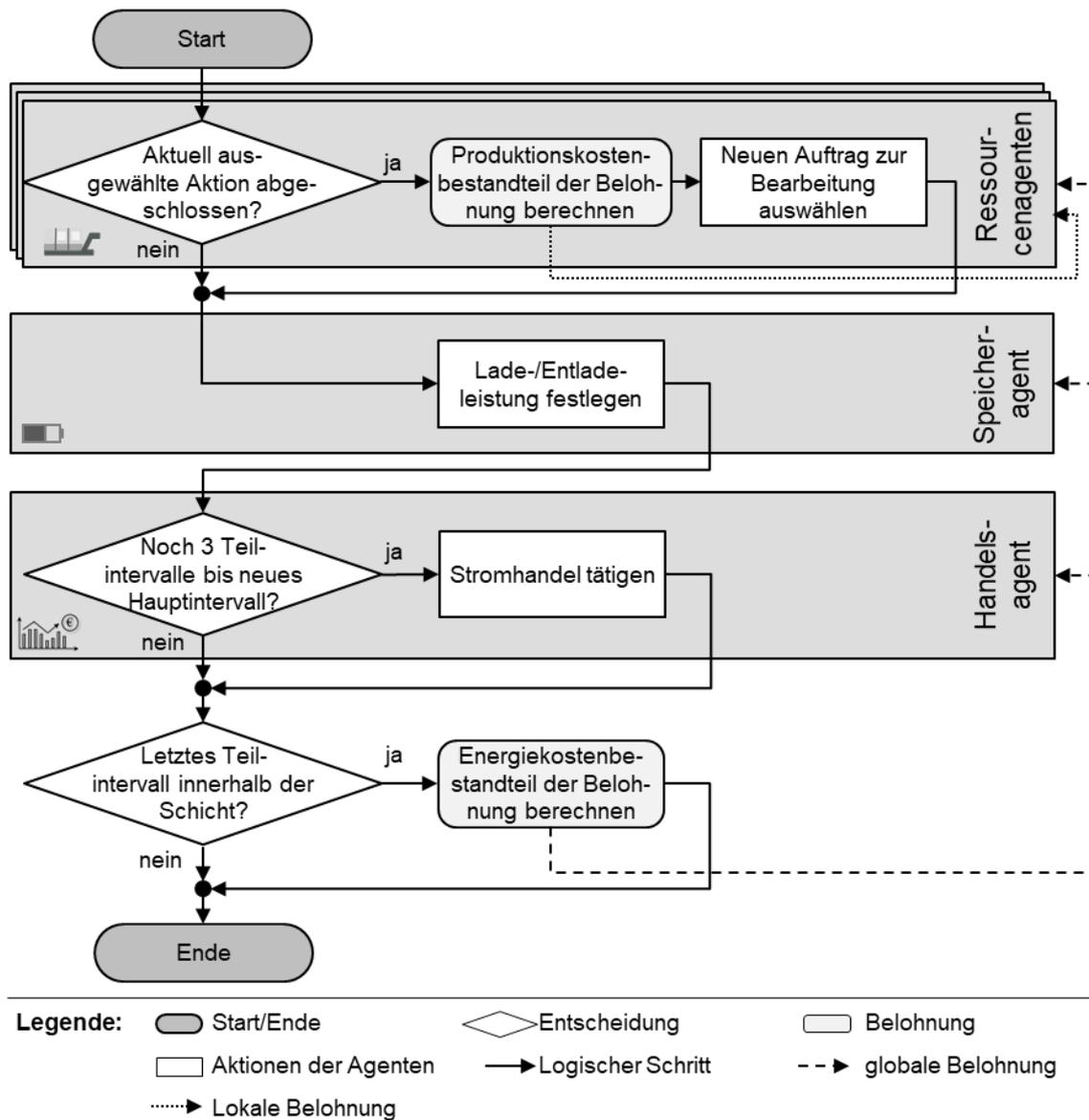


Abbildung 44: Ablauf der Belohnungsvergabe in jedem Zeitschritt (in Anlehnung an ROESCH ET AL. 2020)

7.5 Lernverfahren

Wie in Abschnitt 2.5.3 bereits erwähnt, existieren im Bereich von RL grundsätzlich zahlreiche Algorithmen zum Lernen der optimalen Strategie. Die Auswahl des idealen Verfahrens zur Lösung einer bestimmten Problemstellung erfolgt in der Regel durch Benchmarks. So existieren zahlreiche Untersuchungen zu einem mehr oder weniger breiten Anwendungsspektrum, wie etwa in LOON ET AL. (2019), KRISHNAN ET AL. (2019) oder WANG ET AL. (2019B). Im Kontext von Multi-Agentensystemen im Allgemeinen (OROOJLOOYJADID & HAJINEZHAD 2019, ZHANG ET

AL. 2019) und zur Lösung des JSSP im Besonderen (CUNHA ET AL. 2020) werden modellbasierte Verfahren bisher kaum oder gar nicht eingesetzt. Dies ist unter anderem auf die im Zuge von komplexen, dezentralen Systemen hohe erforderliche Rechenleistung zur Approximation des Modells zurückzuführen (GABEL 2009, FRANÇOIS-LAVET ET AL. 2018, PARK ET AL. 2019). In der vorliegenden Arbeit wird demnach mit PPO (vgl. Abschnitt 2.5.3) ebenfalls auf einen modellfreien Ansatz zurückgegriffen. Im Vergleich zu anderen verbreiteten Ansätzen konnte dieser Algorithmus die besten Ergebnisse im Bereich von diskreten Steuerungsaufgaben erzielen (LOON ET AL. 2019). Die hier vorliegende Problemstellung ist ebenfalls dieser Kategorie zuzuordnen. Darüber hinaus ist PPO aufgrund der schnellen Konvergenz und guten Performance im Allgemeinen ein sehr verbreitetes Lernverfahren.

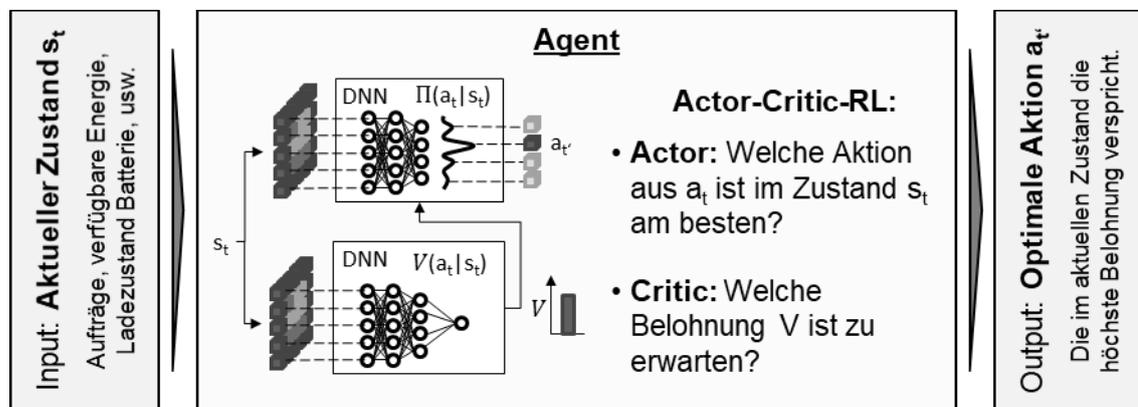


Abbildung 45: Funktionsschema von PPO und Actor-Critic-RL

PPO ist dem Bereich des Deep Reinforcement Learning zuzuordnen und wird in der vorliegenden Arbeit als Actor-Critic implementiert. Während der Actor die Aufgabe hat, die optimale Strategie $\Pi(a_t|s_t)$ in Form einer Wahrscheinlichkeitsverteilung zu approximieren und die Aktionen auszuwählen, nähert der Critic-Teil die Wertefunktion $V(a_t|s_t)$ der vom Actor ausgewählten Aktion an. Im Zuge des Trainings wird die Critic anhand der gesammelten Belohnung trainiert, während der Actor auf Basis der aus der Critic hervorgehenden Wertefunktion optimiert wird. Dadurch weiß der Agent nicht nur basierend auf dem Actor, welche Aktionen auszuwählen sind, sondern kann vielmehr auch die daraus resultierende Belohnung abschätzen. Actor und Critic können sowohl über zwei separate Netze zur Approximation verfügen als auch ein gemeinsames Netz teilen, in dessen letzter Schicht entsprechende Outputs für Actor und Critic ausgegeben werden. Das Konzept des Verfahrens ist in Abbildung 45 am Beispiel einer Actor-Critic mit zwei separaten Netzen dargestellt. Darüber hinaus wird bei der Anwendung von PPO

ein Satz an gesammelten Erfahrungen (engl. Batches) für das Training in sogenannte Mini-Batches zerlegt. Auf diese Weise wird der Gradient der Netze stückweise angepasst. Um dennoch einen stringenten Gradienten-Abstieg zu ermöglichen, wird zudem die Schrittweite jeder Iteration durch den Clipping-Faktor begrenzt. (GRONDMAN ET AL. 2012, SCHULMAN ET AL. 2017)

7.6 Trainingsumgebung

Das Approximationsmodell zur Auswahl der im jeweiligen Zustand besten Aktion – und somit die eigentliche Intelligenz des Systems – steckt in den neuronalen Netzen der RL-Agenten. Um die richtige Entscheidung treffen zu können, müssen diese zuvor mit einem ausreichenden Erfahrungsschatz trainiert werden.

Zu Beginn des Trainings sind die Agenten nicht in der Lage, die richtigen Aktionen auszuwählen. Das Verhalten wird jedoch mit fortschreitendem Training besser. Um negative Auswirkungen auf das reale Produktionssystem zu vermeiden, muss das Agentensystem zuvor auf Basis einer Simulation trainiert werden. Zu diesem Zweck ist der in Abbildung 46 dargestellte Ablauf eingeführt.

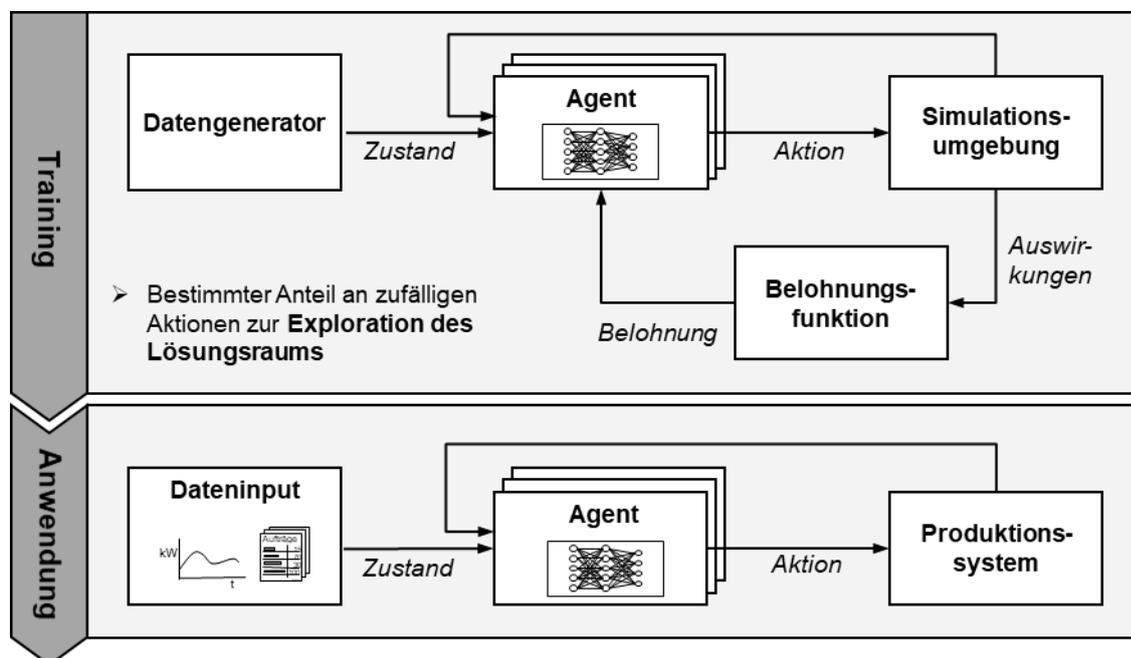


Abbildung 46: Training und Anwendung des Gesamtsystems

Für das Training des Systems wird eine große Anzahl an Trainingsszenarien benötigt. Die Datengrundlage der realen Produktion ist für diese Zwecke in der Regel

nicht ausreichend. Aus diesem Grund wird ein entsprechender Generator eingesetzt, der die vorhandene Informationsgrundlage mit Hilfe von statistischen Zufallsfunktionen erweitert und auf diese Weise nachfolgende Daten generiert:

- **Strompreise am kurzfristigen Strommarkt:** Hier können historische Preisverläufe aus beispielsweise dem letzten Jahr herangezogen werden, die zufällig auf Basis einer Normalverteilung in jedem Zeitschritt verändert werden. Im laufenden Training wird jeweils der reale Preisverlauf eines Tages zufällig gezogen und entsprechend verändert.
- **Erzeugungsleistung und Erzeugungsprognose der FEE:** Die Datengenerierung im Kontext der FEE läuft analog zu den Strompreisen am kurzfristigen Strommarkt ab. Hierbei ist jedoch auf die in der Realität auftretende Abweichung zwischen prognostizierter und tatsächlich erzeugter Energiemenge zu achten. Um das System möglichst realitätsnah abzubilden, ist eine Möglichkeit, die Standardabweichung zu ermitteln und diese bei der Datengenerierung zu berücksichtigen. Die Erzeugungsleistung kann ebenfalls auf Basis eines zufälligen Tages aus dem historischen Datensatz und unter Annahme einer Normalverteilung verändert werden. Zur Generierung der Prognose wird die daraus resultierende Zeitreihe mit der resultierenden statistischen Prognoseabweichung beaufschlagt.
- **Produktionsprogramm:** Für die Variation von Auftragskonstellationen können mit dem jeweiligen Fertigstellungstermin, der vorgegebenen Auslastung und der Verteilung der Auftragsstypen drei wesentliche Eingangsgrößen verändert werden. Falls die Aufträge nicht durch Auftragsstypen, sondern auf Basis von charakteristischen Parametern modelliert werden (vgl. Abschnitt 7.2.2), muss für diese Parameter ebenfalls eine ausreichende Varianz im Training sichergestellt sein. Die entsprechenden Werte werden zufällig variiert und können erneut mit Hilfe der Normalverteilung an die reale Produktionsauslastung angelehnt werden.

Weitere wichtige Bausteine beim Training sind die RL-Agenten; die Simulationsumgebung, in der alle technischen Teilsysteme modelliert sind; die Belohnungsfunktion. Erwähnenswert ist im Bezug zur Simulationsumgebung die Implementierung von stochastischen Ereignissen, wie etwa Maschinenausfälle (vgl. Abschnitt 7.2.3), deren Auftrittswahrscheinlichkeit und Dauer ebenfalls an das reale System angelehnt werden sollten.

Nach erfolgreichem Abschluss des Trainings kann das System auf reale Inputszenarien angewendet werden. Die Agenten interagieren mit dem realen Produktionssystem und sind nicht mehr auf die Belohnungsfunktion angewiesen, denn diese ist lediglich für die iterative Verbesserung im Training von Bedeutung.

7.7 Fazit

Im vorliegenden Kapitel wurde ein System zur integrierten Steuerung der Produktion und des Energiebezugs mittels Reinforcement Learning entwickelt. Dabei konnte auf den Vorarbeiten hinsichtlich der Zielgrößen in der Produktion sowie der Modellierung des Energiebezugs aus den vorausgegangenen Kapiteln 5 und 6 aufgebaut werden. Nachdem eingangs der Aufbau und der Ablauf des Systems beschrieben wurden (Abschnitt 7.1), erfolgte die Modellierung des Produktionssystems (Abschnitt 7.2). Die Produktionsressourcen und die abzuarbeitenden Aufträge sind dabei die zentralen Elemente. Darüber hinaus wurden zusätzlich stochastische Ereignisse abgebildet. Zur Modellierung des Zustands- und Aktionsraumes wurden die Agenten in die drei Klassen Ressourcen-, Speicher- und Marktagent unterteilt (Abschnitt 7.3). Auch die Belohnungsvergabe wurde aufgespaltet (Abschnitt 7.4). So erhalten lediglich die einzelnen Ressourcenagenten die Produktionskosten als direkte Belohnung, während die Energiekosten global an alle Agenten gleichermaßen verteilt werden. Als Lernverfahren wurde in Abschnitt 7.5 mit PPO ein Actor-Critic-Algorithmus eingeführt. Um ein ausreichendes Training der Agenten zu gewährleisten, wurde abschließend in Abschnitt 7.6) der Aufbau der Trainingsumgebung skizziert.

Das Ergebnis dieses Kapitels ist ein System, das in der Lage ist, durch ausreichendes Training die Produktion und den Energiebezug zur Reduzierung der Produktions- und Energiekosten zu steuern. Die Funktionsfähigkeit und der Nutzen der entwickelten Lösung werden im nachfolgenden Kapitel validiert. Insgesamt lernt das Agentensystem im Training, das im Modell abgebildete Produktionssystem zu steuern. Falls sich das reale Produktionssystem aufgrund von äußeren Einflüssen verändert und in Folge dessen vom Modell abweicht, müssen zur Anwendung des Agentensystems die Parameter des Modells entsprechend angepasst und ein erneutes Training durchgeführt werden.

8 Prototypische Umsetzung und Evaluierung

Das vorliegende Kapitel dient dem Zweck, die Funktionsweise des entwickelten Systems nachzuweisen. Zusätzlich soll der Nutzen des Ansatzes anhand von Vergleichen mit ausgewählten Benchmarks untersucht und die Wirtschaftlichkeit des Gesamtsystems bewertet werden. Um diese Ziele zu erreichen, wird zunächst im Abschnitt 8.1 das betrachtete Anwendungsbeispiel näher vorgestellt und in Abschnitt 8.2 die softwareseitige Umsetzung skizziert. Um alle Funktionalitäten des komplexen Gesamtsystems bewerten zu können, werden in Abschnitt 8.3 die betrachteten Evaluierungsszenarien erläutert und die daraus resultierenden Ergebnisse vorgestellt.

8.1 Beschreibung des Anwendungsbeispiels

Die Bewertung des entwickelten Systems zur energieorientierten und kostenbasierten Produktionssteuerung erfordert eine umfangreiche Simulationsumgebung (vgl. Abschnitt 7.1) mit Produktionssystem, den Energiebezugsoptionen und stochastischen Einflüssen. Zu diesem Zweck wird ein skalierbares System zur Evaluierung geschaffen, das sowohl eine unterschiedliche Anzahl an Produktionsressourcen als auch verschiedene Szenarien des Energiebezugs abdecken kann. Auf diese Weise können verschiedene Simulationsstudien mit unterschiedlicher Komplexität mit geringem Aufwand abgebildet werden. Dies ist zur Evaluierung des Gesamtsystems mit allen relevanten Funktionalitäten erforderlich (vgl. Abschnitt 8.3.1). Nachfolgend werden die dafür verwendeten Systemelemente und deren Parameter beschrieben.

8.1.1 Allgemeines

Als Anwendungsbeispiel wird der Produktionsstandort eines mittelständischen Anlagenbauers und Auftragsfertigers für die Präzisionsbearbeitung von metallischen Bauteilen und deren Oberflächen in Südbayern herangezogen. Die Fertigung gliedert sich in die Bereiche Werkzeugbau, Lohnfertigung und Montage. Im Bereich der Lohnfertigung werden auf Kundenwunsch Aufträge auf Basis der Druckfließbläppen-Technologie abgearbeitet, wobei die Taktzeit in der Regel unter einer Minute liegt und pro Takt mehrere Produkte gefertigt werden. Die Kunden fragen in der Regel Standardteile zur Bearbeitung an. Diese werden in Fertigungsaufträge

mit fixer Losgröße aufgeteilt, die wiederum im Zuge des einstufigen Bearbeitungsprozesses abgearbeitet werden. Die Teile werden anschließend direkt an die Kunden verschickt, wo sie Just-in-Time in der dortigen Montage verbaut werden. Die Geometrie und die Größe der zu bearbeitenden Bauteile unterscheiden sich stark, wodurch der Energiebedarf und die Bearbeitungszeit zwischen einzelnen Produkten deutlich variieren können. Der Produktionsstandort hat insgesamt einen Energieverbrauch von etwa 2.000 MWh im Jahr, wobei knapp die Hälfte auf die Lohnfertigung zurückzuführen ist. Eigenerzeugungsanlagen sind aktuell am Standort nicht vorhanden, befinden sich aber genauso wie ein Batteriespeicher in Planung. In Anlehnung an diese realen Gegebenheiten werden nachfolgend die für die Evaluierung des Gesamtsystems erforderlichen Elemente und Parameter abgeleitet.

8.1.2 Produktionssystem

In Anlehnung an die Lohnfertigung des Anwendungsszenarios besteht das einstufige Produktionssystem aus insgesamt fünf Produktionsressourcen, die diskrete Aufträge mit definierter Losgröße bearbeiten. Jede Produktionsressource fertigt dabei zwischen zwei und vier unterschiedliche Produktionsaufträge. Die Maschinenbelegung wurde dabei im Rahmen der vorgelagerten Produktionsplanung bereits festgelegt (vgl. Abschnitt 2.3).

Die Fertigungsaufträge werden im Zuge des in Abschnitt 7.2.1 vorgestellten Modellierungsansatzes durch den Auftragsstyp charakterisiert. Im Zuge dessen wird eine bestimmte Anzahl an gleichen Bauteilen mit identischen Bearbeitungsschritten zusammengefasst. Jeder Auftragsstyp verfügt über eine definierte Bearbeitungsdauer, die zwischen 27 und 42 Minuten variiert. Die innerhalb einer Schicht zu fertigenden Aufträge stehen bereits vorab fest. Mit einer Wahrscheinlichkeit von fünf Prozent können während der laufenden Schicht weitere Eilaufträge hinzukommen. Als weiterer stochastischer Einfluss liegt die Wahrscheinlichkeit einer Störung bei einem Prozent, bei einer durchschnittlichen Dauer von zwölf Minuten und einer Standardabweichung von drei Minuten. Jedem einzelnen Auftrag ist ein individueller Fertigstellungstermin zugeordnet. Wird dieser unterbeziehungsweise überschritten, fallen entsprechend Lager-/Verzugskosten an. Gleichzeitig schwankt auch der durchschnittliche Energiebedarf eines Auftrags zwischen zehn kW und 24 kW. Um einen bestimmten Auftrag bearbeiten zu können, muss die dafür eingesetzte Produktionsressource einen geeigneten Rüstzustand vorweisen. Das Rüsten nimmt je nach Ressource und Auftrag zwischen sechs und zwölf

Minuten in Anspruch und weist außerdem einen im Vergleich zur Bearbeitung reduzierten Energieverbrauch auf. Die einzelnen Parameter des betrachteten Produktionssystems sind in Tabelle 2 zusammengefasst.

Tabelle 2: Parameter des betrachteten Produktionssystems

Kategorie	Parameter	Wert
Produktionsressourcen und Aufträge	Anzahl der Produktionsressourcen	skalierbar
	Anzahl möglicher Auftragsstypen pro Produktionsressource	2 – 4
	Bearbeitungsdauer eines Auftrags	21 – 42 min
	Durchschnittlicher Leistungsbedarf pro Auftrag	20 – 48 kW
	Dauer Rüstvorgang je Auftragsstyp	6 – 12 min
Stochastische Ereignisse	Wahrscheinlichkeit für die Ankunft von Eilaufträgen (normalverteilt)	5 % ($\sigma = 2,5$ %)
	Wahrscheinlichkeit für den Ausfall einer Produktionsressource (pro Zeitschritt)	1 %
	Durchschnittliche Störungsdauer (normalverteilt)	12 min ($\sigma = 3$ min)
Kosten	Verzugskosten	48 – 88 €/h
	Lagerkosten	5,2 – 11 €/h
	Kosten pro Rüstvorgang	2 €

Die Faktoren zur Ermittlung der in Kapitel 6 hergeleiteten relevanten Produktionskosten wurden anhand eines Experteninterviews mit dem Produktionsverantwortlichen des Anwendungsunternehmens ermittelt. So wurden die Verzugskosten für einen Auftrag auf 48 € bis 88 € pro Stunde beziffert. Eine Nutzung verschiedener Versandoptionen ist nicht möglich. Aufgrund des knappen Lagerraums in unmittelbarer Nähe der Produktion werden im Falle einer zu frühen Fertigstellung um mehr als die Termintoleranz von 30 Minuten Lagerkosten in Höhe von etwa zwei bis 5,50 €/h fällig. Darüber hinaus verursacht der Verbrauch von Hilfsstoffen vergleichsweise geringe Kosten in Höhe von etwa zwei € pro Rüstvorgang.

8.1.3 Energiebezugsoptionen

Am realen Produktionsstandort des Anwendungsbeispiels sind keine Eigenerzeugungsanlagen vorhanden. In der Simulation werden, losgelöst vom realen Szenario, zwei Eigenerzeugungsanlagen, basierend auf PV und BHKW, ein stationärer Batteriespeicher sowie der flexible Stromhandel abgebildet.

Als Datenbasis für die modellierte PV-Anlage liegen die realen historischen Erzeugungsdaten einer Aufdachanlage vor, die sich in unmittelbarer Nähe zum Produktionsstandort des Beispielunternehmens befindet. Die erzeugte Energiemenge des historischen Datensatzes wird dabei so skaliert, dass durchschnittlich 25 Prozent des täglichen Energiebedarfs mit PV gedeckt werden können. Außerdem ist eine Erzeugungsprognose integriert, die die zukünftige Erzeugungsleistung mit einer Auflösung von 15 Minuten und einem normalverteilten Fehler bei einer Standardabweichung von zwei Prozent vorhersagt. Daneben wird eine, bereits vorab gekaufte, über die Schicht konstante Energiemenge angenommen. Da deren Kosten außerhalb des Betrachtungshorizonts der Produktionssteuerung liegen, werden diese nicht weiter betrachtet. Die insgesamt resultierende Energieverfügbarkeit ist in Abbildung 47 beispielhaft für eine Schicht dargestellt.

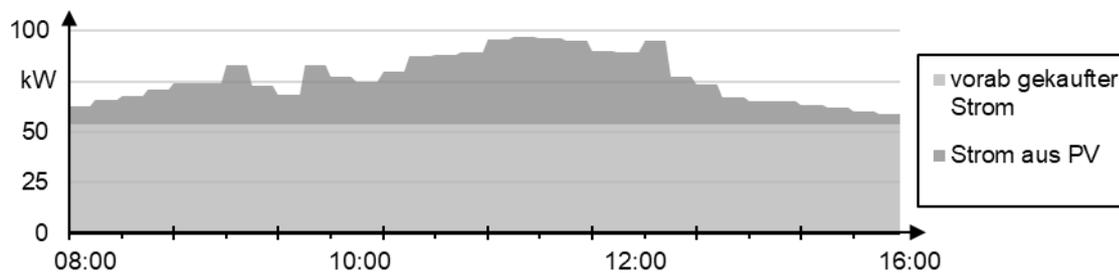


Abbildung 47: Beispielhafte Energieverfügbarkeit für eine Schicht

Neben der fest verfügbaren Energiemenge kann der Energiebedarf außerdem durch das flexibel steuerbare, standorteigene BHKW und den Batteriespeicher angepasst werden. Die Erzeugungsleistung des BHKW kann entsprechend dem ad-hoc Strombezug (vgl. Abschnitt 5.3.2) kurzfristig angepasst werden und verfügt über eine Nennleistung, um theoretisch 20 Prozent des Strombedarfs abzudecken. Die Stromgestehungskosten unter Vollast werden dabei auf 12 Ct/kWh festgelegt. Das BHKW kann lediglich innerhalb des Bereiches von 60 bis 100 Prozent der Nennleistung betrieben werden.

Als stationärer Batteriespeicher wird von einer Lithium-Ionen-Batterie ausgegangen, die in der Lage ist, den theoretischen, maximalen Energiebedarf aller Produktionsressourcen für 45 Minuten zu decken. Dabei wird eine C-Rate von 1 modelliert. Die Gesamtinvestitionen belaufen sich näherungsweise auf 100 € pro kWh Speicherkapazität, wobei diese im Zuge des Degradationsmodells (vgl. Abschnitt 5.2.3) auf jeden Zyklus umgelegt werden. Beispielhaft wird dabei eine Lithium-Ionen-Batterie auf Basis von Mangan-Dioxid verwendet. Die zugehörigen Modellparameter werden aus XU ET AL. (2018) entnommen.

Der letzte Baustein des Energiebezugs ist der kurzfristige Stromhandel. Dabei werden die historischen Preisverläufe am Intraday-Markt für den Zeitraum von Oktober 2018 bis September 2019 herangezogen. Um die Ungenauigkeit einer Strompreisprognose abzubilden, wurden die Preiskurven für den Zustandsraum der Agenten zufällig normalverteilt mit einer Standardabweichung von drei Prozent verändert. Näherungsweise kann Strom von bis zu 20 Prozent des theoretisch maximalen Strombedarfs des Produktionssystems verkauft/gekauft werden. Dabei wird, wie in Abschnitt 5.3.2 dargelegt, lediglich das jeweils nächste 15-Minuten-Intervall für den kurzfristigen Handel herangezogen. Neben den spezifischen Strompreisen werden beim Kauf zusätzliche Gebühren für Steuern und Umlagen aufgeschlagen, die sich auf 7,75 ct/kWh beziffern (BDEW 2020).

Weicht der Energieverbrauch innerhalb eines 15-Minuten-Intervalls von der verfügbaren Energiemenge ab, so sind für diese Lastgangabweichung Strafkosten zu entrichten (vgl. 5.2.4). In den weiteren Szenarien wird ein Toleranzband von \pm fünf Prozent angenommen. Liegt die Abweichung innerhalb dieser Grenzen, so werden keine zusätzlichen Kosten fällig. Andernfalls müssen Strafkosten entrichtet werden (vgl. 2.2.4). In der Realität kann die Höhe der dadurch anfallenden Kosten daher erst im Nachhinein genau beziffert werden und orientiert sich in der Regel an den Ausgleichsenergiepreisen¹ (ZAPF 2017). Aufgrund der starken Preisschwankungen ist diese Ausgleichsenergie mit großem Preisrisiko verbunden. So lag etwa der Mittelwert der Preisspitzen der Ausgleichsenergie im Jahr 2017 bei 440 €/MWh (EDEL & SCHEMM 2018). In der Regel sind die Preise jedoch deutlich niedriger. Derartige Preisspitzen entstehen lediglich in einem Prozent der Zeitintervalle. Um dennoch einen hohen Anreiz zur Vermeidung von Abweichungen zu schaffen, werden im vorliegenden Anwendungsbeispiel die Strafkosten auf einen

¹ Im deutschen Netzverbund als *regelzonenübergreifende einheitliche Bilanzausgleichspreise* (reBAP) bezeichnet.

Prototypische Umsetzung und Evaluierung

Wert von 200 €/MWh festgelegt, die näherungsweise sowohl für eine Unter- als auch eine Überschreitung der Abnahmemenge anfallen.

Die Dimension der einzelnen Komponenten, wie die Kapazität des Speichers oder der Eigenerzeugungsanlagen, ist skalierbar und ermöglicht eine einfache Anpassung an unterschiedliche Szenarien mit einer beliebigen Zahl an Agenten. Die Skalierung für die im weiteren Verlauf betrachteten Evaluierungsszenarien ist daher in Abschnitt 8.3 aufgeführt. Demgegenüber sind die Parameter zur Beschreibung der Energiekosten in allen untersuchten Szenarien unverändert und in Tabelle 3 zusammengefasst.

Tabelle 3: Kostenparameter des Strombezugs

Kategorie	Beschreibung	Wert
Photovoltaik	Preis für Strom aus PV-Anlage	0 ct/kWh
	Abweichung PV-Prognose (normalverteilt)	$\sigma = 2 \%$
BHKW	Preis für Strom des BHKWs bei Voll-Last	10 ct/kWh
	Wirkungsgradabhängiger Korrekturfaktor φ	0,1
Batteriespeicher	Investitionskosten	200 €/kWh
	Gesamtwirkungsgrad der Batterie	0,92
Intraday-Handel	Kostensatz für Steuern und Umlagen	7,75 ct/kWh
	Anteilige Gebühr pro Order	0.01 ct/kWh
	Gebührenpauschale pro Order	0.08 ct
	Abweichung Preisprognose (normalverteilt)	$\sigma = 3 \%$
Lastgang-Abweichung	Zusätzliche Kosten für Lastgangabweichungen	20 ct/kWh
	Toleranzband für Lastgangabweichung	$\pm 5 \%$

8.2 Technische Umsetzung

8.2.1 Softwarestruktur

Zur Implementierung des Softwareprototypen wird als zentraler Baustein für das Agentensystem die Open-Source-Bibliothek *RLlib* verwendet. Eine Simulation auf Basis von *Python* und *C* mit Modellen aller relevanten Komponenten bildet die Umgebung des RL-Systems, wo auch die Belohnung berechnet wird. In jedem Simulationszeitschritt wird der aktuelle Zustand des Systems an *RLlib* übergeben, das wiederum die auszuführende Aktion zurückgibt. Mittels *RLlib* werden die einzelnen Agenten angelegt und gesteuert sowie deren Training durchgeführt. *RLlib* greift wiederum auf die Bibliotheken *Ray* und *TensorFlow* zurück, über die das eigentliche Training der neuronalen Netze durchgeführt wird. Den Dateninput für die Simulation liefert der ebenfalls in *Python* und *C* implementierte Datengenerator, der, wie in Abschnitt 7.6 beschrieben, geeignete Trainingsdaten aus der realen Datenbasis extrahiert. Die Visualisierung der Simulation und des Trainingsverlaufs erfolgt im *TensorBoard*. Ein Überblick über die verwendeten Softwarebausteine und deren Zusammenspiel ist in Abbildung 48 gegeben.

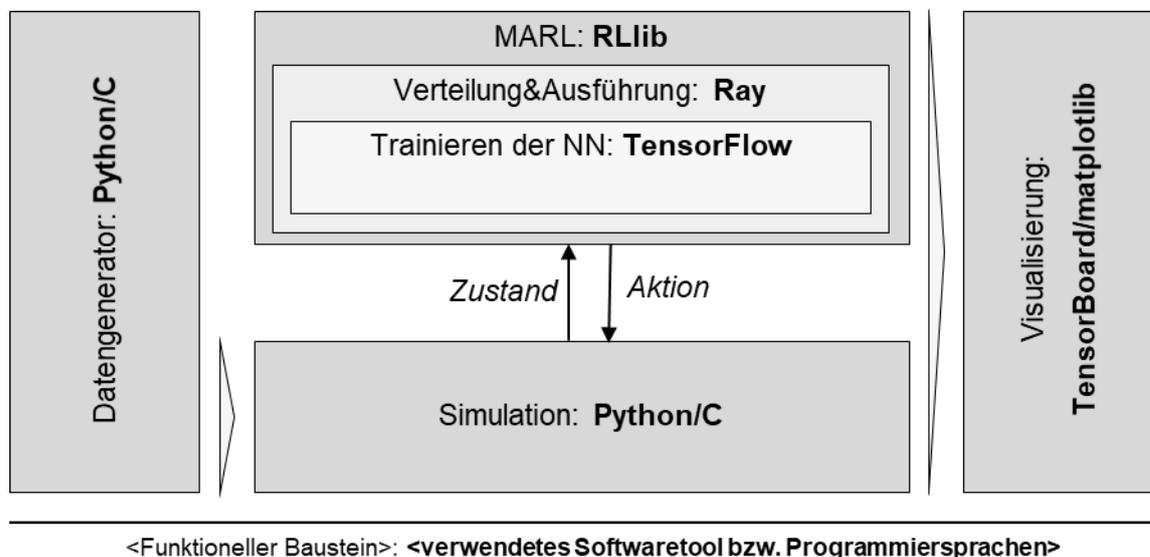


Abbildung 48: Softwarestruktur der technischen Umsetzung

In Kombination mit *Ray* bietet *RLlib* die Möglichkeit, das Training auf verschiedene dezentrale Entitäten, sogenannte *Worker*, zu verteilen (LIANG ET AL. 2017). Auf diese Weise laufen parallel mehrere Simulationen. Das Sammeln der erforderlichen Batches kann deutlich beschleunigt werden. Das Training der Netze findet weiterhin zentral statt. Handelt es sich bei dem angewendeten Lernverfahren um

einen synchronen Ansatz, werden die vollständigen Batches aller Worker gesammelt zum Training übergeben. Im Fall eines asynchronen Ansatzes werden die Netze iterativ trainiert, sobald ein Worker ein fertiges Trainingsbatch übergibt.

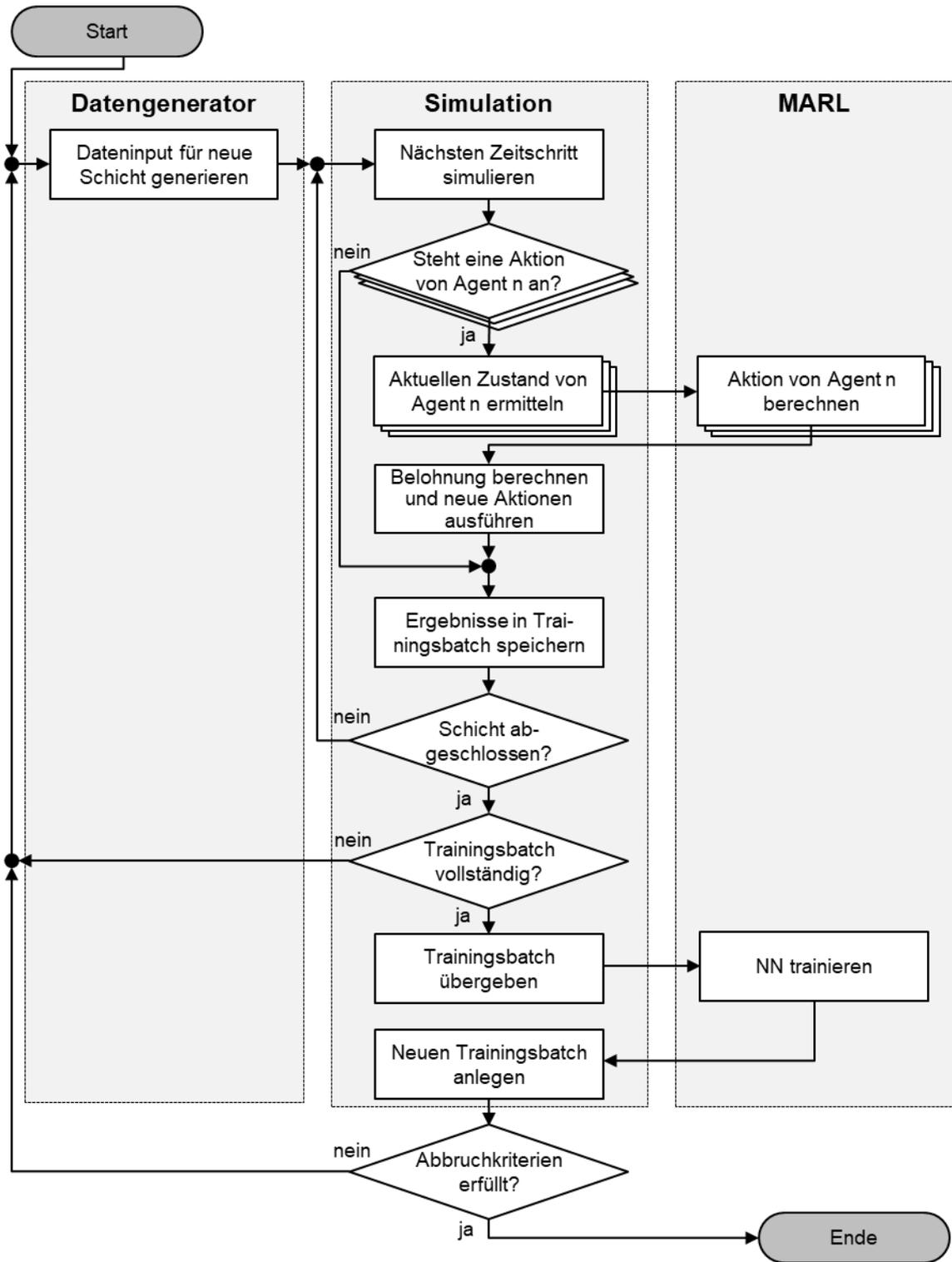
8.2.2 Simulationsablauf und Training

Für das Trainieren der neuronalen Netze müssen zahlreiche Produktionsschichten durchlaufen werden. Der Ablauf dieses Prozesses und die Funktionen der einzelnen Softwarebausteine sind in Abbildung 49 dargestellt.

Der Trainingsdatengenerator wird lediglich zu Beginn einer neuen Trainingschicht aufgerufen, während der größte Teil der Schritte innerhalb der Simulation abläuft. In jedem Zeitschritt werden zunächst die einzelnen Zustände der Agenten berechnet, und darauf aufbauend die Aktionen ausgeführt. Die resultierende Belohnung wird abgeleitet. Wenn ausreichend Trainingserfahrungen gesammelt wurden, erfolgt die Übergabe der Daten in Form des Trainingsbatches an RLlib. Die Größe des Trainingsbatches kann dabei beliebig eingestellt werden und entspricht einem Hyperparameter, der die Trainingsqualität beeinflusst. Das Training läuft so lange, bis die Simulation manuell beendet wird oder ein Abbruchkriterium, beispielsweise in Form einer bestimmten Anzahl an durchgeführten Episoden, erreicht wurde.

Im Falle einer Parallelisierung des Trainings sammeln mehrere Simulationen, die auf sogenannte Worker verteilt werden, die Erfahrungen für den Lernalgorithmus gleichzeitig. Auf diese Weise kann der Trainingsprozess deutlich beschleunigt werden.

Der Trainingsverlauf und -fortschritt werden parallel zur Laufzeit mittels TensorBoard visualisiert. Die Daten werden dabei nach jedem Training aktualisiert. Auf diese Weise kann der Trainingsfortschritt beispielsweise anhand der durchschnittlich pro Episode erzielten Gesamtkosten verfolgt werden. Weitere Visualisierungsmöglichkeiten sind vorhanden, etwa zur Darstellung der durchgeführten Aktionen eines Agenten während einer ausgewählten Episode. Während des Trainings werden darüber hinaus nach einer bestimmten Anzahl an Trainingsiterationen die Gewichte der neuronalen Netze abgespeichert. Auf diese Weise kann der Trainingsfortschritt nachvollzogen werden. Zudem dienen die abgespeicherten Gewichte als Absicherung für den Fall eines Programmabsturzes.



Legende: Start/Ende Entscheidung Operation

Abbildung 49: Trainingsablauf bei der technischen Umsetzung

8.3 Simulationstechnische Evaluierung

Die simulationstechnische Evaluierung erfolgt anhand von *drei Evaluierungsszenarien*, die in Abbildung 50 skizziert werden. Nachfolgend wird zunächst auf die einzelnen *Evaluierungsszenarien* eingegangen, bevor im Anschluss die *Ergebnisse* vorgestellt werden.

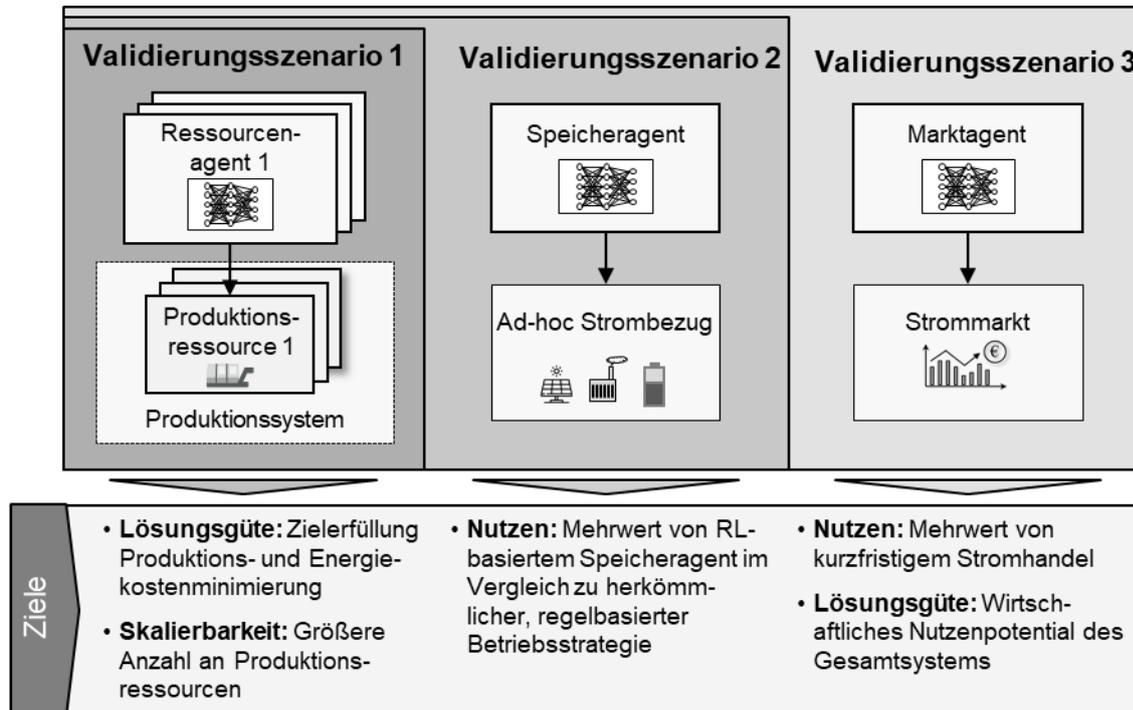


Abbildung 50: Stufenweiser Aufbau der drei Evaluierungsszenarien

8.3.1 Evaluierungsszenarien

Das entwickelte Gesamtsystem setzt sich aus mehreren funktionellen Elementen zusammen. Aus technischer Sicht ist eine wichtige Frage, wie sich die Integration der einzelnen Systembausteine auf die Zielerfüllung sowohl im Hinblick auf die lokale als auch die globale Belohnung auswirkt. Zudem gilt es, den Nutzen der einzelnen Erweiterungen zu untersuchen. Aufgrund der hohen Systemkomplexität können im Zuge einer alleinigen Evaluierung des Gesamtsystems die Funktionsweise der einzelnen Elemente und deren Beitrag zur Lösungsqualität kaum nachvollziehbar überprüft werden. Aus diesem Grund wird nachfolgend die Funktionalität des Gesamtsystems stufenweise untersucht, indem drei geeignete Evaluierungsszenarien herangezogen werden. Zur Einordnung der Ergebnisse werden jeweils geeignete Benchmarks, aus dem Bereich der reaktiven und der prädiktiv-

reaktiven Produktionssteuerungsansätze, eingesetzt. Um die Aussagekraft der Ergebnisse auf andere Evaluierungsszenarien übertragen zu können, basieren alle untersuchten Anwendungsfälle auf der identischen Datengrundlage (siehe Anhang 10.3 und 10.4).

Tabelle 4: Konfiguration der Evaluierungsszenarien

Bezeichnung	Einheit	1a	1b	2	3
Anzahl der Produktionsressourcen	Anzahl	5	15	5	5
Summe des Energiebedarfs*	kWh	ca. 900 kWh	ca. 2700 kWh	ca. 900 kWh	
Vorab gekaufte Strommenge*	kWh	75 % (ca. 675 kWh)	75 % (ca. 2025 kWh)	60 % (ca. 535 kWh)	
Strommenge aus Photovoltaik*	kWh	25 % (ca. 225 kWh)	25 % (ca. 675 kWh)	25 % (ca. 225 kWh)	
Kapazität Batteriespeicher	kWh	0		24 kWh	
Aktionsraum Batteriespeicher	kW	-		-24 bis 24 kW	
Nennleistung BHKW	kW	0		16 kW	
Aktionsraum Marktagent	kWh		-		-60 bis 60 kW

* Angaben pro Schicht, Werte variieren in Abhängigkeit der zufällig gewählten Auslastung

Die Szenarien gliedern sich anhand der wesentlichen funktionellen Elemente Ressourcenagenten, Speicheragent und Marktagent. Sie bauen, wie in Abbildung 50 dargestellt, stufenweise aufeinander auf. Während in Evaluierungsszenario 1 lediglich das Produktionssystem gemeinsam mit den Ressourcenagenten im Fokus steht, werden im Evaluierungsszenario 2 zusätzlich die Eigenerzeugungsanlagen und der Batteriespeicher abgebildet. Im Evaluierungsszenario 3 wird abschließend die Funktion des Gesamtsystems beleuchtet, inklusive des Intraday-Handels per Marktagent. Die Konfiguration der drei Szenarien ist in Tabelle 4 zusammengefasst. Darüber hinaus gibt Abbildung 51 einen Überblick über die für jedes Szenario verwendeten Benchmarks. Im Folgenden werden die einzelnen Szenarien und die jeweils verwendeten Benchmarks näher betrachtet.

	Systemelemente		
	Produktion	Ad-hoc Strombezug	Kurzfristiger Stromhandel
Evaluierungsszenario 1 MARL1: RL-basierte Produktionssteuerung 1a) 5 Ressourcenagenten  1b) 15 Ressourcenagenten	MARL		
EDF: reaktive Produktionssteuerung	EDF		
SA1: reaktiv-prädiktive Produktionssteuerung mittels Simulated Annealing	SA		
Evaluierungsszenario 2: MARL2: + RL-basierter ad-hoc Strombezug 	MARL		
RBS: regelbasierte Betriebsstrategie des Ad-hoc Strombezugs, MARL-basierter Produktionssteuerung	MARL	RBS	
SA2: Steuerung von Produktion und ad-hoc Strombezug mittels Simulated Annealing	SA		
Evaluierungsszenario 3: MARL3: + RL-basierter Stromhandel 	MARL		
MARL2: Vergleich mit MARL-basierter Steuerung ohne aktiven Stromhandel	MARL		
VRS: Vollständig reaktive Steuerung von Produktion und ad-hoc Strombezug	EDF	RBS	
SA3: Steuerung von Produktion, ad-hoc Strombezug und Stromhandel mittels Simulated Annealing	SA		

Legende: MARL: Multi-Agent Reinforcement Learning
SA: Simulated Annealing
VRS: Vollständig reaktive Steuerung

EDF: Earliest Deadline First
RBS: Regelbasierte Betriebsstrategie

Abbildung 51: In den Evaluierungsszenarien verwendete Benchmarks

Evaluierungsszenario 1:

Das Evaluierungsszenario 1 hat zum Ziel, die Funktionsweise der Produktionssteuerung (*MARL1*) zu beurteilen. Es teilt sich in zwei Teilversuche auf. Gegenstand der Untersuchung in *Evaluierungsszenario 1a* (ES1a) ist die Berücksichtigung sowohl der Energie- als auch Produktionskosten und der daraus resultierenden Lösungsqualität. Auf Basis einer bestimmten, vorab gekauften Energiemenge und der Verfügbarkeit von Strom aus einer Photovoltaikanlage fallen Strafkosten an, falls das Produktionssystem seinen Verbrauch nicht an das schwankende Stromangebot anpasst.

Ein weiterer Blick richtet sich auf die Möglichkeit zur Skalierung des Produktionssystems für mehr als die im vorliegenden Anwendungsbeispiel vorhandenen fünf Ressourcen. Aufgrund dessen wird im *Evaluierungsszenario 1b* zusätzlich ein Produktionssystem herangezogen, bestehend aus 15 Ressourcen. Die hinzugekom-

menen Fertigungsmaschinen werden in Anlehnung an die real vorhandenen Stationen parametrisiert. Die detaillierten Auftragsparameter und die Konfiguration des RL-Algorithmus mit zugehörigen Hyperparametern können dem Anhang entnommen werden.

Als Benchmark für die MARL basierte Produktionssteuerung wird mit der Heuristik *Earliest Deadline First (EDF)* und einer Optimierung mittels *Simulated Annealing (SAI)* jeweils ein Repräsentant sowohl der prädiktiv-reaktiven als auch der rein reaktiven Steuerungsansätze (vgl. 2.4.2) herangezogen. Im Zuge von EDF wird stets der Auftrag mit der kürzesten Zeitspanne bis zum erforderlichen Fertigstellungstermin zur Bearbeitung ausgewählt. Die Metaheuristik SA ist ein weitverbreiteter Ansatz im Kontext der Produktionssteuerung (vgl. Abschnitt 2.4.4). Im vorliegenden Anwendungsfall wird jeweils zu Schichtbeginn und nach Auftreten eines stochastischen Ereignisses eine Neuplanung angestoßen. Der mögliche Zustandsraum der einzelnen Produktionsressourcen wurde direkt an die für das MARL-System geltenden Randbedingungen angelehnt. Die Zustandsänderungen erfolgen zufällig durch eine Veränderung der Reihenfolge und des Startzeitpunktes von Aufträgen. Verändert sich aufgrund dessen die Anzahl der Rüstzeiten in einer Schicht, so werden für diesen Zeitraum entsprechende Stand-by-Zeiten entnommen beziehungsweise eingefügt. Die Implementierung erfolgt in Anlehnung an VAN LAARHOVEN & AARTS (1992) in der Programmiersprache C. Details zur Parametrierung können Anhang 10.6 entnommen werden.

Evaluierungsszenario 2:

Im Fokus des *Evaluierungsszenarios 2* steht die integrierte Steuerung ad-hoc Strombezugs zusammen mit der Produktion (*MARL2*). Zu diesem Zweck wird das bestehende Produktionssystem aus Szenario 1a um einen Lithium-Ionen-Speicher modelliert, mit einer Kapazität von 24 kWh und Eigenerzeugung auf Basis eines BHKWs, mit einer Leistung von 16 kW. Es wird angenommen, dass der Batteriespeicher dank einer Klimatisierung auf einer konstanten Temperatur von 20° C gehalten wird. Die Speichermodellierung erfolgt auf Basis der Modellbeschreibung in Abschnitt 5.2.3. Die verwendeten Parameter sind in Anhang 10.1 dargelegt.

Um den alleinigen Nutzen einer RL basierten Batteriesteuerung beurteilen zu können, wird darüber hinaus ein Vergleich mit einer *regelbasierten Betriebsstrategie (RBS)* angestellt. Dabei werden in der Regel statische Schwellwerte für die vom Netz bezogene Leistung definiert, ab deren Über- oder Unterschreitung die Batte-

rie geladen/entladen wird. RBS stellen den aktuellen Stand der Technik beim Einsatz von Batteriespeichern in der Praxis dar (WENIGER ET AL. 2014). Darüber hinaus werden stromgeführte BHKWs in der Regel mit möglichst vielen Volllaststunden betrieben, ohne dabei ins Netz zurück zu speisen (LECHNER ET AL. 2018). Im verwendeten RBS-Ansatz wird daher primär versucht, die Batterie so zu steuern, dass die Lastgangabweichung ausgeglichen und das BHKW mit maximaler Leistung betrieben werden kann. Ist eine Lastabweichung zu groß, um von der Batterie und der Volllast des BHKWs ausgeglichen zu werden, erfolgt eine entsprechende Anpassung der BHKW-Leistung. Die Logik dieser RBS ist im Anhang 10.5 als Pseudocode beschrieben. Die Produktionsressourcen werden in diesem Szenario genauso wie in VS1 per RL gesteuert.

Zur Einordnung der Wirksamkeit des Batterieagenten wird zusätzlich der bereits in *Evaluierungsszenario 1* eingesetzte SA-Algorithmus um einen Batteriespeicher zu SA2 erweitert und auf diese Weise das Gesamtsystem prädiktiv-reaktiv gesteuert. Dadurch kann die Lösungsqualität des Gesamtsystems eingeordnet werden.

Evaluierungsszenario 3:

Ziel des *Evaluierungsszenario 3* ist im Allgemeinen die Bewertung der Wirksamkeit des Gesamtsystems im Sinne der Zielgrößenerreichung und im Besonderen die Evaluierung des Nutzens des Marktagenten. Zu diesem Zweck wird das System aus *Evaluierungsszenario 2* um den kurzfristigen Stromhandel mit Hilfe eines Marktagenten erweitert (*MARL3*). Wie in Abschnitt 5.2.5 erläutert, ist der Marktagent in der Lage, neun Minuten vor Beginn eines neuen Hauptintervalls kurzfristig Strom zu handeln. Die maximal handelbare Strommenge ist bei 15 kWh und damit pro Hauptintervall auf eine Leistung von 60 kW begrenzt. Zur grundlegenden Evaluierung des Mehrwerts, der durch den Marktagenten generiert werden kann, wird der *MARL3*-Ansatz zudem mit den Ergebnissen von *MARL2* aus *Evaluierungsszenario 2* verglichen, das ohne kurzfristigen Stromhandel agiert.

Zur Bewertung des wirtschaftlichen Nutzenpotenzials des Gesamtsystems wird mit einer vollständig reaktiven Steuerung (*VRS*) ein Benchmark herangezogen, der in dieser Form schon heute in der Praxis eingesetzt werden kann und auf regelbasierten Logiken beruht. Dabei erfolgt eine Steuerung der Produktionsressourcen mittels EDF. Der ad-hoc Strombezug erfolgt auf Basis der in *Evaluierungsszenario 2* angewendeten regelbasierten Betriebsstrategie (RBS).

Um darüber hinaus eine Bewertung des Gesamtsystems zu ermöglichen, wird der bereits in den vorherigen *Evaluierungsszenarien* entwickelte *SA-Algorithmus* um

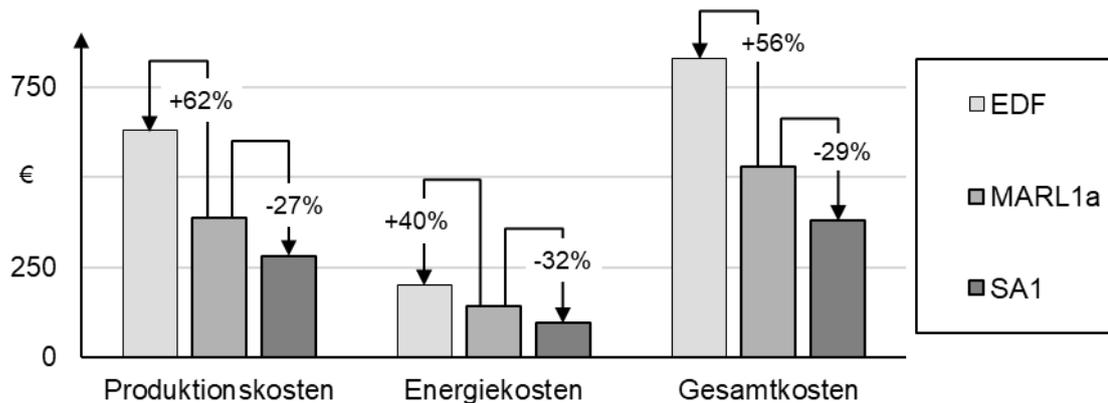
den kurzfristigen Stromhandel am Intraday-Markt zu SA3 erweitert. Auf diese Weise ist eine Abschätzung der Lösungsqualität des Gesamtsystems möglich.

8.3.2 Ergebnisse

Im nachfolgenden Abschnitt sind die in den Evaluierungsszenarien erzielten Ergebnisse zusammengefasst. Nähere Informationen zu den verwendeten Evaluierungsepisoden können den Anhängen 10.2 bis 10.6 entnommen werden. Für alle Berechnungen wurde die identische virtuelle Maschine mit 16x 2.4 GHz CPU und 32 GB RAM verwendet. Alle MARL-Ansätze wurden 3000 Episoden lang trainiert.

Evaluierungsszenario 1a:

Die Ergebnisse des Evaluierungsszenarios 1a sind in Abbildung 52 dargestellt. Dabei wurde der entwickelte MARL-Ansatz für zehn Evaluierungsepisoden von jeweils einer Schicht und fünf Ressourcenagenten mit den in Abschnitt 8.3.1 beschriebenen Ansätzen EDF und SA1 verglichen.



Erforderliche Rechenzeit für eine Umplanung:

EDF: < 1 Sekunde

MARL1a: ~ 1 Sekunde

SA1: bis zu 1,5 Stunden

Abbildung 52: Durchschnittliche resultierende Kosten in Evaluierungsszenario 1a

Es zeigt sich, dass die durchschnittlichen Gesamtkosten im Vergleich zu EDF deutlich reduziert werden können. Gleichzeitig schneidet MARL1a insgesamt um 29 Prozent schlechter ab als die nahoptimale Lösung von SA1. In Anbetracht der Rechenzeit erzielen die drei Verfahren ebenfalls sehr unterschiedliche Ergebnisse. Während EDF und MARL innerhalb von weniger als einer Sekunde eine Entscheidung treffen können, benötigt die Berechnung eines neuen Produktionsplans mit-

tels SA1 etwa 1,5 Stunden, was die Anwendung von SA im Kontext der Produktionssteuerung in der Praxis deutlich limitiert. Für das vorausgehende Trainieren von MARL1 wurden etwa 48 Stunden benötigt.

Die beispielhaften Verläufe von *verfügbarer* und *verbrauchter Energiemenge* in einer Produktionsschicht sind zusammen mit dem Verhalten eines Agenten in Abbildung 53 dargestellt. Der Ressourcenagent nimmt einen Stand-by-Zustand an, obwohl sich offene Aufträge in der Warteschlange befinden, die es noch in der laufenden Schicht abzuarbeiten gilt. Dies trägt dazu bei, eine kostspielige Lastgangabweichung zu verhindern. Die Agenten lernen demnach, die Produktionsressourcen temporär abzuschalten, um eine Minimierung der Energiekosten zu ermöglichen. Die Agenten können durch ihre erlernte Strategie auch für kurzfristiges Lastmanagement eingesetzt werden. Diese Fähigkeit wurde in Abschnitt 3.4 als Handlungsbedarf identifiziert.

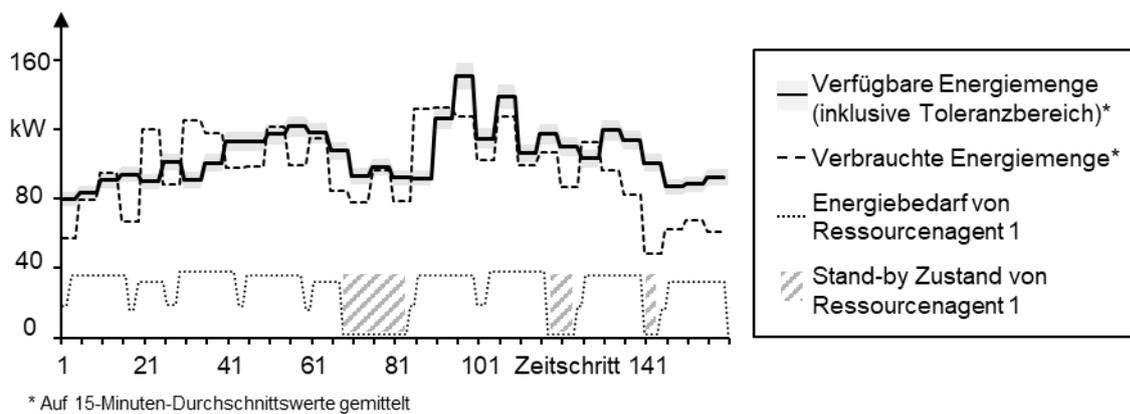
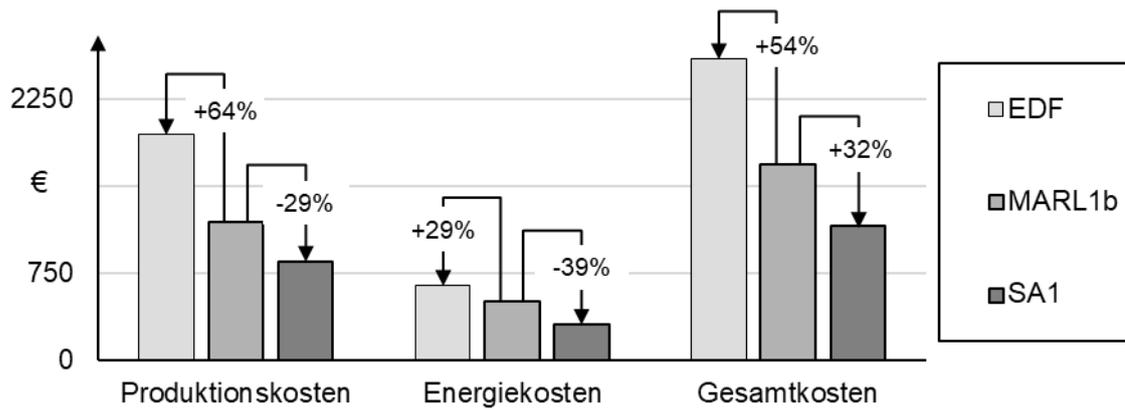


Abbildung 53: Verhalten eines Ressourcenagenten in Evaluierungsszenario 1a (in Anlehnung an Roesch et al. 2019)

Evaluierungsszenario 1b

Die prototypische Skalierung der entwickelten Produktionssteuerung auf 15 Produktionsressourcen führt zu den in Abbildung 54 dargestellten Ergebnissen. Erneut wurde *MARL1b* mit den beiden Benchmark-Verfahren für zehn identische Evaluierungsschichten angewendet. Das vorgelagerte Training des MARL-Systems nahm dabei etwa eine Woche in Anspruch.



Erforderliche Rechenzeit für eine Umplanung:

EDF: < 1 Sekunde

MARL1b: ~ 2 Sekunde

SA1: bis zu 5 Stunden

Abbildung 54: Durchschnittliche resultierende Kosten in Evaluierungsszenario 1b

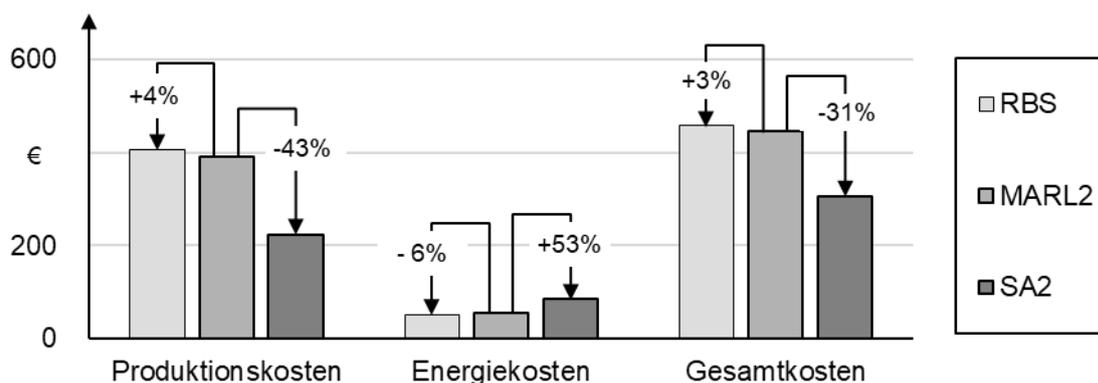
Trotz des deutlichen Anstiegs der Systemkomplexität schneidet *EDF* durchschnittlich im Bereich Energiekosten und im Hinblick auf die Produktionskosten um gut die Hälfte schlechter ab als der entwickelte *MARL*-Ansatz. Die nahoptimale Lösung von *SA1* ist durchschnittlich um 33 Prozent besser. Für eine Umplanung wird mit diesem Ansatz allerdings eine Rechenzeit von etwa fünf Stunden benötigt. Auffällig ist, dass sich der Abstand zwischen den resultierenden Produktionskosten in *MARL1b* und *SA1* mit 33 Prozent vergleichsweise nahe an den Ergebnissen des Evaluierungsszenarios 1a mit fünf Agenten (29 Prozent) einordnet. Im Hinblick auf die Energiekosten ist die Differenz mit 42 Prozent deutlich gestiegen. Dies kann als Hinweis dafür gewertet werden, dass es mit wachsender Ressourcenzahl für die Agenten auch zunehmend schwieriger wird, die gemeinschaftliche Minimierung der globalen Belohnung zu erreichen. Diese Beobachtung kann auf das bereits in Abschnitt 7.4.1 diskutierte „credit assignment problem“ zurückzuführen sein. Es ist demnach für einen Agenten mit zunehmender Agentenanzahl herausfordernd, die unmittelbaren Auswirkungen seiner Aktionen auf die globale Belohnung abzuschätzen. Daraus lässt sich schließen, dass neben der Rechenzeit möglicherweise auch die Minimierung der globalen Belohnung eine technische Grenze für die Skalierbarkeit des Ansatzes sein kann. Insgesamt ist das System mit 15 Ressourcenagenten dennoch in der Lage, die Kosten maßgeblich zu minimieren.

Evaluierungsszenario 2

Im Evaluierungsszenario 2 wird das entwickelte *MARL*-System zur integrierten Steuerung der Produktion und des Ad-hoc Strombezugs (*MARL2*) eingesetzt. Die

in Abbildung 55 dargestellten Ergebnisse zeigen die vergleichsweise hohe Lösungsgüte von *RBS*, bei dem eine regelbasierte Betriebsstrategie von Speicher und BHKW mit der RL basierten Steuerung der Ressourcenagenten (MARL1) aus VS1 gepaart wird. Während dieser Ansatz durchschnittlich um sechs Prozent niedrigere Energiekosten als MARL2 erreicht, fallen die Produktionskosten im Vergleich zu einer RL basierten Speichersteuerung von MARL2 um vier Prozent höher aus. Insgesamt lässt sich durch MARL2 eine Reduzierung der Gesamtkosten um etwa drei Prozent erzielen. Der Einsatz eines RL basierten Batterieagenten ermöglicht offensichtlich eine bessere Kooperation der Agenten. Die Ressourcenagenten sind auf diese Weise in der Lage, die Produktionskosten im Vergleich zu einer RBS des Speichers zusätzlich zu reduzieren. Diese Verbesserung geht allerdings zu Lasten der Energiekosten, die im Durchschnitt etwas höher ausfallen.

Demgegenüber können durch eine Steuerung des Gesamtsystems mittels SA2 die Gesamtkosten um etwa ein weiteres Drittel reduziert werden. Diese Verbesserung kommt durch die Verringerung der Produktionskosten um 43 Prozent zustande. Die Energiekosten liegen dagegen um über 50 Prozent höher als im vergleichbaren MARL2-Ansatz. SA2 scheint somit ein Kostenminimum zu finden, indem der Fokus auf die Reduzierung der Produktionskosten fällt und etwas höhere Energiekosten in Kauf genommen werden. Die Strategie von MARL2 resultiert dagegen in einem lokalen Optimum, das niedrigere Energiekosten aber gleichzeitig deutlich höhere Produktionskosten nach sich zieht.



Erforderliche Rechenzeit für eine Umplanung:

RBS: < 1 Sekunde

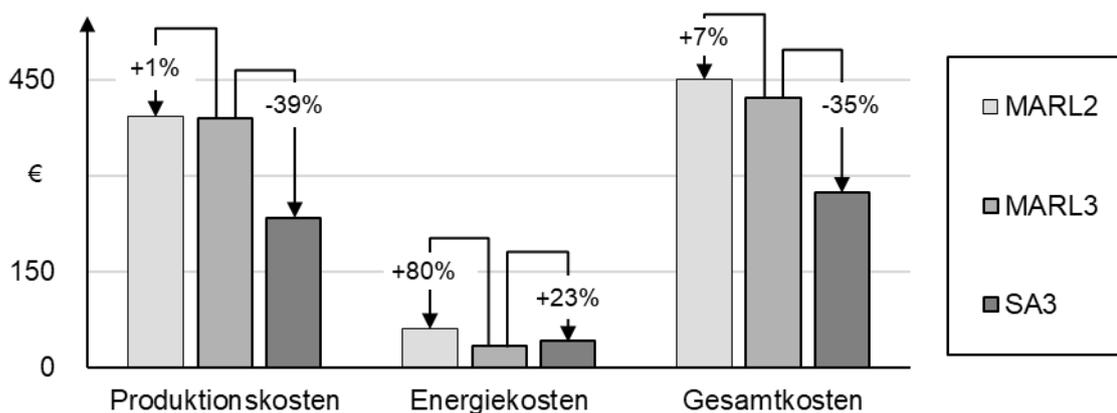
MARL2: ~ 1 Sekunde

SA2: bis zu 2 Stunden

Abbildung 55: Durchschnittliche resultierende Kosten in Evaluierungsszenario 2

Evaluierungsszenario 3

Im Zuge des abschließenden Evaluierungsszenarios 3 werden die Funktionsweise des Gesamtsystems und der zusätzliche Nutzen, der durch den kurzfristigen Stromhandel mittels Marktagent entsteht, anhand der Benchmark-Verfahren bewertet (vgl. Abbildung 56). Eine Einschätzung im Hinblick auf das zusätzliche Potenzial zur Kostenreduzierung durch den Einsatz des Marktagenten gibt insbesondere die Gegenüberstellung mit den Ergebnissen des Laufs mit *MARL2*. Der alleinige Unterschied dieser Systemkonfiguration ist, dass *MARL2* im Vergleich zu *MARL3* keinen Marktagenten verwendet und daher kein kurzfristiger Stromhandel möglich ist (vgl. Evaluierungsszenario 2). Die untersuchten zehn Evaluierungsszenarien zeigen, dass *MARL3* vergleichbare Produktionskosten erzielt (vgl. Abbildung 56). Demgegenüber stehen signifikant niedrigere Energiekosten. Ohne Marktagenten fallen die Kosten durchschnittlich 80 Prozent höher aus. Insgesamt bleibt ein Kostenvorteil von *MARL3* in Höhe von sieben Prozent der Gesamtkosten.



Erforderliche Rechenzeit für eine Umplanung:

RBS: < 1 Sekunde

MARL2: ~ 1 Sekunde

SA2: bis zu 2,5 Stunden

Abbildung 56: Durchschnittliche resultierende Kosten im Evaluierungsszenario 3

Die Gegenüberstellung mit der nahoptimalen Lösung von *SA3* vermittelt einen Eindruck über die Lösungsqualität des Gesamtsystems. Hier lässt sich festhalten, dass die durchschnittlichen Kosten von *MARL3* um 35 Prozent höher liegen. Ähnlich wie im Evaluierungsszenario 2 ist dieser Kostenvorteil alleinig auf die Produktionskosten zurückzuführen, während die Energiekosten mit *SA3* um durchschnittlich 23 Prozent ansteigen. Somit erreicht auch *MARL3* ein lokales Optimum, mit vergleichsweise hohen Produktionskosten und niedrigen Energiekosten. Dennoch wird ein klarer Vorteil dieser Lösung bei der Betrachtung der erforderli-

chen Rechenzeit deutlich. Während MARL3 nach einem knapp einwöchigen Training auf ein stochastisches Ereignis innerhalb von etwa einer Sekunde reagieren kann, benötigt SA3 für die Neuberechnung innerhalb einer laufenden Episode bis zu drei Stunden.

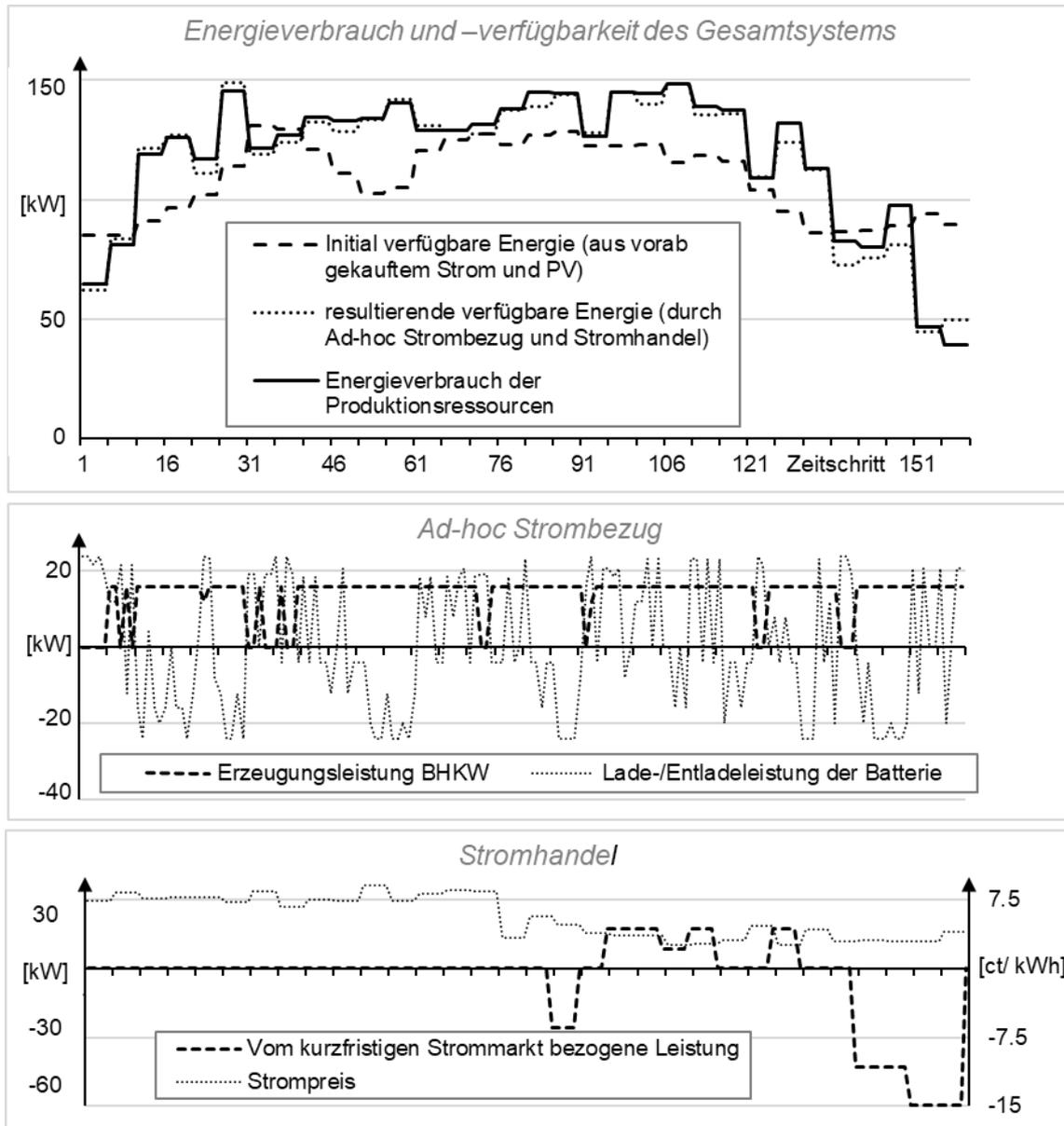


Abbildung 57: Systemverhalten in einer beispielhaften Episode
(in Anlehnung an ROESCH ET AL. 2018)

Abbildung 57 gibt einen Einblick in das Verhalten des mit MARL3 gesteuerten Gesamtsystems für eine beispielhafte Episode. Die Gegenüberstellung von Energieverbrauch und -verfügbarkeit zeigt deutlich, dass durch die RL basierte Steuerung des Ad-hoc Strombezugs und des Stromhandels die verfügbare Energiemenge an den tatsächlichen Verbrauch adaptiert werden kann und auf diese Weise

Lastgangabweichungen weitestgehend verhindert werden. Während die Batterie flexibel geladen und entladen wird und dadurch zahlreichen Zyklen-Wechsel ausgesetzt ist, wird das BHKW überwiegend kontinuierlich mit der kosteneffizienten Nennleistung betrieben. Der Marktagent wird in der zweiten Hälfte der Episode aktiv, indem zunächst Strom in mehreren Intervallen gekauft und gegen Ende eine größere Energiemenge verkauft wird.

Die vollständig reaktive Steuerung (VRS) der Produktion und des ad-hoc Strombezugs ohne aktiven Stromhandel stellt einen rein regelbasierten Ansatz mit geringer Komplexität und niedrigem Implementierungsaufwand dar. Aus diesem Grund wird angenommen, dass ein derartiges System schon heute in der Praxis eingesetzt werden kann. Daher wird dieser Ansatz als Referenz zur Beurteilung des Gesamtsystems herangezogen. Die erzielten Ergebnisse mit SA waren zwar deutlich besser. Allerdings ist aufgrund der langen Rechenzeit von bis zu 2,5 Stunden für eine Umplanung ein derartiger Lösungsalgorithmus im Hinblick auf die erforderliche Reaktionsfähigkeit nicht geeignet. Insgesamt kann dabei mit dem entwickelten MARL-Gesamtsystem für die untersuchten Szenarien eine Kostenreduzierung um knapp die Hälfte erzielt werden (vgl. Abbildung 58). Bezogen auf den vorliegenden Anwendungsfall bedeutet dies eine Kosteneinsparung von 272 € pro Schicht im Vergleich zu einer VRS.

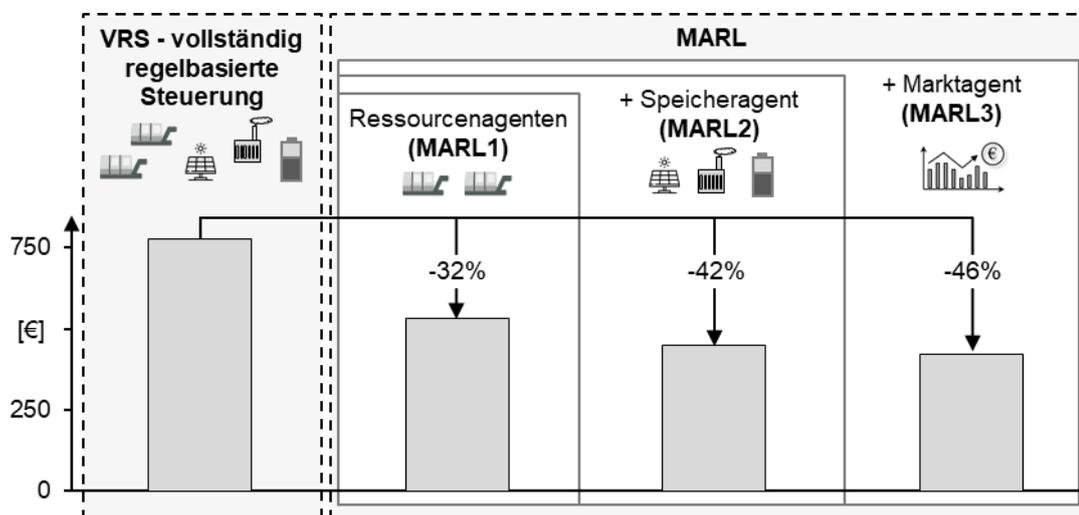


Abbildung 58: Gesamtkostenreduzierung durch Anwendung des entwickelten Systems im Vergleich zu einer vollständig regelbasierten Steuerung

Zusammenfassung

Die untersuchten Evaluierungsszenarien zeigen, dass jeder einzelne der drei Agententypen – Ressourcen-, Batterie- und Marktagent – einen Beitrag zur weiteren

Reduzierung der Gesamtkosten leistet (vgl. Abbildung 58). Zusätzlich geben die Untersuchungen mit 15 Ressourcenagenten (Evaluierungsszenario 1b) einen Hinweis darauf, dass die Skalierung des Systems auf größere Anwendungsbeispiele möglich ist. Insgesamt schneidet MARL in den betrachteten Szenarien durchwegs deutlich besser ab als vergleichbare, regelbasierte Ansätze. Die reaktiv-prädiktiven Benchmark-Lösungen mittels SA übertreffen hingegen in allen untersuchten Szenarien MARL um 29 bis 35 Prozent. Die langen Rechenzeiten für eine Umplanung von bis zu 2,5 Stunden stehen dagegen im Widerspruch zu der erforderlichen Reaktionsfähigkeit und schränken den realen Einsatz eines derartigen Lösungsalgorithmus ein. Dagegen ist im Zuge von MARL bei vorangegangenem Training eine Ausführung der Steuerung innerhalb von Sekunden möglich, was ein deutlicher Vorteil im Vergleich zu dem angewendeten reaktiv-prädiktiven Steuerungsansatz ist.

Im Zuge der Untersuchungen wurde das entwickelte System zur Steuerung eines Simulationsmodells eingesetzt. Um den dargelegten Nutzen auf ein reales System übertragen zu können, ist die Grundvoraussetzung, dass die für das Training verwendete Simulationsumgebung das tatsächliche Produktionssystem in der Realität ausreichend genau abbildet. Zur Überprüfung dieser Prämisse können Daten des Produktionsablaufs im realen System zur Validierung des Modells herangezogen werden, was jedoch im Rahmen dieser Untersuchung nicht ohne weiteres möglich war. Darüber hinaus wurde die entwickelte Lösung im Rahmen der vorliegenden Arbeit ausschließlich zur Steuerung des gezeigten Anwendungsfalls eingesetzt, bei dem es sich um ein einstufiges Produktionssystem handelt. Die Ergebnisse aus der Literatur zur Steuerung von Produktionssystemen, wie etwa JIMÉNEZ (2012) oder WASCHNECK ET AL. (2018), legen nahe, dass die entwickelte Lösung auch auf mehrstufige Produktionssysteme übertragen werden kann. Im Hinblick auf die Kostenstruktur in Unternehmen kann das System durch die parametrierbaren Kostenfaktoren für eine breite Zahl an Unternehmen eingesetzt werden. Es ist zu erwarten, dass der Nutzen für Anwendungsfälle, bei denen die Energiekosten einen hohen Anteil an den Gesamtkosten einnehmen, durch das entwickelte System ansteigt. Daneben kann sich der Mehrwert des kurzfristigen Stromhandels erhöhen, falls der prognostizierte Anstieg von Strompreisschwankungen am Intraday-Markt eintritt.

8.4 Nutzen und Wirtschaftlichkeitsbetrachtung

Aufbauend auf der simulationstechnischen Evaluierung wird im Folgenden das sich aus der Anwendung des Steuerungssystems ergebende wirtschaftliche Potenzial bewertet. Zu diesem Zweck wird nach BUSSE VON COLBE ET AL. (2015) die statische Investitionsrechnung zur Ermittlung der Amortisationszeit für die Implementierung und Nutzung des Steuerungssystems herangezogen. Um die Rentabilität beziffern zu können, muss auf der einen Seite der Aufwand für die Umsetzung des Systems quantifiziert werden. Auf der anderen Seite ist die durch den Einsatz der Steuerung erzielbare Einsparung abzuschätzen. Die Größenordnung dieses Potenzials geht aus den Ergebnissen in Abschnitt 8.3.2 hervor.

Der Großteil der anfallenden Investition ist auf die Personalaufwände zur Programmierung und Installation des Gesamtsystems zurückzuführen. Da es sich um ein reines Softwaresystem handelt, sind keine nennenswerten Investitionen in Hardware erforderlich. Ferner wird davon ausgegangen, dass ein Energiemanagementsystem zur Online-Erfassung von Energiedaten und eine digitale Vernetzung des Produktionssystems bereits vorhanden sind. Insgesamt lassen sich die erforderlichen Tätigkeiten in sieben einzelne Positionen gliedern, deren jeweiliger Aufwand in Tabelle 5 abgeschätzt ist.

Um alle relevanten Informationen und Daten aus dem Produktionssystem inklusive der Energieerzeugung anzubinden, sind im ersten Schritt der Aufbau eines geeigneten Datenmodells und die Implementierung der erforderlichen Schnittstellen erforderlich (*P1*), etwa zur Verlinkung der Informationen über zu bearbeitende Aufträge und Energieverbräuche. Ein zentraler Baustein ist der Aufbau der Simulationsumgebung, bei dem die einzelnen Agenten angelegt, der Ablauf der Simulation und die Vergabe der Belohnung implementiert werden (*P2*). Die Modelle des Produktionssystems (*P3*) sowie der Eigenerzeugungsanlagen und des Stromspeichers (*P4*) müssen anschließend in die entwickelte Simulationsumgebung integriert werden. Auf dieser Basis kann die damit lauffähige Simulation an RLlib angebunden, der Zustands- und Aktionsraum können konfiguriert werden (*P5*). Die Feinjustierung der Parameter erfolgt im Zuge des Hyperparameter-Tunings anhand einer strukturierten experimentellen Parametersuche (*P6*). Parallel zu den letzten Schritten kann mit dem Aufbau der Benutzerschnittstelle begonnen werden, wobei die Visualisierung der Ergebnisse im Vordergrund stehen sollte (*P7*).

Prototypische Umsetzung und Evaluierung

Tabelle 5: Aufstellung der durch die Implementierung anfallenden Personalkosten

Einmalige Investitionskosten			
Personalkosten:			
Position	Bezeichnung	Personen- tage (PT)	Kosten (900.00 €/PT)
P1	Datenmodell und Schnittstellen	20	18.000 €
P2	Aufbau Simulationsumgebung	30	27.000 €
P3	Datenerhebung und Modellierung Produktionssystem	35	31.500 €
P4	Modellierung Eigenerzeugungsanlagen und Energiespeicher	10	9.000 €
P5	Konfiguration und Anbindung von Rllib	15	13.500 €
P6	Parametrierung und Tuning	20	18.000 €
P7	Benutzerschnittstelle	5	4.500 €
Summe Personalkosten:			121.500 €
Sachkosten:			
Position	Bezeichnung	Kosten	
S1	Einmalige Gebühr für Marktzugang	5.000 €	
Summe Sachkosten:			5.000 €
Summe einmalige Investitionskosten:			126.500 €

Um am kurzfristigen Stromhandel teilzunehmen, sind des Weiteren einmalige Gebühren für den Marktzugang zu entrichten. In Abhängigkeit davon, ob das Unternehmen selbst direkt an der Strombörse agiert oder indirekt über einen Aggregator oder Broker am Markt teilnimmt, fallen die Gebühren unterschiedlich hoch aus. Für die Nutzenkalkulation wird von letzterem Szenario ausgegangen, also dem Handel mittels eines Service-Brokers, wodurch einmalige Kosten in Höhe von 5.000 € anfallen (EPEX SPOT 2019).

Um die Amortisationszeit zu berechnen, sind neben den einmaligen Investitionskosten die Nettoeinsparungen pro Jahr zu berechnen, welche sich aus der Differenz zwischen der erzielbaren Kostenreduzierung und den jährlichen Betriebskosten zusammensetzen. Ein wesentlicher Einflussparameter zur Bewertung der jährlichen Betriebskosten ist die Häufigkeit, mit der ein erneutes Training des Gesamtsystems angestoßen werden muss, da dies einen Zugriff auf entsprechende Rechenkapazitäten erfordert. Dieser Faktor hängt direkt von der Frequenz ab, mit der

sich Randbedingungen im Produktionssystem ändern. So können beispielsweise neue Auftragsstypen oder geänderte Verfügbarkeiten und Störungswahrscheinlichkeiten eine Anpassung der Modellparameter und damit ein erneutes Training erforderlich machen. Darüber hinaus hat auf Seiten der erzielbaren Kostenreduzierung die jährliche Nutzungsdauer des Systems einen wesentlichen Einfluss auf die Wirtschaftlichkeit. Aus diesen beiden Gründen werden für die weitere Untersuchung folgende Szenarien gebildet:

- **Nutzungsszenarien:**
 - *Nutzungsszenario 1:* Einschichtbetrieb an 260 Arbeitstagen im Jahr
 - *Nutzungsszenario 2:* Zweischichtbetrieb an 260 Arbeitstagen im Jahr
- **Trainingsszenarien:**
 - *Trainingsszenario a:* Gelegentliche Änderungen im Produktionssystem machen durchschnittlich 0,5 Mal pro Woche ein erneutes Training erforderlich
 - *Trainingsszenario b:* Häufige Änderungen im Produktionssystem machen durchschnittlich 1,5 Mal pro Woche ein erneutes Training erforderlich

Aus der Kombination der beiden Nutzungs- und Trainingsszenarien gehen vier Fälle hervor, die eine erste Einordnung des Einflusses der jährlichen Nutzungsdauer und des erforderlichen Trainingsaufwands zulassen. Während sich die Trainingsszenarien auf die jährlichen Betriebskosten auswirken, beeinflussen die Nutzungsszenarien maßgeblich die erzielbare Kostenreduzierung.

Die jährlichen Betriebskosten setzen sich aus insgesamt drei Kostenfaktoren zusammen (vgl. Tabelle 6). Da zur technischen Umsetzung des Gesamtsystems ausschließlich Open-Source-Softwarebibliotheken verwendet werden, fallen im Betrieb keine Lizenzgebühren oder Ähnliches an. Dagegen wird für das Training der RL-Agenten eine ausreichende Rechnerinfrastruktur benötigt. Falls nicht bereits bestehende Ressourcen genutzt werden können, ist die in den meisten Fällen einfachste Lösung, auf die Dienste eines kommerziellen Anbieters für Cloud-Rechenleistung (*BI*) zurückzugreifen. Zur Nutzung einer virtuellen Maschine mit einem Arbeitsspeicher über 100 GB und zwei GPUs werden etwa drei € pro Stunde veranschlagt.² Für die Berechnung der beiden Trainingsszenarien wird eine durchschnittlicher Zeitbedarf pro, woraus eine Gesamttrainingsdauer von 36 (*Trainingsszenario a*) bzw. 108 (*Trainingsszenario b*) Stunden pro Woche resultiert. Darüber

² Die Preise basieren auf Angaben eines international führenden Anbieters von Cloud-Rechenleistung.

Prototypische Umsetzung und Evaluierung

hinaus fallen zusätzliche jährliche Kosten für den Marktzugang an der Strombörse (B2), die im Zuge des Handels mit einem Broker pro Jahr 5.000 € betragen, und die Systemwartung (B3) an (EPEX SPOT 2019).

Tabelle 6: Geschätzte Nettoeinsparung pro Jahr

Nettokosteneinsparung pro Jahr					
Jährliche Betriebskosten					
Position	Trainings-szenario	Beschreibung	Preis	Dauer	Kosten
B1	a	Cloud-Rechenleistung für 50 Betriebswochen im Jahr	3,00 € pro h	36 h pro Woche	5.400 €
	b			108 h pro Woche	16.200 €
B2	alle	Gebühr für Marktzugang			5.000 €
B3	alle	Systemwartung			7.000 €
Summe jährlicher Betriebskosten					
Trainings-szenario a:		17.400 €		Trainings-szenario b	
				28.200 €	
Erzielbare Kostenreduzierung					
pro Schicht	Nutzungs-szenarien	Schichten pro Tag	Arbeitstage pro Jahr	Gesamt	
272.00 €	1	1	260	70.720 €	
	2	2		141.440 €	
Gesamte Nettokosteneinsparung pro Jahr:					
Szenario 1a:		53.320 €		Szenario 2a:	
				124.040 €	
Szenario 1b:		42.520 €		Szenario 2b:	
				113.240 €	

Auf Seiten der Kosteneinsparung wird der auf MARL basierende Ansatz wie in Evaluierungsszenario 3 einer regelbasierten Steuerung des Gesamtsystems ohne aktiven Stromhandel gegenübergestellt. Aufgrund der langen Rechenzeit wird die Anwendung eines prädiktiven Lösungsansatzes, wie etwa des im Evaluierungsszenario verwendeten Simulated Annealing, kategorisch ausgeschlossen. Im Vergleich zum regelbasierten Ansatz ergibt sich nach Abschnitt 8.3.2 im vorliegenden Anwendungsbeispiel eine Kosteneinsparung von durchschnittlich 272 € pro Schicht. Bei geschätzten 260 Arbeitstagen pro Jahr im Einschichtbetrieb (*Nutzungsszenario 1*) ergibt sich eine jährliche Einsparung in Höhe von 70.720 €. Demgegenüber liegt die erzielbare Kostenreduzierung im Falle des Zweischichtbetriebs von *Nutzungsszenario 2* bei 141.440 €. Alle laufenden Kosten und die gesamte erzielbare Nettokostenreduzierung sind in Tabelle 6 zusammengefasst.

Insgesamt bewegt sich die nach KRUSCHWITZ (2005) berechnete Amortisationszeit unter den getroffenen Annahmen für das Nutzungsszenario im Einschichtbetrieb zwischen 2,4 und 3 Jahren. Im Falle eines Zweischichtbetriebs liegt der Wert je nach Trainingsszenario im Bereich von 1,0 und 1,1 Jahren (vgl. Abbildung 59). Die Auswertung macht deutlich, dass der wirtschaftliche Nutzen wesentlich von der jährlichen Nutzungszeit abhängt. Zudem steigen in Produktionssystemen mit häufigen Änderungen der Randbedingungen, wie etwa neuen Auftragsstypen, Verfügbarkeiten oder Störungswahrscheinlichkeiten, die laufenden Kosten, da dies ein erneutes Training und somit zusätzliche Ausgaben für Rechenleistung erfordert.

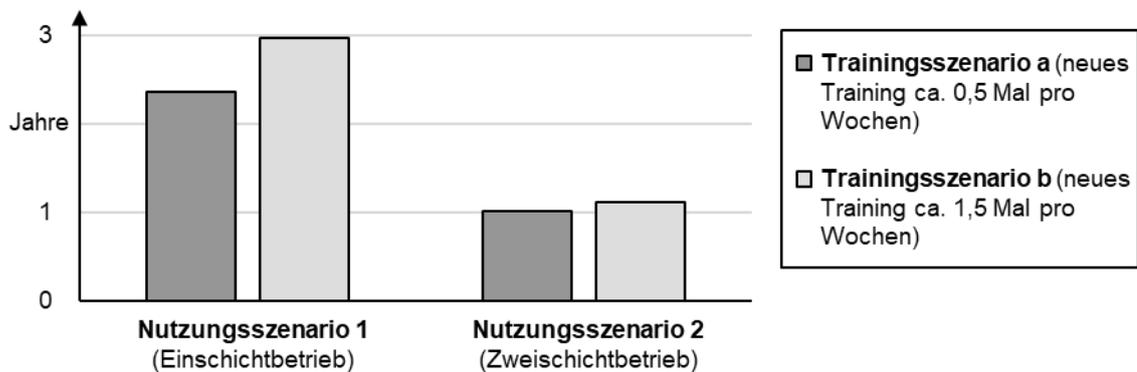


Abbildung 59: Resultierende Amortisationszeit für die betrachteten Szenarien

Es bleibt festzuhalten, dass der dargelegte erzielbare Erlös und damit die Amortisationszeit lediglich das Potential der für das Anwendungsbeispiel herangezogenen Randbedingungen widerspiegeln. Zur Beurteilung des Nutzens im Falle von geänderten Einflussgrößen wie geringere Strompreisschwankungen oder höheren Produktionskosten ist eine spezifische Bewertung erforderlich. Im Allgemeinen können potentiell zusätzliche Skaleneffekte erzielt werden, etwa durch die Nutzung des Ansatzes für ein umfangreiches Produktionssystem, mit einer größeren Anzahl an Ressourcen und höherem Energiebedarf. Die Kosten für die Installation und den Betrieb der Eigenerzeugungsanlagen und des Batteriespeichers wurden in der durchgeführten Bewertung nicht berücksichtigt. Deren wirtschaftlich optimale Auslegung erfordert eine spezifische Betrachtung. Das im Zuge der vorliegenden Arbeit entwickelte System ermöglicht lediglich eine bessere Nutzung dieser Energiebezugsoptionen.

9 Zusammenfassung und Ausblick

9.1 Zusammenfassung

Bei der Eindämmung des anthropogenen Klimawandels spielen die Transformation und Dekarbonisierung der Stromerzeugung eine entscheidende Rolle. Die damit einhergehende Substitution konventioneller Kraftwerke durch Wind- und PV-Anlagen zieht aufgrund ihrer witterungsabhängigen Erzeugungsleistung zunehmende Strompreisschwankungen nach sich und führt im Allgemeinen zu steigenden Strompreisen in Deutschland. Demgegenüber kann auf Basis von dezentralen Kleinkraftwerken Strom zu attraktiven Konditionen vor Ort produziert werden. Vor diesem Hintergrund bietet sich für Industrieunternehmen die Möglichkeit, durch die gezielte Steuerung des Strombedarfs die Preisschwankungen zu ihrem wirtschaftlichen Vorteil zu nutzen und gleichzeitig den Stromeigenverbrauch zu erhöhen. Dabei können einerseits stationäre Batteriespeicher eingesetzt werden. Andererseits ist die intelligente Steuerung der Produktion ein wichtiger Baustein, um den Energieverbrauch gezielt anzupassen. Da derartige Ansätze die logistischen Zielgrößen der Produktion, wie etwa Durchlaufzeit oder Termintreue, negativ beeinflussen können, sind potenzielle Energiekosteneinsparungen mit den Auswirkungen auf die logistischen Zielgrößen abzuwägen. Insgesamt sind die Produktion, Energiemarkt und -erzeugung stochastischen Ereignissen ausgesetzt, wodurch eine integrierte und reaktionsschnelle Steuerung dieser Elemente erforderlich ist.

Ausgehend von dieser Motivation lässt sich konstatieren, dass bestehende Ansätze zur Produktionssteuerung keine Integration von Eigenerzeugung, Batteriespeicher und kurzfristigem Stromhandel ermöglichen. Darüber hinaus geraten herkömmliche Steuerungsverfahren aufgrund der komplexen Problemstellung an ihre Grenzen. Um eine Vergleichbarkeit mit den Energiekosten sicherzustellen, sind zudem die logistischen Zielgrößen in ein gesamtkostenbasiertes Zielsystem zu überführen. Aus diesen Gründen ist das Ziel der vorliegenden Arbeit die Entwicklung eines *Systems zur energieorientierten und kostenbasierten Produktionssteuerung*, das in der Lage ist, sowohl den Energiebezug als auch die Produktion integriert zu steuern.

Ausgehend von dieser Zielsetzung wurden im Rahmen der vorliegenden Arbeit zunächst die relevanten Grundlagen des Energiebezugs produzierender Unterneh-

men, der Produktionsplanung und -steuerung sowie der existierenden Steuerungsverfahren aufgezeigt. Anschließend erfolgte eine Analyse des aktuellen Stands der Forschung. Der Fokus lag auf der operativen Nutzung von Energieflexibilität in der Produktion. Darüber hinaus erfolgte eine detaillierte Betrachtung der bestehenden Ansätze zur kostenbasierten Produktionssteuerung. Aus den daraus generierten Erkenntnissen konnte im nächsten Schritt der Handlungsbedarf identifiziert werden. Die Anforderungen ließen sich ableiten.

Auf Basis der definierten Anforderungen wurde anschließend mit Reinforcement Learning das für den vorliegenden Anwendungsfall am besten geeignete Steuerungsverfahren ausgewählt. Da es sich dabei um einen agentenbasierten Ansatz handelt, wurde eine dezentrale Systemarchitektur mit mehreren Ressourcenagenten sowie jeweils einem Speicher- und Marktagenten konzipiert. Die Agenten handeln durch Aktionen autonom innerhalb der Umgebung, die sich aus dem Produktionssystem, den Eigenerzeugungsanlagen, dem Speicher und dem kurzfristigen Stromhandel zusammensetzt. Der aktuelle Zustand der Umgebung dient als Input für die Agenten, die durch ihre Aktionen eine Maximierung ihrer Belohnung in Form von Produktions- und Energiekosten anstreben.

Zur *Integration des kurzfristigen Energiebezugs* in die Produktionssteuerung wurden im nächsten Schritt die relevanten Bezugsoptionen kategorisiert und analysiert. Auf Basis der durchgeführten Modellierung konnte ein Steuerungsschema abgeleitet werden. Im Zuge dessen handelt ein Marktagent mit einer Vorlaufzeit von neun Minuten an der Strombörse, während der ad-hoc Strombezug, bestehend aus Eigenerzeugung und Batteriespeicher, mit Hilfe eines Speicheragenten kurzfristig gesteuert wird. Um eine Vergleichbarkeit zwischen Energiekosten und den logistischen Zielgrößen herzustellen, wurde darüber hinaus ein *gesamtkostenbasiertes Zielsystem* hergeleitet, in dem alle für die Produktionssteuerung relevanten Kosten abgebildet sind. Als Ausgangsbasis diente hierbei die durchgeführte Literaturrecherche. Die identifizierten Kostenfaktoren wurden abschließend modelliert und in ein generisches Datenmodell überführt.

Diese Vorarbeiten bildeten die Basis für die Entwicklung der *integrierten Produktionssteuerung mittels Reinforcement Learning*. Neben der Modellierung des Produktionssystems erfolgte die Konzeptionierung der Zustands- und Aktionsräume der einzelnen Agenten sowie der Belohnungsvergabe. Zudem wurde ein geeignetes Lernverfahren ausgewählt und eine passende Lernumgebung geschaffen.

Das entwickelte System wurde abschließend prototypisch umgesetzt und anhand eines realen Anwendungsbeispiels validiert. Bei der technischen Umsetzung war

die Software-Bibliothek RLlib das zentrale Element für das Lernen und Trainieren der Agenten, während die Umgebung in Form einer Simulation eigens entwickelt wurde. Um die Funktionalitäten aller Systemelemente nachvollziehbar untersuchen zu können, erfolgte die Evaluierung stufenweise mittels drei Szenarien. Insgesamt zeigen die Ergebnisse, dass der entwickelte Ansatz in der Lage ist, sowohl die Energie- als auch die Produktionskosten zu minimieren und dabei ein komplexes System aus Produktionsressourcen, Stromerzeugungsanlagen und aktivem Stromhandel zu steuern. Während die Lösungsgüte im Vergleich zu einer Berechnung mittels der Metaheuristik Simulated Annealing abfällt, ist die kurze Reaktionszeit ein wesentlicher Vorteil für den Einsatz des Systems in der Produktionssteuerung. Mit Earliest-Deadline-First konnte zudem ein verbreiteter reaktiver Steuerungsansatz deutlich übertroffen werden.

9.2 Ausblick

Im Rahmen der vorliegenden Arbeit wurde ein Ansatz zur integrierten Steuerung von Produktion, Eigenerzeugung, Batteriespeicher und kurzfristigem Stromhandel entwickelt. Das System ist in der Lage, die Energie- und die relevanten Produktionskosten zu minimieren.

Der Betrachtungshorizont der entwickelten Steuerung beschränkt sich lediglich auf einen Tag. Die Produktionsplanung wurde im Zuge der Arbeit genauso wenig wie die Möglichkeiten zum Stromhandel außerhalb des Intraday-Markts berücksichtigt. Um einen maximalen Nutzen zu generieren, ist jedoch die Produktionsplanung mit der entwickelten Produktionssteuerung zu verzahnen. Die zeitlich vorgelagerten Strommärkte sind in die Beschaffungsstrategie zu integrieren. Dabei gilt es, den für die Produktionssteuerung entwickelten kostenbasierten Ansatz auf die Produktionsplanung zu übertragen. Auf diese Weise kann der Handlungsspielraum für die Produktionssteuerung situationsabhängig angepasst werden, beispielsweise unter Berücksichtigung der Strompreisprognose. Dies verspricht eine effizientere Vermarktung des Energieflexibilitätpotenzials.

Als Lösungsverfahren wurde mit MARL ein Ansatz ausgewählt, der aktuell Gegenstand intensiver Forschungen ist. Insbesondere die Vergleiche mit der Metaheuristik Simulated Annealing haben dabei weitere Optimierungspotenziale dieses Ansatzes aufgezeigt. Obwohl mit einer Actor-Critic-Variante von Proximal Policy Optimization ein Lernalgorithmus ausgewählt wurde, der zum aktuellen

Zeitpunkt für diskrete Aktionsräume in vielen Benchmarks hervorragend abschneidet, kann das System teilweise lokale Optima nicht überwinden. Neue Lernverfahren sowie Methoden zu einer besseren Ausbalancierung der Exploration und zum Ausnutzen erfolgsversprechender Aktionen (engl. Exploitation) beim Training können einen wichtigen Beitrag leisten, die Lösungsgüte des Systems weiter zu verbessern.

Der alleinige Fokus der vorliegenden Arbeit liegt auf der elektrischen Energie. Andere Energieformen, wie etwa Wärme, wurden nicht betrachtet. Im Kontext der intelligenten Steuerung von Produktion und Eigenerzeugung auf Basis von KWK kann durch die integrierte Betrachtung von Wärme die Effizienz des Systems möglicherweise weiter gesteigert werden. Zudem wurde im Hinblick auf den Energiebedarf einzelner Aufträge zur Vereinfachung ein konstanter Leistungsverbrauch angenommen. Bearbeitungsschritte, deren Energiebedarf sich während der Bearbeitungsdauer stark ändert, können auf diese Weise nicht realitätsnah abgebildet werden. Eine Weiterentwicklung des Systems zur Berücksichtigung von auftragsindividuellen Lastverläufen erscheint daher vielversprechend.

10 Anhang

10.1 Parameter des Batteriemodells

Zur Parametrierung des in Abschnitt 5.2.3 hergeleiteten Batteriemodells wurden die in Tabelle 7 aufgeführten Werte verwendet. Die Ermittlung der Zyklen erfolgte anhand der Python-Bibliothek *rainflow 3.0.0*.

Tabelle 7: Parameter des Batteriemodells

Parameter	Formelzeichen	Wert
Wirkungsgrad beim Laden beziehungsweise Entladen der Batterie	η	0,96
SOC-Stress Koeffizient	k_σ	1,04
Referenz SOC-Wert	σ_{ref}	0,5
Empirische Koeffizienten für die DoD bedingte Alterung	$k_{\delta 1}; k_{\delta 2}; k_{\delta 3}$	$1,5 \cdot 10^5; -0,501;$ $-1,23 \cdot 10^5$
Empirischer Koeffizient für die kalendrische Alterung	k_t	$4,14 \cdot 10^{-10}$
Referenztemperatur	T_{ref}	25° C
Empirischer Koeffizient für die temperaturbedingte Alterung	k_T	$6,93 \cdot 10^{-2}$

10.2 Konfiguration MARL

Tabelle 8: Verwendete Parameter in RLib für MARL-3 (Version 0.8.0)

Parameter	Wert
Fcnet_hiddens	[128, 1024] (no LSTM or shared layers)
simple_optimizer	True
sgd_mini_batchsize	128
sample_batch_size	2000
train_batch_size	40000
use_gae	True
lambda	1
kl_coeff	0.2
num_sgd_iter	30
learning_rate	5.00E-05
vf_loss_coeff	1

entropy_coeff	0
clip_param	0.3
vf_clip_param	10
grad_clip	None
kl_target	0.01
batch_mode	truncate_episodes
observation_filter	NoFilter
distributed_data_parallel_optimizer	False

10.3 Produktionsdaten der Evaluierungsepisoden

Die zugehörigen Produktionsdaten der für die Benchmarks herangezogenen Evaluierungsszenarien sind in den Tabellen 9 bis 19 aufgelistet. Für das Evaluierungsszenario 1b mit 15 Agenten wurden diese Daten mit zwei weiteren, analog generierten Datensätzen erweitert, die sich lediglich im Hinblick auf die Fertigstellungstermine und Verteilung der Aufträge unterscheiden. Bei den fett markierten Aufträgen handelt es sich um Eilaufträge. Die Fertigstellungstermine werden durch Anzahl der Drei-Minuten-Zeitschritte seit Schichtbeginn ausgedrückt.

Tabelle 9: Daten von Evaluierungsepisoden 1

Ressource	Auftragstyp	Anzahl	Fertigstellungstermine
0	0	6	53, 73, 93, 114, 129, 159
	1	5	41, 85 , 125, 136, 156
	2	2	81 , 113
1	3	4	41, 103, 153, 159
	4	4	75, 84, 104 , 132
	5	3	78, 102, 123
2	6	2	59, 135, 159
	7	5	53, 72 , 94, 115, 127
	8	4	64, 92, 108, 156
3	9	8	46, 105 , 119, 120 , 126, 128, 153, 157
	10	6	63, 72, 89, 100, 111, 159
4	11	2	74 , 140
	12	3	98, 156, 159
	13	1	107
	14	3	52, 74, 86

Tabelle 10: Daten von Evaluierungsepisoden 2

Ressource	Auftragstyp	Anzahl	Fertigstellungstermine
0	0	4	69, 75, 85, 134
	1	4	82, 93 , 101, 114
	2	5	70, 125, 140, 147, 159
1	3	4	54, 103, 122, 122
	4	4	42, 69, 88, 97
	5	4	114, 122, 159, 159
2	6	3	54, 116, 139
	7	6	64, 77, 118, 119, 159, 159
	8	3	95, 127, 148
3	9	4	74, 87, 114, 124
	10	9	43, 57, 78, 92, 117, 121, 129, 151, 159
4	11	2	61, 159
	12	4	66, 85, 129, 145
	13	2	62, 112
	14	2	76 , 118

Tabelle 11: Daten von Evaluierungsepisoden 3

Ressource	Auftragstyp	Anzahl	Fertigstellungstermine
0	0	5	53, 56, 64, 95, 159
	1	4	81, 141, 146, 159
	2	5	71, 100, 126, 131, 146
1	3	3	58, 98, 105
	4	5	90, 104, 148, 156, 159
	5	4	57, 65, 95 , 132
2	6	4	103, 115, 152, 159
	7	8	48, 58, 69, 108, 125, 128, 130, 159
	8	0	
3	9	6	69, 79, 87, 96, 145, 155
	10	9	61 , 74, 87, 83, 119, 128, 143, 145, 159
4	11	1	124
	12	3	118, 148, 159
	13	3	66, 97 , 143
	4	4	49, 60, 116, 139

Tabelle 12: Produktionsaufträge von Evaluierungsepisoden 4

Ressource	Auftragstyp	Anzahl	Fertigstellungstermine
0	0	2	91, 108, 138
	1	5	53, 81 , 86, 96 , 129, 159
	2	4	41, 42, 88, 126
1	3	1	45
	4	7	77, 102, 110, 128, 151 , 158, 159
	5	4	41, 100, 149, 152
2	6	5	63, 131, 136, 137 , 145
	7	2	68, 101
	8	3	96, 99, 159
3	9	5	57, 58, 91, 91, 136
	10	10	66, 69, 98, 98 , 101, 119, 125, 133, 136, 159
4	11	4	55, 75, 88, 154
	12	1	152
	13	2	124, 159
	14	5	67, 86, 110, 138, 159

Tabelle 13: Produktionsaufträge von Evaluierungsepisoden 5

Ressource	Auftragstyp	Anzahl	Fertigstellungstermine
0	0	8	65, 74, 81, 97 , 128, 125, 146, 159
	1	1	44
	2	3	108, 109, 156
1	3	3	84, 136, 149
	4	5	58, 84, 99, 111, 159
	5	3	57, 125, 159
2	6	2	150, 159
	7	7	60, 69, 72 , 99, 131, 144, 153
	8	3	59, 97, 114
3	9	6	46, 49, 74, 88, 109, 123
	10	8	46, 109, 111, 119, 126 , 131, 149, 159
4	11	5	64, 65, 110, 151, 154
	12	2	125, 159
	13	3	103, 119, 125
	14	0	

Tabelle 14: Produktionsaufträge von Evaluierungsepisoden 6

Ressource	Auftragstyp	Anzahl	Fertigstellungstermine
0	0	6	60, 78, 99 , 107, 119, 137
	1	3	120, 128, 155
	2	6	63, 70, 84, 97, 140 , 159
1	3	2	83, 153
	4	6	95, 97, 116, 125, 131, 140
	5	5	39, 55, 106 , 137, 141
2	6	4	47, 55, 81, 127
	7	4	127 , 131, 159
	8	4	53, 81, 139, 149
3	9	4	48, 64, 95, 121, 126 , 159
	10	9	44, 46, 85, 102, 116, 120, 145, 148, 159
4	11	1	152
	12	1	65, 102, 107, 126, 159
	13	2	73, 81, 122, 153
	14	5	78, 124

Tabelle 15: Produktionsaufträge von Evaluierungsepisoden 7

Ressource	Auftragstyp	Anzahl	Fertigstellungstermine
0	0	7	60, 86, 87, 87, 116, 135, 159
	1	2	151, 151
	2	5	29, 52, 102 , 123, 145
1	3	3	42, 59, 87
	4	4	66, 67, 143, 159
	5	5	116, 126, 137, 146, 159
2	6	4	55, 102, 122, 154
	7	3	59, 84, 109
	8	3	64, 96 , 159
3	9	10	45, 48, 68 , 106, 117, 124, 125 , 128, 143, 159
	10	6	46, 61, 80, 100, 123, 129
4	11	4	61, 79
	12	1	87
	13	3	81, 116, 139, 159
	14	3	56, 58 , 146, 153, 159

Tabelle 16: Produktionsaufträge von Evaluierungsepisoden 8

Ressource	Auftragstyp	Anzahl	Fertigstellungstermine
0	0	6	64, 83, 94, 135, 152, 154
	1	6	49, 84, 100, 112, 141, 159
	2	3	68, 79, 128
1	3	1	159
	4	4	71, 86 , 119, 123
	5	5	70, 76, 119, 122, 138
2	6	3	41, 91, 159
	7	2	107, 121
	8	5	68, 106 , 111, 124, 147
3	9	9	60, 60, 79, 98, 106 , 128, 140, 142, 146
	10	6	57, 83 , 87, 87, 131, 159
4	11	1	65
	12	5	67, 102, 131, 135, 159
	13	2	109, 159
	14	3	53, 108, 151

Tabelle 17: Produktionsaufträge von Evaluierungsepisoden 9

Ressource	Auftragstyp	Anzahl	Fertigstellungstermine
0	0	5	53, 56, 78 , 99, 121
	1	6	80, 84, 111, 127, 147, 159
	2	2	48, 140
1	3	1	94
	4	2	54, 117
	5	6	69, 72, 108, 128, 151, 159
2	6	3	74, 119 , 140, 159
	7	3	72, 72, 159
	8	4	81, 83, 95, 153
3	9	11	41, 59, 68, 91, 100, 109 , 110, 124, 146, 153, 159
	10	5	73, 103, 106, 138, 152
4	11	3	43, 157, 159
	12	3	80, 98 , 142
	13	2	78, 79
	14	3	62, 81, 142

Tabelle 18: Produktionsaufträge von Evaluierungsepisoden 10

Ressource	Auftragstyp	Anzahl	Fertigstellungstermine
0	0	4	50, 107, 108 , 159
	1	4	67, 74, 91, 138
	2	6	97, 107, 123, 139, 147 , 159
1	3	1	43
	4	5	68, 104 , 115, 123, 159
	5	4	47, 109, 152, 159
2	6	5	53, 96, 97, 129, 134
	7	2	123 , 159
	8	3	49, 91, 158
3	9	4	89, 139, 142, 159
	10	9	52, 78, 79, 83 , 98, 108, 109, 124 , 141
4	11	2	122
	12	2	59, 132,
	13	2	91, 110, 159 , 159
	14	3	54, 111, 151

Tabelle 19: Parameter der verwendeten Auftragstypen

Auftrags-Typ	Bearbeitungs-dauer [min]	Lagerkosten [€/h]	Verzugskosten [€/h]	Rüstdauer [min]	Kosten pro Rüstvorgang [€]	Leistungsbedarf für Bearbeitung [kW]	Leistungsbedarf für Rüsten [kW]
0	27	10	80	9	8	48	24
1	21	6	48	6	8	33.6	16.8
2	36	10	80	6	8	48	24
3	39	11	88	6	8	36	18
4	21	8	64	6	8	32	16
5	42	10	80	9	8	38	19
6	36	9	72	6	8	27.2	13.6
7	27	7	56	9	8	32	16
8	30	9	72	12	8	22.4	11.2
9	21	10	80	6	8	19.2	9.6
10	21	10	80	12	8	16.8	8.4
11	27	5	88	9	8	34.2	17.1
12	42	8	64	6	8	32.4	16.2
13	21	9	72	6	8	36	18
14	42	6	48	9	8	28.8	14.4

10.4 Energiedaten der Evaluierungsepisoden

In Tabelle 20 sind die für die Evaluierungsszenarien 1a, 2 und 3 verwendeten PV-Daten abgebildet. Für das Evaluierungsszenario 1b wurden diese Erzeugungsdaten um den Faktor 3 skaliert. Daneben wurden für das Evaluierungsszenario 3 die in Tabelle 21 aufgeführten Strompreisverläufe verwendet. Tabelle 22 zeigt die vorab gekaufte elektrische Leistung der einzelnen Evaluierungsszenarien, die konstant über die gesamte Episode zur Verfügung steht.

Tabelle 20: Verwendete PV-Daten für die Evaluierungsszenarien [in kW]

Zeitschritt	Episode									
	1	2	3	4	5	6	7	8	9	10
1	0	11	24	3	15	35	5	0	11	5
2	0	19	25	8	14	18	3	0	10	5
3	2	27	27	14	16	23	10	1	11	5
4	8	24	29	15	19	19	18	4	10	8
5	15	18	31	16	25	29	23	6	15	8
6	10	16	31	18	27	53	35	10	12	9
7	20	20	35	12	23	20	61	14	10	10
8	36	29	35	12	22	10	60	15	11	11
9	39	40	35	13	6	38	52	23	14	11
10	23	38	36	19	19	33	39	25	18	12
11	18	45	38	42	13	14	27	39	26	17
12	34	47	36	36	49	33	25	45	22	20
13	41	38	41	49	17	65	48	36	9	23
14	51	40	42	44	22	65	53	31	24	31
15	49	36	42	87	33	24	59	29	28	44
16	116	51	42	66	56	48	50	28	24	36
17	69	60	42	45	37	14	56	49	38	40
18	21	34	42	51	28	19	60	50	22	42
19	15	44	42	29	48	64	51	58	49	52
20	23	50	42	73	48	64	49	52	40	43
21	15	47	44	62	41	65	55	48	53	48
22	15	50	30	36	51	47	41	48	38	50
23	42	46	27	35	22	47	46	43	25	51
24	52	52	31	30	51	46	45	47	35	47
25	42	39	27	27	47	28	30	43	38	43
26	31	15	37	25	36	16	19	34	46	45
27	52	13	35	25	52	52	5	69	63	40
28	56	26	31	17	47	57	5	58	73	37
29	32	28	27	14	40	16	6	48	70	34
30	10	8	18	16	18	11	7	7	42	30
31	8	1	21	14	26	11	16	3	29	23
32	4	6	20	7	17	12	9	2	19	23

Tabelle 21: In Evaluierungsszenario 3 angenommene Strompreise

Zeitschritt	Episode									
	1	2	3	4	5	6	7	8	9	10
1	5.7	5.7	3.3	5.3	3.8	4.9	7.3	4.6	4.0	4.8
2	5.6	5.8	6.0	5.4	2.0	5.0	8.2	4.3	3.9	4.2
3	4.9	3.8	5.8	3.8	1.8	5.7	7.6	3.7	4.2	4.1
4	5.9	3.6	5.8	3.7	3.2	5.4	7.8	2.4	4.6	4.1
5	6.3	4.4	3.9	4.4	2.6	6.5	7.7	3.0	8.6	3.5
6	6.2	3.8	4.8	3.8	3.4	5.3	7.2	1.6	6.2	4.2
7	7.1	3.2	6.6	3.2	2.9	9.2	8.3	1.5	3.2	3.9
8	6.6	3.0	5.6	3.0	1.5	7.5	6.7	1.7	1.9	3.2
9	6.8	0.5	6.1	0.5	4.7	5.8	7.5	2.5	3.0	4.8
10	6.5	3.1	7.5	2.9	3.3	8.2	7.3	2.9	5.9	5.8
11	6.6	2.2	5.9	2.2	2.8	6.1	9.0	3.1	3.2	4.1
12	6.5	2.6	6.1	2.7	1.9	7.3	7.4	3.2	15.3	3.6
13	6.4	3.4	5.1	3.5	4.7	7.4	8.2	3.9	4.8	3.3
14	6.5	3.1	4.4	3.2	4.0	6.8	8.5	4.5	6.9	2.8
15	6.5	3.5	4.7	3.4	3.0	6.3	8.4	4.1	6.3	6.1
16	5.8	2.5	5.1	2.5	3.1	6.7	3.3	5.0	3.5	3.6
17	8.6	8.4	5.4	8.8	5.0	6.2	5.6	5.3	6.6	4.3
18	7.9	2.8	3.9	2.7	3.4	4.2	4.7	4.2	4.0	3.1
19	7.7	2.6	1.1	2.5	3.0	3.6	3.8	4.2	3.4	3.1
20	7.9	1.8	0.5	1.7	3.0	3.6	3.5	3.9	2.2	3.0
21	6.2	2.5	2.6	2.3	3.6	2.9	3.5	2.8	4.0	4.2
22	5.7	1.2	3.6	1.2	3.9	5.5	2.5	3.1	3.6	3.3
23	5.2	1.0	1.7	1.0	1.9	8.6	2.6	2.7	2.9	3.7
24	4.1	-0.2	3.3	-0.2	1.0	5.9	3.1	3.3	2.0	3.4
25	6.1	0.8	3.9	0.8	3.6	6.0	4.7	4.2	2.0	2.7
26	4.2	0.6	3.4	0.6	2.4	7.5	2.6	3.7	7.9	3.6
27	4.4	1.4	2.9	1.5	2.3	6.6	4.2	3.7	4.5	3.7
28	4.0	2.9	3.7	3.1	2.9	5.8	2.9	1.8	2.9	5.1
29	3.8	1.0	2.9	1.0	3.6	5.6	3.0	1.9	2.6	3.0
30	4.3	1.1	3.4	1.1	3.1	6.1	2.9	2.6	3.0	2.8
31	5.8	3.0	3.1	2.9	2.3	6.5	2.9	3.5	2.7	1.9
32	6.1	3.0	3.9	2.9	3.7	5.8	4.0	2.8	3.6	3.2
Durchschnitt:	6.0	2.8	4.3	2.7	3.0	6.1	5.5	3.3	4.5	3.7

Tabelle 22: Vorab gekaufte Energiemenge in den Evaluierungsepisoden [in kW]

Evaluierungs- szenario	Episoden									
	1	2	3	4	5	6	7	8	9	10
1a	72.8	76.8	81.6	73.2	75.2	82.8	81.6	74	71.6	69.6
1b	218.4	230.4	244.8	219.6	225.6	248.4	244.8	222	214.8	208.8
2	72.8	76.8	81.6	73.2	75.2	82.8	81.6	74	71.6	69.6
3	72.8	76.8	81.6	73.2	75.2	82.8	81.6	74	71.6	69.6

10.5 Regelbasierte Betriebsstrategie von Eigenerzeugung und Speicher (RBS)

Input:	$P_{V,t}$ # im aktuellen Zeitschritt verbrauchte Energiemenge $P_{PV,t}$ # im aktuellen Zeitschritt aus PV verfügbare Energiemenge $P_{gek,t}$ # für den aktuellen Zeitschritt bereits gekaufte Energiemenge $P_{BHKW,N}$ # Nennleistung des BHKW $P_{BHKW,min}$ # minimal mögliche Erzeugungsleistung in Teillast $P_{Bmax,t}$ # maximale Entladeleistung der Batterie im aktuellen Zeitschritt $P_{Bmin,t}$ # maximale Ladeleistung der Batterie im aktuellen Zeitschritt
Output:	$P_{BHKW,t}$ # Erzeugungsleistung des BHKW im aktuellen Zeitschritt t $P_{Bat,t}$ # Lade- bzw. Entladeleistung der Batterie im aktuellen Zeitschritt t
<pre> 1: if $P_{V,t} - (P_{PV,t} + P_{gek,t} + P_{Bmin,t}) < P_{BHKW,min}$ 2: $P_{BHKW,t} = 0$ 3: else if $P_{V,t} - (P_{PV,t} + P_{gek,t} + P_{Bmin,t}) > P_{BHKW,N}$ 4: $P_{BHKW,t} = P_{BHKW,N}$ 5: else 6: $P_{BHKW,t} = P_{V,t} - (P_{PV,t} + P_{gek,t} + P_{Bmin,t})$ 7: end 8: 9: if $P_{BHKW,t} == 0$ 10: if $P_{Bmin,t} > P_{V,t} - (P_{PV,t} + P_{gek,t} + P_{Bmin,t})$ 11: $P_{Bat,t} = P_{Bmin,t}$ 12: else 13: $P_{Bat,t} = P_{V,t} - (P_{PV,t} + P_{gek,t})$ 14: end 15: else if $P_{Bmax,t} < P_{V,t} - (P_{PV,t} + P_{gek,t} + P_{BHKW,t})$ 16: $P_{Bat,t} = P_{Bmax,t}$ 17: else 18: $P_{Bat,t} = P_{V,t} - (P_{PV,t} + P_{gek,t} + P_{BHKW,t})$ 19: end </pre>	

Abbildung 60: Pseudo-Code der regelbasierten Steuerung von Eigenbezug und Speicher

10.6 Parametrierung von Simulated Annealing

Tabelle 23: Parametrierung von Simulated Annealing

Parameter	Evaluierungsszenario			
	1a (SA1)	1b (SA1)	2 (SA2)	3 (SA3)
Temperatur	10	10	10	10
Temperatur_min	0,001	0,001	0,001	0,001
alpha	0,9999	0,9999	0,9999	0,9999
Iterations	100	100	100	100
Berechnungsdauer je Umplanung [h]	1,5	5	2	2,5

11 Studienarbeiten

Im Rahmen dieser Dissertation entstanden in den Jahren von 2017 bis 2020 unter wesentlicher wissenschaftlicher, fachlicher und inhaltlicher Anleitung des Autors die im Folgenden aufgeführten studentischen Arbeiten, in denen verschiedene Fragestellungen zur Energieflexibilität in der Produktion untersucht wurden und deren Ergebnisse in Teilen in das vorliegende Dokument eingeflossen sind. Der Autor dankt allen Studierenden für ihr Engagement bei der Unterstützung dieser wissenschaftlichen Arbeit.

- Blenk, Anna: Echtzeitfähige Produktionssteuerung basierend auf einem Multi-Agenten-System mit Reinforcement Learning (Masterarbeit, Universität Ulm).
- Christoph, Max: Integration von Eigenenergieerzeugungsanlagen und Batteriespeichern in ein Produktionssystem mittels Reinforcement Learning (Bachelorarbeit, Universität Augsburg).
- Linder, Christian: Entwicklung eines Systems zur energieorientierten Produktionssteuerung auf Basis kooperierender Agenten und Reinforcement-Learning (Masterarbeit, Hochschule München).
- Lohr, Lukas: Development of an Operation Strategy for Battery Storages at Industrial Sites using Reinforcement Learning (Bachelorarbeit, Technische Universität München).
- Priesack, Verena: Multiagentenbasierte dynamische Produktionssteuerung mittels Reinforcement Learning (Masterarbeit, Technische Universität München).
- Rudolf, Andreas: Automatisierter Stromhandel am Intraday-Markt zur Optimierung der Stromkosten von Industrieunternehmen mittels Reinforcement Learning (Bachelorarbeit, Hochschule Augsburg).

12 Literaturverzeichnis

50HERTZ 2019

50HERTZ TRANSMISSION GMBH: Internetplattform zur Vergabe von Regelleistung. <https://www.regelleistung.net/ext/static/market-information>. Zuletzt abgerufen am: 04.12.2019.

ABELE & REINHART 2011

Abele, E.; Reinhart, G.: Zukunft der Produktion. Herausforderungen, Forschungsfelder, Chancen. München: Carl Hanser Fachbuchverlag 2011. ISBN: 978-3-44642-8058.

ABELE ET AL. 2012

Abele, E.; Schrems, S.; Schraml, P.: Energieeffizienz in der Fertigungsplanung. Frühzeitige Abschätzung des Energieverbrauchs von Produktionsmaschinen in der Mittel- und Großserienfertigung. wt Werkstattstechnik online 102 2012 1/2, S. 38-42.

ACKER 2011

Acker, I. J.: Methoden zur mehrstufigen Ablaufplanung in der Halbleiterindustrie. Dissertation Universität Hohenheim. Wiesbaden: Gabler Verlag / Springer Fachmedien Wiesbaden GmbH 2011. ISBN: 978-3-8349-2921-1.

ACRICOLA ET AL. 2015

Acricola, A.; Seidl, H.; Heuke, R.: Regelleistungserbringung aus dezentralen Energieanlagen. Analyse des weiteren Handlungsbedarfs der dena-Plattform Systemdienstleistungen. Berlin: Deutsche Energie-Agentur GmbH (dena) 2015.

ADAM 1998

Adam, D.: Produktions-Management. 9. Auflage. Wiesbaden: Gabler 1998. ISBN: 978-3-409-69117-8.

ADDISU ET AL. 2017

Addisu, A.; Badis, H.; George, L.; Courbin, P.: An efficient production scheduling based on queuing theory in systems with synchronous part transfer during a demand response event, 2017 IEEE International Conference on Smart Grid Communications (SmartGridComm). Dresden, Deutschland, 23.-27.19.2017, S. 546-552. ISBN: 978-1-5386-0943-9.

AGGARWAL ET AL. 1973

Aggarwal, S. C.; Wyman, F. P.; McCarl, B. A.: An investigation of a cost-

based rule for job-shop scheduling. International Journal of Production Research 11 1973 3, S. 247-261.

AGHELINJAD ET AL. 2016

Aghelinejad, M.; Ouazene, Y.; Yalaoui, A.: Machine and production scheduling under electricity time varying prices, 2016 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). Bali, Indonesien, 03.-12.4.2016, S. 992-996. ISBN: 978-1-5090-3665-3.

AKTURK ET AL. 1999

Akturk, M.S.; Gorgulu, E.: Match-up scheduling under a machine breakdown. European Journal of Operational Research 112 1999 1, S. 81-97.

ANDRESEN ET AL. 2017

Andresen, L.; Badeda, J.; Bauer, F.; Dötsch, C.; Friedrich, A.: Energiespeicher. VDI Statusreport. Verein Deutscher Ingenieure. [1. Auflage] . Düsseldorf: VDI Verein Deutscher Ingenieure e.V. Oktober 2017. ISBN: 978-3-931-38491-3.

APPELFELLER & BUCHHOLZ 2011

Appelfeller, W.; Buchholz, W.: Supplier Relationship Management. Strategie, Organisation und IT des modernen Beschaffungsmanagements. 2. vollständig überarbeitete und erweiterte Auflage. Wiesbaden: Gabler Verlag / Springer Fachmedien Wiesbaden GmbH Wiesbaden 2011. ISBN: 978-3-8349-1809-3.

ARTIGUES ET AL. 2013

Artigues, C.; Lopez, P.; Haït, A.: The energy scheduling problem: Industrial case-study and constraint propagation techniques. International Journal of Production Economics 143 2013 1, S. 13-23.

BANKHOFER 2001

Bankhofer, Udo: Industrielles Standortmanagement. Aufgabenbereiche, Entwicklungstendenzen und problemorientierte Lösungsansätze. Wiesbaden, s.l.: Deutscher Universitätsverlag 2001. ISBN: 978-3-8244-7511-7.

BARBATI ET AL. 2012

Barbati, M.; Bruno, G.; Genovese, A.: Applications of agent-based models for optimization problems: A literature review. Expert Systems with Applications 39 2012 5, S. 6020-6028.

BATISTA ABIKARRAM ET AL. 2019

Batista Abikarram, J.; McConky, K.; Proano, R.: Energy cost minimization

for unrelated parallel machine scheduling under real time and demand charge pricing. *Journal of Cleaner Production* 208 2019, S. 232-242.

BAUERNHANSL ET AL. 2014

Bauernhansl, T.; Homple, M.; Vogel-Heuser, B. (Hrsg.): *Industrie 4.0 in Produktion, Automatisierung und Logistik. Anwendung, Technologien und Migration*. Wiesbaden: Springer Vieweg 2014. ISBN: 978-3-658-04681-1.

BDEW 2018

Bundesverband der Energie- und Wasserwirtschaft e. V. (BDEW): *BDEW-Strompreisanalyse Mai 2018. Haushalte und Industrie*. Berlin: BDEW 2018.

BDEW 2020

Bundesverband der Energie- und Wasserwirtschaft e. V. (BDEW): *BDEW-Strompreisanalyse Januar 2020. Haushalte und Industrie*. Berlin: BDEW 2020.

BEIER 2017

Beier, J.: *Simulation approach towards energy flexible manufacturing systems*. Dissertation Technische Universität Braunschweig. Cham: Springer International Publishing 2017. ISBN: 978-3-319-46638-5.

BEIER ET AL. 2015

Beier, J.; Thiede, S.; Herrmann, C.: *Increasing energy flexibility of manufacturing systems through flexible compressed air generation*. *Procedia CIRP* 37 2015, S. 18-23.

BEIER ET AL. 2016

Beier, J.; Neef, B.; Thiede, S.; Herrmann, C.: *Integrating on-site renewable electricity generation into a manufacturing system with intermittent battery Storage from Electric Vehicles*. *Procedia CIRP* 48 2016, S. 483-488.

BEIER ET AL. 2017

Beier, J.; Thiede, S.; Herrmann, C.: *Energy flexibility of manufacturing systems for variable renewable energy supply integration. Real-time control method and simulation*. *Journal of Cleaner Production* 141 2017, S. 648-661.

BELLMAN & DREYFUS 2015

Bellman, R. E.; Dreyfus, S. E.: *Applied dynamic programming*. Princeton: Princeton University Press 2015. ISBN: 978-0-691-62542-3.

BERG & BORCHERT 2014

Berg, M.; Borchert, S.: *Strategischer Energieeinkauf: Der Energieeinkauf zwischen liberalisierten Märkten und einer wechselhaften Energiepolitik in*

Deutschland. Bundesverband Materialwirtschaft, Einkauf und Logistik e.V. (BME). Frankfurt am Main 2014.

BERNER ET AL. 2019

Berner, C.; Brockman, G.; Brooke, C.; Cheung, V.; Debiak, P.; Dennison, C.; Farhi, D.; Fischer, Q.; Hashme, S.; Hesse, C.; Józefowicz, R.; Gray, S.; Olsson, C.; Pachocki, J.; Petrov, M.; Pondé de Oliveira Pinto, H.; Raiman, J.; Salimans, T.; Schlatter, J.; Schneider, J.; Sidor, S.; Sutskever, I.; Tang, J.; Wolski, F.; Zhang, S.: Dota 2 with Large Scale Deep Reinforcement Learning. ArXiv abs/1912.06680 (2019).

BERTSCH ET AL. 2017

Bertsch, J.; Fridgen, G.; Sachs, T.; Schöpf, M.; Schweter, H.; Sitzmann, A.: Ausgangsbedingungen für die Vermarktung von Nachfrageflexibilität: Status-Quo-Analyse und Metastudie. Bayreuth 2017.

BIEL 2017

Biel, K.: Multi-stage production planning with special consideration of energy supply and demand. Dissertation Technische Universität Darmstadt 2017.

BIEL & GLOCK. 2016

Biel, K.; Glock, C. H.: Systematic literature review of decision support models for energy-efficient production planning. Computers & Industrial Engineering 101 2016, S. 243-259.

BIEL ET AL. 2018

Biel, K.; Zhao, F.; Sutherland, J. W.; Glock, C. H.: Flow shop scheduling with grid-integrated onsite wind power using stochastic MILP. International Journal of Production Research 56 2018 5, S. 2076-2098.

BIERWIRTH 1993

Bierwirth, C.: Flowshop Scheduling mit parallelen Genetischen Algorithmen. Eine problemorientierte Analyse genetischer Suchstrategien. Wiesbaden: Deutscher Universitätsverlag 1993. ISBN: 978-3-8244-2051-3.

BLACKSTONE ET AL. 1982

Blackstone, J. H.; Phillips, D. T.; Hogg, G. L.: A state-of-the-art survey of dispatching rules for manufacturing job shop operations. International Journal of Production Research 20 1982 1, S. 27-45.

BLESSING & CHAKRABARTI 2009

Blessing, L. T.M.; Chakrabarti, A.: DRM, a Design Research Methodology. London: Springer London 2009. ISBN: 978-1-84882-587-1.

BMJV 2014

Bundesministerium für Justiz und für Verbraucherschutz (BMJV): Bundesgesetzblatt I/33 Art. 1: Gesetz für den Vorrang Erneuerbarer Energien (Erneuerbare-Energien-Gesetz – EEG). Köln: Bundesanzeiger 2014.

BMW i 2010

Bundesministerium für Wirtschaft und Technologie (BMWi): Energiekonzept für eine umweltschonende, zuverlässige und bezahlbare Energieversorgung. Berlin: BMW i 2010.

BMW i 2014

Bundesministerium für Wirtschaft und Technologie (BMWi): Potenzial- und Kosten-Nutzen-Analyse zu den Einsatzmöglichkeiten von Kraft-Wärme-Kopplung (Umsetzung der EU-Energieeffizienzrichtlinie) sowie Evaluierung des KWKG im Jahr 2014. Berlin: BMW i 2014.

BMW i 2018

Bundesministerium für Wirtschaft und Technologie (BMWi): Gesamtausgabe der Energiedaten - Datensammlung des BMW i. Berlin: BMW i 2018.

BNETZA & BKARTA 2015

Bundesnetzagentur (BNetzA) und Bundeskartellamt (BKartA): Monitoringbericht 2017. Elektrizitätsmarkt - Großhandel. Bonn: BNetzA 2017.

BNETZA 2011

Bundesnetzagentur (BNetzA): Monitoringbericht 2011. Bonn: BNetzA 2011.

BNETZA 2016

Bundesnetzagentur (BNetzA): Leitfaden zur Eigenversorgung. Bonn: BNetzA 2016.

BNETZA 2017

Bundesnetzagentur (BNetzA): Monitoringbericht 2017. Bonn: BNetzA 2017.

BNETZA 2019

Bundesnetzagentur (BNetzA): Monitoringbericht 2019. Bonn: BNetzA 2019.

BOLAY ET AL. 2014

Bolay, S.; Flechtner, J.; Otto, C.: Faktenpapier Eigenerzeugung von Strom. Rahmenbedingungen | Trends | Beispiele. Deutscher Industrie- und Handelskammertag (DIHK); Bundesverband der Energie-Arbeitnehmer e. V. (VEA). Berlin 2014.

BOLAY ET AL. 2017

Bolay, S.; Bullmann, T.; Hegner, M.: Faktenpapier Energiespeicher.

Rechtsrahmen | Geschäftsmodelle | Forderungen. Bundesverband Energiespeicher e. V. (BVES); Deutscher Industrie- und Handelskammertag (DIHK). Berlin 2017.

BRACKEL 2009

Brackel, T.: Adaptive Steuerung flexibler Werkstattfertigungssysteme. Nutzung moderner Informations- und Kommunikationstechnologien zur effizienten Produktionssteuerung unter Echtzeitbedingungen. Dissertation Universität Paderborn. Wiesbaden: Gabler 2009. ISBN: 978-3-8349-8066-3.

BRÄUNINGER ET AL. 2012

Bräuninger, M.; Schulze, S.; Neubach, C.; Quante, N.; ter Steege, L.: Konsequenzen der Energiewende. UniCredit Bank AG. München 2012.

BRINZER 2005

Brinzer, B.: Produktionsregelung für die variantenreiche Serienfertigung. Dissertation Universität Stuttgart. Heimsheim: Jost-Jetter 2005. ISBN: 978-3-936947-58-8.

BRUGGER ET AL. 2017

Brugger, M.; Götz, G.; Richter, C.; Reinhart, G.: Study on Energy Efficiency and Flexibility in Machinery and Production. Applied Mechanics and Materials 871 2017, S. 103-114.

BRÜNDLINGER ET AL. 2018

Bründlinger, T.; König, J.; Frank, O.; Gründig, D.; Jugel, C.; Kraft, P.; Krieger, O.; Mischinger, S.; Prein, P.; Seidl, H.; Siegemund, S.; Stolte, C.; Teichmann, M.; Willke, J.; Wolke, M.: dena-Leitstudie Integrierte Energiewende. Impulse für die Gestaltung des Energiesystems bis 2050. Ergebnisbericht und Handlungsempfehlungen. Deutsche Energie-Agentur GmbH (dena). Berlin 2018.

BRUZZONE ET AL. 2012

Bruzzone, A.A.G.; Anghinolfi, D.; Paolucci, M.; Tonelli, F.: Energy-aware scheduling for improving manufacturing process sustainability: A mathematical model for flexible flow shops. CIRP Annals - Manufacturing Technology 61 2012 1, S. 459-462.

BULLINGER ET AL. 2009

Bullinger, H.-J.; Spath, D.; Warnecke, H.-J.; Westkämper, E.: Handbuch Unternehmensorganisation. Berlin, Heidelberg: Springer Berlin Heidelberg 2009. ISBN: 978-3-540-72136-9.

BUSSE ET AL. 2017

Busse, J.; Kuprat, T.; Ambrosi, K.: Maschinenbelegungsplanung unter Berücksichtigung volatiler Energiepreise. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb 112 2017 5, S. 327-331.

BUSSE VON COLBE ET AL. 2015

Busse von Colbe, W.; Lassmann, G.; Witte, F.: Investitionstheorie und Investitionsrechnung. 4. vollständig überarbeitete Auflage . Berlin, Heidelberg: Springer 2015. ISBN: 978-3-662-46987-3.

CARIDI & SIANESI 2000

Caridi, M.; Sianesi, A.: Multi-agent systems in production planning and control: An application to the scheduling of mixed-model assembly lines. International Journal of Production Economics 68 2000 1, S. 29-42.

CARIDI & CAVALIERI 2004

Caridi, M.; Cavalieri, S.: Multi-agent systems in production planning and control: an overview. Production Planning & Control 15 2004 2, S. 106-118.

CHANDRASHEKAR & SAHIN. 2014

Chandrashekar, G.; Sahin, F.: A survey on feature selection methods. Computers & Electrical Engineering 40 2014 1, S. 16-28.

CHO & WYSK 1993

Cho, H.; Wysk, R. A.: A robust adaptive scheduler for an intelligent workstation controller. International Journal of Production Research 31 1993 4, S. 771-789.

CHOOBINEH & MOHAGHEGHI 2016

Choobineh, M.; Mohagheghi, S.: Optimal Energy Management in an Industrial plant using on-site generation and demand scheduling. IEEE Transactions on Industry Applications 52 2016 3, S. 1945-1952.

CRUZ ET AL. 2019

Cruz S., L.; Ryashentseva, D.; Lüder, A.; Vogel-Heuser, B.: Cyber-Physical Production Systems architecture based on multi-agent's design pattern comparison of selected approaches mapping four agent patterns. International Journal of Advanced Manufacturing Technology 105 2019 9, S. 4005-4034.

CUNHA ET AL.

Cunha, B.; Madureira, A. M.; Fonseca, B.; Coelho, D.: Deep Reinforcement Learning as a job shop scheduling solver: A Literature Review. In: Hybrid intelligent systems. 18th International Conference on Hybrid Intelligent

Systems (HIS 2018). Cham, Schweiz: Springer 2020, S. 350-359.
ISBN: 978-3-030-14346-6.

DABABNEH ET AL. 2018

Dababneh, F.; Li, L.; Shah, R.; Haefke, C.: Demand Response-Driven Production and Maintenance Decision-Making for Cost-Effective Manufacturing. *Journal of Manufacturing Science and Engineering* 140 2018 6, S. 61008.

DANGELMAIER 2009

Dangelmaier, W.: *Theorie der Produktionsplanung und -steuerung*. Berlin, Heidelberg: Springer Berlin Heidelberg 2009. ISBN: 978-3-642-00632-6.

DANNHAUSER 2000

Dannhauser, M.: *Kostenrechnung und Produktionsplanung und -steuerung. Ansatz für ein integriertes Fertigungscontrolling*. Gabler Edition Wissenschaft . Wiesbaden: Deutscher Universitätsverlag 2000. ISBN: 978-3-8244-7142-3.

DAW ET AL. 2005

Daw, N. D.; Niv, Y.; Dayan, P.: Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature neuroscience* 8 2005 12, S. 1704-1711.

DAYAN & NIV 2008

Dayan, P.; Niv, Y.: Reinforcement learning: the good, the bad and the ugly. *Current opinion in neurobiology* 18 2008 2, S. 185-196.

DESTATIS 2016

Statistisches Bundesamt (destatis): *Datensatz: Erhebung über die Energieverwendung von 2005, 2010 und 2015*. Wiesbaden: destatis 2016.

DEWEY 2014

Dewey, D.: Reinforcement Learning and the reward engineering principle. 2014 Association for the Advancement of Artificial Intelligence (AAAI) Spring Symposium Series 2014.

DIETTERICH ET AL. 2018

Dietterich, T. G.; Trimponias, G.; Chen, Z.: Discovering and removing Exogenous state variables and rewards for Reinforcement Learning. *The thirty-fifth International Conference on Machine Learning (ICML) 2018*, S. 1261-1269.

DIHK 2017

Deutscher Industrie- und Handelskammertag (DIHK): Faktenpapier Strompreise in Deutschland 2017. Bestandteile, Entwicklungen, Strategien. Berlin: DIHK 2017.

DIHK 2019

Deutscher Industrie- und Handelskammertag (DIHK): Strombeschaffung und Stromhandel. Hintergründe, Herausforderungen, Hinweise. Berlin: DIHK 2019.

DOMSCHKE ET AL. 2015

Domschke, W.; Drexl, A.; Klein, R.; Scholl, A.: Einführung in Operations Research. 9. überarbeitete und verbesserte Auflage Berlin, Heidelberg: Springer Gabler 2015. ISBN: 978-3-662-48215-5.

DÖRMER & GÜNTHER 2013

Dörmer, J.; Günther, H.-O.: Produktionsprogrammplanung bei variantenreicher Fließproduktion. Untersucht am Beispiel der Automobilentmontage. Dissertation Technische Universität Berlin. Wiesbaden: Springer Gabler 2013. ISBN: 978-3-658-02092-7.

DOYA ET AL. 2002

Doya, K.; Samejima, K.; Katagiri, K.-i.; Kawato, M.: Multiple model-based reinforcement learning. Neural computation 14 2002 6, S. 1347-1369.

DVOŘÁK & HAVEL 2012

Dvořák, M.; Havel, P.: Combined heat and power production planning under liberalized market conditions. Applied Thermal Engineering 43 2012, S. 163-173.

EDEL & SCHEMM 2018

Edel, P.; Schemm, R.: Einfluss von Preisspitzen im Regelenergiemarkt auf die Risikovorsorge von Stromvertrieben. "et" - Energiewirtschaftliche Transferfragen 2018 05/2018, S. 63-65.

EICHLER 1993

Eichler, H.: Ökosystem Erde. Der Störfall Mensch; eine Schadens- und Vernetzungsanalyse. Mannheim: BI-Taschenbuchverl. 1993. ISBN: 978-3-4111-0291-4.

EILON & CHOWDHURY 1976

Eilon, S.; Chowdhury, I. O.: Due dates in job shop scheduling. International Journal of Production Research 14 1976 2, S. 223-237.

ELLER 2015

Eller, D.: Integration erneuerbarer Energien mit Power-to-Heat in Deutschland. Potentiale zur Nutzung von Stromüberschüssen in Fernwärmenetzen. Wiesbaden: Springer Vieweg 2015. ISBN: 978-3-6581-0560-0.

EPEX SPOT 2018

EPEX SPOT: Intraday market with delivery on the German TSO zone. EPEX SPOT 2018. <https://www.epexspot.com/en/product-info/intradaycontinuous/germany>. Zuletzt abgerufen am: 11.12.2018.

EPEX SPOT 2019

EPEX SPOT: EPEX SPOT Price List. EPEX Spot. <https://www.epexspot.com/en/fees/pricelist>. Zuletzt abgerufen am: 10.10.2019.

ERNST & YOUNG 2016

Ernst & Young GmbH Wirtschaftsprüfungsgesellschaft (Ernst & Young): Wirtschaft unter Strom. Wie Unternehmen sich unabhängiger und kostengünstiger selbst mit Energie versorgen. Ernst & Young GmbH 2016.

EUROSTAT 2018

EUROSTAT: Electricity prices for non-household consumers - bi-annual data 2018. http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nrg_pc_205&lang=en. Zuletzt abgerufen am: 2.08.2018.

FANG ET AL. 2011

Fang, K.; Uhan, N.; Zhao, F.; Sutherland, J. W.: A new approach to scheduling in manufacturing for power consumption and carbon footprint reduction. *Journal of Manufacturing Systems* 30 2011 4, S. 234-240.

FATTAHI & FALLAHI 2010

Fattahi, P.; Fallahi, A.: Dynamic scheduling in flexible job shop systems by considering simultaneously efficiency and stability. *CIRP Journal of Manufacturing Science and Technology* 2 2010 2, S. 114-123.

FAZLI KHALAF & WANG 2018

Fazli Khalaf, A.; Wang, Y.: Energy-cost-aware flow shop scheduling considering intermittent renewables, energy storage, and real-time electricity pricing. *International Journal of Energy Research* 42 2018 12, S. 3928-3942.

FERNANDEZ ET AL. 2013

Fernandez, M.; Li, L.; Sun, Z.: "Just-for-Peak" buffer inventory for peak electricity demand reduction of manufacturing systems. *International Journal of Production Economics* 146 2013 1, S. 178-184.

FISCHER 2007

Fischer, A.: Modellbasierte Wirkbeschreibung von Prioritätsregeln. Zugl.: Hannover, Univ., Diss., 2007. Garbsen: PZH Produktionstechn. Zentrum 2007. ISBN: 9783939026501.

FLEITER 2013

Fleiter, Tobias: Energieverbrauch und CO₂-Emissionen industrieller Prozesstechnologien. Einsparpotenziale, Hemmnisse und Instrumente. Fraunhofer-Institut für Systemtechnik und Innovationsforschung. Stuttgart: Fraunhofer-Verl. 2013. ISBN: 978-3-839-60515-8. (ISI-Schriftenreihe "Innovationspotenziale").

FLOREA ET AL.

Florea, A.; Luttamus, J.; Postelnicu, C.; Lastra, J. L.: Agent-Based Control of Operational Conditions for Smart Factories: The Peak Load Management Scenario. In: Proceedings of the 6th International Conference on Industrial Applications of Holonic and Multi-Agent Systems (HoloMAS 2013). Prag, Tschechische Republik, 26-28.08.2013. Berlin, Heidelberg: Springer Berlin Heidelberg 2013, S. 83-94.

FRANÇOIS-LAVET ET AL. 2018

François-Lavet, V.; Henderson, P.; Islam, R.; Bellemare, M. G.; Pineau, J.: An Introduction to Deep Reinforcement Learning. Foundations and Trends in Machine Learning 11 2018 3-4, S. 219-354.

FRENCH 1986

French, S.: Sequencing and scheduling. An introduction to the mathematics of the job-shop. Chicester: Ellis Horwood 1986. ISBN: 0470272295.

GABEL & RIEDMILLER 2007

Gabel, T.; Riedmiller, M.: On a Successful Application of Multi-Agent Reinforcement Learning to Operations Research Benchmarks. Honolulu, Hawaii, USA, 2007.

GABEL 2009

Gabel, T.: Multi-Agent Reinforcement Learning Approaches for Distributed Job-Shop Scheduling Problems. Dissertation Universität Osnabrück 2009.

GAO ET AL. 2015

Gao, B.; Liu, X.; Zhang, W.; Tang, Y.: Autonomous Household Energy Management Based on a Double Cooperative Game Approach in the Smart Grid. Energies 8 2015 7, S. 7326-7343.

GELLINGS & CHAMBERLIN 1993

Gellings, C. W.; Chamberlin, J. H.: Demand-side management. Concepts and methods. 2. edition. Lilburn, Ga.: Fairmont Press 1993. ISBN: 978-0132049757.

GESINO 2011

Gesino, A. J.: Power reserve provision with wind farms. Grid integrations of wind power. Kassel: Kassel Univ. Press 2011. ISBN: 978-3-86219-022-5.

GLOBAL FOOTPRINTING NETWORK 2020

Global Footprinting Network: Earth Overshoot Day 2020.
<https://www.overshootday.org/>. Zuletzt abgerufen am: 26.04.2020.

GONG ET AL. 2015

Gong, X.; Pessemier, T. de; Joseph, W.; Martens, L.: An Energy-Cost-Aware Scheduling Methodology for Sustainable Manufacturing. *Procedia CIRP* 29 2015, S. 185-190.

GONG ET AL. 2017A

Gong, X.; van der Wee, M.; Pessemier, T.; Verbrugge, S.; Colle, D.; Martens, L.; Joseph, W.: Integrating labor awareness to energy-efficient production scheduling under real-time electricity pricing. An empirical study. *Journal of Cleaner Production* 168 2017, S. 239-253.

GONG ET AL. 2017B

Gong, X.; van der Wee, M.; Pessemier, T. de; Verbrugge, S.; Colle, D.; Martens, L.; Joseph, W.: Energy- and Labor-aware Production Scheduling for Sustainable Manufacturing: A Case Study on Plastic Bottle Manufacturing. *Procedia CIRP* 61 2017, S. 387-392.

GONG ET AL. 2018

Gong, X.; Liu, Y.; Lohse, N.; Pessemier, T. de; Martens, L.; Joseph, W.: Energy and Labor Aware Production Scheduling for Industrial Demand Response Using Adaptive Multi-objective Memetic Algorithm. *IEEE Transactions on Industrial Informatics* 2018, S. 1.

GRABL 2015

Graßl, M.: Bewertung der Energieflexibilität in der Produktion. Zugl.: München, Techn. Univ., Diss., 2015. München: Utz 2015. ISBN: 978-3-8316-4476-6.

GRONDMAN ET AL. 2012

Grondman, I.; Busoniu, L.; Lopes, G. A.; Babuska, R.: A Survey of Actor-Critic Reinforcement Learning: Standard and Natural Policy Gradients.

IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 42 2012 6, S. 1291-1307.

GROSCH ET AL. 2019

Grosch, B.; Weitzel, T.; Panten, N.; Abele, E.: A metaheuristic for energy adaptive production scheduling with multiple energy carriers and its implementation in a real production system. *Procedia CIRP* 80 2019, S. 203-208.

GROBE BÖCKMANN 2014

Große Böckmann, M.: Senkung der Produktionskosten durch Gestaltung eines Energiereglerkreis-Konzeptes. Dissertation Technische Hochschule Aachen. 1. Aufl. . Aachen: Apprimus-Verl. 2014. ISBN: 978-3-86359-246-2.

GÜNTHER 2015

Günther, M.: Energieeffizienz durch erneuerbare Energien. Möglichkeiten, Potenziale, Systeme. Wiesbaden: Springer Vieweg 2015. ISBN: 978-3-658-06752-6.

GUO ET AL. 2006

Guo, Z. X.; Wong, W. K.; Leung, S.Y.S.; Fan, J. T.; Chan, S. F.: Mathematical model and genetic optimization for the job shop scheduling problem in a mixed- and multi-product assembly environment. A case study based on the apparel industry. *Computers & Industrial Engineering* 50 2006 3, S. 202-219.

GUPTA ET AL. 1989

Gupta, Y.; Gupta, M.; Bector, C.: A review of scheduling rules in flexible manufacturing systems. *International Journal of Computer Integrated Manufacturing* 2 1989 6, S. 356-377.

HAARNOJA ET AL. 2018

Haarnoja, T.; Zhou, A.; Abbeel, P.; Levine, S.: Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. 2018.

HACKSTEIN 1989

Hackstein, R.: Produktionsplanung und -steuerung (PPS). Ein Handbuch für die Betriebspraxis. 2., überarb. Aufl. . Düsseldorf: VDI-Verl. 1989. ISBN: 3-18-400924-6.

HANDELSBLATT 2017

Handelsblatt: Anzahl der Stunden mit negativen Strompreisen in Deutschland in den Jahren 2008 bis 2017. Düsseldorf: Handelsblatt 2017.

HANSMANN

Hansmann, K.-W.: Industrielles Management. 8. völlig überarbeitete und erweiterte Auflage. München: Oldenbourg Wissenschafts Verlage GmbH 2006. ISBN: 9783486580587.

HAOUASSI ET AL. 2016

Haouassi, M.; Desdouits, C.; Giroudeau, R.; Le Pape, C.: Production scheduling with a piecewise-linear energy cost function, 2016 IEEE Symposium Series on Computational Intelligence (SSCI). Athens, Greece, 12/6/2016 - 12/9/2016, S. 1-8. ISBN: 978-1-5090-4240-1.

HENDERSON ET AL. 2018

Henderson, P.; Islam, R.; Bachman, P.; Pineau, J.; Precup, D.; Meger, D.: Deep Reinforcement Learning that Matters. Proceedings of the Thirtieth-Second AAAI Conference on Artificial Intelligence 32 2018 1.

HENSELER 1998

Henseler, H.: Aktive Ablaufplanung mit Multi-Agenten. Zugl.: Oldenburg, Univ., Diss., 1997. Sankt Augustin: Infix 1998. ISBN: 3896011804.

HERRMANN 2006

Herrmann, J. W.: Handbook of Production Scheduling. Boston, MA: Springer Science+Business Media Inc 2006. ISBN: 978-0-387-33117-1.

HESSEL 2008

Hessel, V.: Energiemanagement. Maßnahmen zur Verbrauchs- und Kostenreduzierung, Förderprogramme, Vorschriften. Erlangen: Publicis Publ. 2008. ISBN: 978-3-89578-327-2.

HOITSCH 1993

Hoitsch, H.-J.: Produktionswirtschaft. Grundlagen einer industriellen Betriebswirtschaftslehre. 2. völlig überarbeitete und erweiterte Auflage. München: Vahlen 1993. ISBN: 978-3-8006-1619-0.

HOLDACK 2015

HOLDACK: Hinweise und Empfehlungen zur Direktvermarktung aus technischer Sicht. Bedarfsgerechte Stromerzeugung mit Biogas-BHKW. Dorfmark: Dreyer & Bosse Kraftwerke GmbH 2015.

https://www.leuphana.de/fileadmin/user_upload/Forschungseinrichtungen/professuren/finanzierung-finanzwirtschaft/files/24_03_2015_Direktvermarktung_D_B_BHKW_Dorfmark_Uni_Lueneburg.pdf. Zuletzt abgerufen am: 21.11.2019.

HORSCH 2015

Horsch, J.: Kostenrechnung. Klassische und neue Methoden in der Unternehmenspraxis. 2. Auflage. Wiesbaden: Springer Gabler 2015. ISBN: 978-3-6580-7311-4.

HU & HU 2018

Hu, Z.; Hu, G.: A multi-stage stochastic programming for lot-sizing and scheduling under demand uncertainty. *Computers & Industrial Engineering* 119 2018, S. 157-166.

HUNEKE & LENCK. 2016

Huneke, F.; Lenck, T.: Kurzanalyse zur Stromerzeugung bei Netzbedingter Abregelung Erneuerbarer Energien. Berlin: Energy Brainpool 2016.

IEA 2017

Internationale Energieagentur (IEA): CO2 Emissions From Fuel Combustion 2017. Paris: IEA 2017.

ISLAM ET AL. 2018

Islam, M. M.; Zhong, X.; Xiong, H.; Sun, Z.: Optimal scheduling of manufacturing and onsite generation systems in over-generation mitigation oriented electricity demand response program. *Computers & Industrial Engineering* 115 2018, S. 381-388.

JADERBERG ET AL. 2019

Jaderberg, M.; Czarnecki, W. M.; Dunning, I.; Marris, L.; Lever, G.; Castañeda, A. G.; Beattie, C.; Rabinowitz, N. C.; Morcos, Ari S.; Ruderman, A.; Sonnerat, N.; Green, T.; Deason, L.; Leibo, J. Z.; Silver, D.; Hassabis, D.; Kavukcuoglu, K.; Graepel, T.: Human-level performance in 3D multiplayer games with population-based reinforcement learning. *Science* 364 2019 6443, S. 859-865.

JARASS ET AL. 2009

Jarass, L.; Obermair, G. M.; Voigt, W.: Windenergie. Zuverlässige Integration in die Energieversorgung. 2. Auflage. Springer-Verlag 2009. ISBN: 978-3-540-85253-7.

JIMÉNEZ 2012

Jiménez, Y. M.: A Generic Multi-Agent Reinforcement Learning Approach for Scheduling Problems. PhD Thesis. Vrije Universiteit Brussel 2012.

KAEBLING ET AL. 1996

Kaelbling, L. P.; Littman, M. L.; Moore, A. W.: Reinforcement Learning: A Survey. *Journal of artificial intelligence research* 1996, S. 237-285.

KALTSCHMITT ET AL. 2013

Kaltschmitt, M.; Streicher, W.; Wiese, A.: Erneuerbare Energien: Systemtechnik, Wirtschaftlichkeit, Umweltaspekte: Springer Berlin Heidelberg 2013. ISBN: 978-3-540-28205-1.

KEIL ET AL. 2016

Keil, P.; Schuster, S. F.; Wilhelm, J.; Travi, J.; Hauser, A.; Karl, R. C.; Jossen, A.: Calendar Aging of Lithium-Ion Batteries. Journal of The Electrochemical Society 163 2016 9, S. A1872-A1880.

KELLER 2018

Keller, F.: Methodik zur energiebezugsorientierten Auftragsplanung. Dissertation Technische Universität München. München: Utz 2018.

KELLER & REINHART 2016

Keller, F.; Reinhart, G.: Energy Supply Orientation in Production Planning Systems. Procedia CIRP 40 2016, S. 244-249.

KHODKE & BHONGADE 2013

Khodke, P. M.; Bhongade, A. S.: Real-time scheduling in manufacturing system with machining and assembly operations: a state of art. International Journal of Production Research 51 2013 16, S. 4966-4978.

KIM ET AL. 2017

Kim, S.; Meng, C.; Son, Y.-J.: Simulation-based machine shop operations scheduling system for energy cost reduction. Simulation Modelling Practice and Theory 77 2017, S. 68-83.

KLETTI & SCHUMACHER 2014

Kletti, J.; Schumacher, J.: Die perfekte Produktion. Manufacturing Excellence durch Short Interval Technology (SIT). 2. Auflage. Berlin: Springer Vieweg 2014. ISBN: 978-3-662-45441-1.

KÖHLER ET AL. 2018

Köhler, A.; Baron, Y.; Bulach, W.; Vogel, M.; Behrendt, S.; Degel, M.; Krauß, N.; Buchert, M.: Studie: Ökologische und ökonomische Bewertung des Ressourcenaufwands - Stationäre Energiespeichersysteme in der industriellen Produktion. Berlin: VDI ZRE 2018.

KOLLER ET AL. 2013

Koller, M.; Borsche, T.; Ulbig, A.; Andersson, G.: Defining a degradation cost function for optimal control of a battery energy storage system, 2013 IEEE Grenoble PowerTech. Grenoble, France, 16.-20.06.2013, S. 1-6. ISBN: 978-1-4673-5669-5.

KONDZIELLA ET AL. 2013

Kondziella, H.; Brod, K.; Bruckner, T.; Olbert, S.; Mes, F.: Stromspeicher für die „Energiewende“ – eine aktorsbasierte Analyse der zusätzlichen Speicherkosten. Zeitschrift für Energiewirtschaft 37 2013 4, S. 249-260.

KONSTANTIN 2009

Konstantin, P.: Praxisbuch Energiewirtschaft. Energieumwandlung, -transport und -beschaffung im liberalisierten Markt. 2. bearbeitete und aktualisierte Auflage. Berlin, Heidelberg: Springer Berlin Heidelberg 2009. ISBN: 978-3-540-78592-7.

KÖRNER ET AL. 2019

Körner, M.-F.; Bauer, D.; Keller, R.; Rösch, M.; Schlereth, A.; Simon, P.; Bauernhansl, T.; Fridgen, G.; Reinhart, G.: Extending the Automation Pyramid for Industrial Demand Response. Procedia CIRP 81 2019, S. 998-1003.

KOST ET AL. 2018

Kost, C.; Schammugam, S.; Jülch, V.; Ngyen, H.-T.; Schlegl, T.: Stromgestehungskosten erneuerbarer Energien. Freiburg: Fraunhofer ISE 2018.

KRISHNAN ET AL. 2012

Krishnan, M.; Chinnusamy, T. R.; Karthikeyan, T.: Performance Study of Flexible Manufacturing System Scheduling Using Dispatching Rules in Dynamic Environment. Procedia Engineering 38 2012, S. 2793-2798.

KRISHNAN ET AL. 2019

Krishnan, S.; Borojerdian, B.; Fu, W.; Faust, A.; Reddi, V. J.: Air Learning: An AI Research Platform for Algorithm-Hardware Benchmarking of Autonomous Aerial Robots. 2019.

KRUSCHWITZ 2005

Kruschwitz, L.: Investitionsrechnung. 10., überarbeitete und erweiterte Auflage. München: Oldenbourg 2005. ISBN: 978-3-1103-7627-2.

KUHNLE ET AL. 2019

Kuhnle, A.; Schäfer, L.; Stricker, N.; Lanza, G.: Design, Implementation and Evaluation of Reinforcement Learning for an Adaptive Order Dispatching in Job Shop Manufacturing Systems. Procedia CIRP 81 2019, S. 234-239.

KUMAR & RAJOTIA 2006

Kumar, M.; Rajotia, S.: Cost based scheduling in a job shop environment. The International Journal of Advanced Manufacturing Technology 30 2006 11-12, S. 1144-1153.

KURBEL 2013

Kurbel, K.: Enterprise resource planning and supply chain management. Functions, business processes and software for manufacturing companies. Dordrecht: Springer 2013. ISBN: 978-3-642-31573-2.

KÜSTER ET AL. 2013

Küster, T.; Lützenberger, M.; Freund, D.; Albayrak, S.: Distributed Evolutionary Optimisation for Electricity Price Responsive Manufacturing using Multi-Agent System Technology. International Journal on Advances in Intelligent Systems 2013 vol 6, S. 27-40.

LANGE ET AL. 2012

Lange, S.; Riedmiller, M.; Voigtlander, A.: Autonomous reinforcement learning on raw visual input data in a real world application, 2012 International Joint Conference on Neural Networks (IJCNN 2012 - Brisbane). Brisbane, Australia, 6/10/2012 - 6/15/2012, S. 1-8. ISBN: 978-1-4673-1490-9.

LANGER ET AL. 2014

Langer, T.; Schlegel, A.; Stoldt, J.; Putz, M.: A Model-based Approach to Energy-saving Manufacturing Control Strategies. Procedia CIRP 15 2014, S. 123-128.

LARESGOITI ET AL. 2015

Laresgoiti, I.; Käbitz, S.; Ecker, M.; Sauer, D. U.: Modeling mechanical degradation in lithium ion batteries during cycling: Solid electrolyte inter-phase fracture. Journal of Power Sources 300 2015, S. 112-122.

LE & PANG 2013

Le, C. V.; Pang, C. K.: Fast reactive scheduling to minimize tardiness penalty and energy cost under power consumption uncertainties. Computers & Industrial Engineering 66 2013 2, S. 406-417.

LECHNER ET AL. 2018

Lechner, R.; O'Connell, N.; Meierhofer, T.; Brautsch, M.: Entwicklung, Umsetzung und Bewertung optimierter Monitoring-, Betriebs- und Regelstrategien für Blockheizkraftwerke. Abschlussbericht. Bearbeitungszeitraum: 11/2014 12/2016 Aktenzeichen: SWD-10.08.18.7-14. Stuttgart: Fraunhofer IRB Verlag 2018. ISBN: 978-3-7388-0083-8.

LEE ET AL. 2017

Lee, S.; Do Chung, B.; Jeon, H. W.; Chang, J.: A dynamic control approach for energy-efficient production scheduling on a single machine under time-

varying electricity pricing. *Journal of Cleaner Production* 165 2017, S. 552-563.

LEITÃO 2009

Leitão, P.: Agent-based distributed manufacturing control. A state-of-the-art survey. *Engineering Applications of Artificial Intelligence* 22 2009 7, S. 979-991.

LI ET AL. 2018A

Li, S.; Liu, F.; Zhou, X.: Multi-objective energy-saving scheduling for a permutation flow line. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* 232 2018 5, S. 879-888.

LI ET AL. 2018B

Li, X.; Xing, K.; Zhou, M.; Wang, X.; Wu, Y.: Modified Dynamic Programming Algorithm for Optimization of Total Energy Consumption in Flexible Manufacturing Systems. *IEEE Transactions on Automation Science and Engineering* 2018, S. 1-15.

LIANG ET AL. 2017

Liang, E.; Liaw, R.; Moritz, P.; Nishihara, R.; Fox, R.; Goldberg, K.; Gonzalez, J. E.; Jordan, M. I.; Stoica, Ion: *RLlib: Abstractions for Distributed Reinforcement Learning*. 2017.

LIAW ET AL. 2003

Liaw, B. Y.; Roth, E.P.; Jungst, R. G.; Nagasubramanian, G.; Case, H. L.; Doughty, D. H.: Correlation of Arrhenius behaviors in power and capacity fades with cell impedance and heat generation in cylindrical lithium-ion cells. *Journal of Power Sources* 119-121 2003, S. 874-886.

LILLICRAP ET AL. 2015

Lillicrap, T. P.; Hunt, J. J.; Pritzel, A.; Heess, N.; Erez, T.; Tassa, Y.; Silver, D.; Wierstra, D.: Continuous control with deep reinforcement learning. 2015.

LIN ET AL. 2012

Lin, L.; Zeyi, S.; Zhijun, T.: Real Time Electricity Demand Response for Sustainable Manufacturing Systems: Challenges and a Case Study. 20 - 24 Aug. 2012, Seoul, Korea. Institute of Electrical and Electronics Engineers; IEEE International Conference on Automation Science and Engineering; CASE. Piscataway: IEEE 2012. ISBN: 978-1-467-30430-6.

LIU ET AL. 2014

Liu, Y.; Dong, H.; Lohse, N.; Petrovic, S.; Gindy, N.: An investigation into

minimising total energy consumption and total weighted tardiness in job shops. *Journal of Cleaner Production* 65 2014, S. 87-96.

LIU ET AL. 2015

Liu, D.-R.; Li, H.-L.; Wang, D.: Feature selection and feature learning for high-dimensional batch reinforcement learning: A survey. *International Journal of Automation and Computing* 12 2015 3, S. 229-242.

LIU ET AL. 2020

Liu, M.; Yang, X.; Chu, F.; Zhang, J.; Chu, C.: Energy-oriented bi-objective optimization for the tempered glass scheduling. *Omega* 90 2020, S. 101995.

LÖDDING 2008

Lödning, H.: *Verfahren der Fertigungssteuerung*. Berlin, Heidelberg: Springer Berlin Heidelberg 2008. ISBN: 978-3-540-76859-3.

LÖDDING 2013

Lödning, H.: *Handbook of manufacturing control. Fundamentals, description, configuration*. Heidelberg: Springer 2013. ISBN: 978-3-642-24457-5.

LÖDDING ET AL. 2012

Lödning, H.; Nyhuis, P.; Schmidt, M.; Kuyumcu, A.: Modelling lateness and schedule reliability. How companies can produce on time. *Production Planning & Control* 25 2012 1, S. 59-72.

LOON ET AL. 2019

Loon, K. W.; Graesser, L.; Cvitkovic, M.: *SLM Lab: A Comprehensive Benchmark and Modular Software Framework for Reproducible Deep Reinforcement Learning*. 2019.

LU ET AL. 2018

Lu, R.; Hong, S. H.; Zhang, X.: A Dynamic pricing demand response algorithm for smart grid. Reinforcement learning approach. *Applied Energy* 220 2018, S. 220-230.

LÜERS ET AL. 2015

Lüers, S.; Wallasch, A.-K.; Rehfeldt, K.: *Kostensituation der Windenergie an Land in Deutschland*. Varel: Deutsche WindGuard GmbH 2015.

LUDWIG ET AL. 2018

Ludwig, M.; Korotkiewicz, K.; Dahlmann, B.; Zdrallek, M.; Derksen, C.: Agent-based Grid Automation in Distribution Grids: Experiences Under Real Field Conditions. In: Aim (Hrsg.): *Proceedings of CIRED Workshop 2018* 2018.

LUO ET AL. 2013

Luo, H.; Du, B.; Huang, G. Q.; Chen, H.; Li, X.: Hybrid flow shop scheduling considering machine electricity consumption cost. *International Journal of Production Economics* 146 2013 2, S. 423-439.

MACKENSEN ET AL. 2017

Mackensen, R.; Saint-Drenan, Y.-M.; Jost, D.; Fritz, R.; Asanalieva, N.; Widdel, M.; Hahler, M.: *Regelenergie durch Wind- und Photovoltaikparks. Abschlussbericht. Fraunhofer IWES. Kassel 2017.*

MALKIS & BROY 2018

Malkis, A.; Broy, M.: *Logische und Methodische Grundlagen der Programm- und Systementwicklung: Datenstrukturen, funktionale, sequenzielle und objektorientierte Programmierung. Springer Fachmedien Wiesbaden, 2019.*

MANN ET AL. 2017

Mann, M. E.; Rahmstorf, S.; Kornhuber, K.; Steinman, B. A.; Miller, S. K.; Coumou, D.: Influence of Anthropogenic Climate Change on Planetary Wave Resonance and Extreme Weather Events. *Scientific Reports* 7 2017, S. 45242-45253.

MARTINEZ-LASERNA ET AL. 2018

Martinez-Laserna, E.; Herrera, V.; Gandiaga, I.; Milo, A.; Sarasketa-Zabala, E.; Gaztañaga, H.: Li-Ion Battery Lifetime Model's Influence on the Economic Assessment of a Hybrid Electric Bus's Operation. *World Electric Vehicle Journal* 9 2018 2, S. 28.

MCKINSEY&COMPANY 2017

McKinsey&Company: *Electrifying insights: How automakers can drive electrified vehicle sales and profitability 2017.*

MENDIL ET AL. 2018

Mendil, M.; Domenico, A.; Heiries, V.; Caire, R.; Hadjsaid, N.: Battery-Aware Optimization of Green Small Cells: Sizing and Energy Management. *IEEE Transactions on Green Communications and Networking* 2 2018 3, S. 635-651.

MERTENS & WIENDAHL 1992

Mertens, P.; Wiendahl, H.-P.: *PPS im Wandel. Kundenorientierung und Wirtschaftlichkeit durch innovative PPS-Lösungen. Gesellschaft für Management und Technologie; Friedrich-Alexander-Universität Erlangen-*

Nürnberg; Technische Universität Hannover; Technische Universität München; Fachtagung PPS im Wandel. München: Gesellschaft für Management und Technologie-Verl.-KG 1992. ISBN: 389-4150394.

MEZIANE & TAGHEZOUT 2018

Meziane, M. E.; Taghezout, N.: Predictive Reactive Approach for Energy-Aware Scheduling and Control of Flexible Manufacturing Processes. International Journal of Information Systems and Supply Chain Management 11 2018 4, S. 43-62.

MILBERG 1992

Milberg, J.: Werkzeugmaschinen - Grundlagen. Zerspantechnik, Dynamik, Baugruppen und Steuerungen. Berlin, Heidelberg, s.l.: Springer Berlin Heidelberg 1992. ISBN: 978-3-662-10915-1.

MILLNER 2010

Millner, A.: Modeling Lithium Ion battery degradation in electric vehicles, 2010 IEEE Conference on Innovative Technologies for an Efficient and Reliable Electricity Supply. Waltham, MA, USA, 27.-92.09.2010, S. 349-356. ISBN: 978-1-4244-6076-2.

MINNE ET AL. 2019

Minne, L.; Lisheng, W.; Haitham, B. A.; Jun, W.: Multi-View Reinforcement Learning. 33rd Conference on Neural Information Processing System 2019.

MNIH ET AL. 2013

Mnih, V.; Kavukcuoglu, K.; Silver, D.; Graves, A.; Antonoglou, I.; Wierstra, D.; Riedmiller, M.: Playing Atari with Deep Reinforcement Learning. 2013.

MOSLER 2017

Mosler, A.: Integrierte Unternehmensplanung. Anforderungen, Lösungen und Echtzeitsimulation im Rahmen von Industrie 4.0. Wiesbaden: Springer Gabler 2017. ISBN: 978-3-658-08752-4.

MÜLLER ET AL. 2009

Müller, E.; Engelmann, J.; Löffler, T.; Strauch, J.: Energieeffiziente Fabriken planen und betreiben. Berlin, Heidelberg: Springer Berlin Heidelberg 2009. ISBN: 978-3-642-31945-7.

MUTH & THOMPSON 1963

Muth, J.; Thompson, G.: Industrial Scheduling. Dordrecht, The Netherlands: Kluwer Academic Publishers 1963.

NASERI & KUZGUNKAYA 2010

Nasari, E.; Kuzgunkaya, O.: Cost-based rescheduling in a flexible manufacturing system using filtered-beam search. IIE Annual Conference and Expo 2010.

NEBEL-WENNER ET AL. 2019

Nebel-Wenner, M.; Reinhold, C.; Wille, F.; Nieße, A.; Sonnenschein, M.: Distributed multi-objective scheduling of power consumption for smart buildings. *Energy Inform* 2, 28 (2019).

NEUGEBAUER ET AL. 2012

Neugebauer, R.; Putz, M.; Schlegel, A.; Langer, T.; Franz, E.; Lorenz, S.: Energy-Sensitive Production Control in Mixed Model Manufacturing Processes. Dornfeld D., Linke B. (eds) *Leveraging Technology for a Sustainable World 2012*, S. 399-404.

NGUYEN & NGUYEN 2016

Nguyen, M. Y.; Nguyen, D. M.: A New Framework of Demand Response for Household Customers Based on Advanced Metering Infrastructure Under Smart Grids. *Electric Power Components and Systems* 44 2016 2, S. 165-171.

NGYEN ET AL. 2018

Nguyen, D. T.; Kumar, A.; Lau, H. C.: Credit assignment for collective multi-agent RL with global rewards. *NIPS'18: Proceedings of the 32nd International Conference on Neural Information Processing System 2018*, S. 8113-8124.

NIEHUES 2017

Niehues, M. R.: Adaptive Produktionssteuerung für Werkstattfertigungssysteme durch fertigungsbegleitende Reihenfolgebildung. Dissertation Technische Universität München. München: Utz 2017.

NOURI ET AL. 2018

Nouri, M.; Bekrar, A.; Trentesaux, D.: Towards Energy Efficient Scheduling and Rescheduling for Dynamic Flexible Job Shop Problem. *IFAC-PapersOnLine* 51 2018 11, S. 1275-1280.

NYHUIS 2008

Nyhuis, P.: *Beiträge zu einer Theorie der Logistik*. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg 2008. ISBN: 978-3-540-75642-2.

NYHUIS & WIENDAHL 2009

Nyhuis, P.; Wiendahl, H.-P.: Fundamentals of production logistics. Theory, tools and applications. Berlin: Springer 2009. ISBN: 978-3-540-34211-3.

NYHUIS & MÜNZBERG 2010

Nyhuis, P.; Münzberg, B.: Konfiguration der Fertigungssteuerung. Berücksichtigung unternehmensspezifischer Anforderungen. wt Werkstattstechnik online 100 2010 4, S. 285-290.

OPENAI 2018

OPENAI: A Taxonomy of RL Algorithms 2018. https://spinningup.openai.com/en/latest/spinningup/rl_intro2.html#citations-below. Zuletzt abgerufen am: 10.11.2019.

OROOJLOOYJADID & HAJINEZHAD 2019

OroojlooyJadid, A.; Hajinezhad, D.: A Review of Cooperative Multi-Agent Deep Reinforcement Learning. ArXiv abs/1908.03963 2019.

ORTEGA-VAZQUEZ 2014

Ortega-Vazquez, M. A.: Optimal scheduling of electric vehicle charging and vehicle-to-grid services at household level including battery degradation and price uncertainty. IET Generation, Transmission & Distribution 8 2014 6, S. 1007-1016.

OUELHADJ & PETROVIC 2009

Ouelhadj, D.; Petrovic, S.: A survey of dynamic scheduling in manufacturing systems. Journal of Scheduling 12 2009 4, S. 417-431.

OU-YANG & LIOU 1998

Ou-Yang, C.; Liou, M. E.: A decision analysis method for a tardy WIP from the production cost point of view. International Journal of Systems Science 29 1998 4, S. 373-388.

PANWALKAR & ISKANDER 1977

Panwalkar, S. S.; Iskander, W.: A Survey of Scheduling Rules. Operations Research 25 1977 1, S. 45-61.

PARK ET AL. 2019

Park, Y. J.; Cho, Y. S.; Kim, S. B.: Multi-agent reinforcement learning with approximate model learning for competitive games. PloS one 14 2019 9.

PECHMANN ET AL. 2012

Pechmann, A.; Schöler, I.; Hackmann, R.: Energy Efficient and Intelligent Production Scheduling – Evaluation of a New Production Planning and

Scheduling Software. Leveraging Technology for a Sustainable World
2012.

PILGRAM ET AL. 2014

Pilgram, T.; Karger, M.: Teilnahme von Erdgas-BHKW am Regelenergie-
markt. Gas-Plus-Technologien 2014 Juni 2014, S. 2-4.

PLINKE ET AL. 2015

Plinke, W.; Rese, M.; Utzig, B. P.: Industrielle Kostenrechnung. Eine Ein-
führung. 8. Auflage. Berlin, Heidelberg: Springer Vieweg 2015.
ISBN: 978-3-66246-853-1.

POLOTSKI ET AL. 2017

Polotski, V.; Kenne, J.-P.; Gharbi, A.: Production and setup policy opti-
mization for hybrid manufacturing–remanufacturing systems. International
Journal of Production Economics 183 2017, S. 322-333.

POPP ET AL. 2016

Popp, R. S.-H.; Liebl, C.; Zaeh, M. F.: Energy Flexible Machine tool Com-
ponents – An Investigation of Capabilities. Procedia CIRP 57 2016,
S. 692-697.

PRIORE ET AL. 2001

Priore, P.; La Fuennte, D.; Gomez, A.; Puente, J.: A review of machine
learning in dynamic scheduling of flexible manufacturing systems. Artificial
Intelligence for Engineering Design, Analysis and Manufacturing 15
2001 3, S. 251-263.

PRIORE ET AL. 2014

Priore, P.; Gómez, A.; Pino, R.; Rosillo, R.: Dynamic scheduling of manu-
facturing systems using machine learning: An updated review. Artificial In-
telligence for Engineering Design, Analysis and Manufacturing 28 2014 01,
S. 83-97.

PUDLIK ET AL. 2015

Pudlik, M.; Sensfuß, F.; Winkler, J.: Leitstudie Strommarkt Arbeitspaket 4:
Welche Faktoren beeinflussen die Entwicklung des Marktwerts der Erneuer-
baren Energien? Berlin 2015.

PUTERMAN 2009

Puterman, M. L.: Markov Decision Processes. Discrete Stochastic Dynamic
Programming. Hoboken: John Wiley & Sons Inc 2009. ISBN: 978-0-471-
72782-8.

RAGER 2008

Rager, M.: Energieorientierte Produktionsplanung. Analyse, Konzeption und Umsetzung. Wiesbaden: Betriebswirtschaftlicher Verlag Dr. Th. Gabler / GWV Fachverlage GmbH 2008. ISBN: 978-3-8350-5569-8.

RAGER ET AL. 2015

Rager, M.; Gahm, C.; Denz, F.: Energy-oriented scheduling based on Evolutionary Algorithms. Computers & Operations Research 54 2015, S. 218-231.

REGELLEISTUNG.NET 2018

Regelleistung.net: Markt für Regelleistung in Deutschland 2018.
<https://www.regelleistung.net/ext/static/market-information>. Zuletzt abgerufen am: 27.03.2020

REHBERGER ET AL. 2017

Rehberger, S.; Spreiter, L.; Vogel-Heuser, B.: An agent-based approach for dependable planning of production sequences in automated production systems. Automatisierungstechnik (at) 2017 65 (11), S. 766-778.

REHWINKEL 1978

Rehwinkel, G.: Erfolgsorientierte Reihenfolgeplanung. Grundfragen. Wiesbaden: Gabler 1978. ISBN: 978-3-409-34541-5.

REICH & REPPICH 2018

Reich, G.; Reppich, M.: Regenerative Energietechnik. Überblick über ausgewählte Technologien zur nachhaltigen Energieversorgung. 2. Auflage. Wiesbaden: Springer Vieweg 2018. ISBN: 978-3-658-20607-9.

REINHART 2017

Reinhart, G. (Hrsg.): Handbuch Industrie 4.0. Geschäftsmodelle, Prozesse, Technik. München: Hanser 2017. ISBN: 978-3-446-44989-3.

REINHART ET AL. 2012A

Reinhart, G.; Reinhardt, S.; Graßl, M.: Bewertung der Energieflexibilität von Produktionssystemen. wt Werkstattstechnik online 2012 102, S. 622-628.

REINHART ET AL. 2012B

Reinhart, G.; Reinhardt, S.; Graßl, M.: Energieflexible Produktionssysteme. Einführungen zur Bewertung der Energieeffizienz von Produktionssystemen. wt Werkstattstechnik online 102 2012 9, S. 622-628.

REUTER 1995

Reuter, B.: Vernetzte administrative Inseln. Prozeßorientierte Ablauf- und

Aufbauorganisation. Wiesbaden: Gabler Verlag 1995. ISBN: 978-3-663-05904-2.

ROESCH ET AL. 2018

Roesch, M.; Berger, C.; Braunreuther, S.; Reinhart, G.: Cost-model for Energy-oriented Production Control. 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), 2018, S. 158-162

ROESCH ET AL. 2019

Roesch, M.; Linder, C.; Bruckdorfer, C.; Hohmann, A.; Reinhart, G.: Industrial Load Management using Multi-Agent Reinforcement Learning for Rescheduling. Proceedings of the 2019 Second International Conference on Artificial Intelligence for Industries (AI4I), Laguna Hills, CA, USA, 25–27 September 2019; S. 99–102.

ROESCH ET AL. 2020

Roesch, M.; Linder, C.; Zimmermann, R.; Rudolf, A.; Hohman, A.; Reinhart, G.: Smart Grid for Industry Using Multi-Agent Reinforcement Learning. Applied Sciences. 2020, 10 (19), 6900.

RUBAIEE ET AL. 2018

Rubaiee, S.; Cinar, S.; Yildirim, M. B.: An Energy-Aware Multiobjective Optimization Framework to Minimize Total Tardiness and Energy Cost on a Single-Machine Nonpreemptive Scheduling. IEEE Transactions on Engineering Management 2018, S. 1-16.

RUIZ DUARTE ET AL. 2020

Ruiz Duarte, J. L.; Fan, N.; Jin, T.: Multi-process production scheduling with variable renewable integration and demand response. European Journal of Operational Research 281 2020 1, S. 186-200.

SABUNCUOGLU & BAYIZ 2000

Sabuncuoglu, I.; Bayiz, M.: Analysis of reactive scheduling problems in a job shop environment. European Journal of Operational Research 126 2000 3, S. 567-586.

SAFARI ET AL. 2009

Safari, M.; Morcrette, M.; Teyssot, A.; Delacourt, C.: Multimodal Physics-Based Aging Model for Life Prediction of Li-Ion Batteries. Journal of The Electrochemical Society 156 2009 3, S. A145.

SAUER ET AL. 2019

Sauer, A.; Abele, E.; Buhl, H. U.: Energieflexibilität in der deutschen Industrie. Ergebnisse aus dem Kopernikus-Projekt - Synchronisierte und energieadaptive Produktionstechnik zur flexiblen Ausrichtung von Industrieprozessen auf eine fluktuierende Energieversorgung (SynErgie). Stuttgart: Fraunhofer Verlag 2019. ISBN: 978-3-8396-1479-2.

SCHÄFER ET AL. 2016

Schäfer, R.; Chankov, S.; Bendul, J.: What is Really “On-Time”? A Comparison of Due Date Performance Indicators in Production. *Procedia CIRP* 52 2016, S. 124-129.

SCHAUMANN & SCHMITZ 2010

Schaumann, G.; Schmitz, K. W.: Kraft-Wärme-Kopplung. 4. vollständig bearbeitete und erweiterte Auflage. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg 2010. ISBN: 978-3-642-01425-3.

SCHENK 2010

Schenk, M.: Instandhaltung technischer Systeme. Methoden und Werkzeuge zur Gewährleistung eines sicheren und wirtschaftlichen Anlagenbetriebs. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg 2010. ISBN: 978-3-642-03949-2.

SCHMIDT ET AL. 2014

Schmidt, M.; Bertsch, S.; Nyhuis, P.: Schedule compliance operating curves and their application in designing the supply chain of a metal producer. *Production Planning & Control* 25 2014 2, S. 123-133.

SCHUH 2011

Schuh, G.: Handbuch Produktion und Management. 2. vollständig neu bearbeitete und erweiterte Auflage. Berlin: Springer Vieweg 2011. ISBN: 978-3-642-12529-4.

SCHUH & STICH 2012

Schuh, G.; Stich, V.: Produktionsplanung und -steuerung. 4. überarbeitete Auflage. Berlin, Heidelberg: Springer Vieweg 2012. ISBN: 978-3-64225-423-9. (VDI-Buch).

SCHUH & FUß 2015

Schuh, G.; Fuß, C.: ProSense. Ergebnisbericht des BMBF-Verbundprojektes; hochauflösende Produktionssteuerung auf Basis kybernetischer Unterstützungssysteme und intelligenter Sensorik. 1. Auflage. Aachen: Apprimus Verlag 2015. ISBN: 978-3-8635-9346-9.

SCHULMAN ET AL. 2017

Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O.: Proximal Policy Optimization Algorithms. 2017. Verfügbar unter: <http://arxiv.org/abs/1707.0634>

SCHULTZ 2018

Schultz, C. M.: System zur energieorientierten Produktionssteuerung in der auftragsbezogenen Fertigung. Dissertation Technische Universität München. München: Utz 2018.

SCHULTZ & MERTENS 2000

Schultz, J.; Mertens, P.: Untersuchung wissensbasierter und weiterer ausgewählter Ansätze zur Unterstützung der Produktionsfeinplanung - ein Methodenvergleich. *Wirtschaftsinformatik* 42 2000 1, S. 56-65.

SCHULTZ ET AL. 2015

Schultz, C.; Sellmaier, P.; Reinhart, G.: An Approach for Energy-oriented Production Control Using Energy Flexibility. *Procedia CIRP* 29 2015, S. 197-202.

SCHULTZ ET AL. 2016A

Schultz, C.; Braun, S.; Braunreuther, S.; Reinhart, G.: Integration of load management into an energy-oriented production control. 14th Global Conference on Sustainable Manufacturing 2016.

SCHULTZ ET AL. 2016B

Schultz, C.; Braunreuther, S.; Reinhart, G.: Method for an Energy-oriented Production Control. *Procedia CIRP* 48 2016, S. 248-253.

SCHULTZ ET AL. 2017

Schultz, C.; Bayer, C.; Roesch, M.; Braunreuther, S.; Reinhart, G.: Integration of an automated load management in a manufacturing execution system, 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). Singapore, 10.12.2017 - 13.12.2017, S. 494-498. ISBN: 978-1-5386-0948-4.

SCHUMACHER & WÜRFEL 2015

Schumacher, I.; Würfel, P.: Strategien zur Strombeschaffung in Unternehmen. Energieeinkauf optimieren, Kosten senken. Wiesbaden: Springer Gabler 2015. ISBN: 978-3-658-07421-0.

SHAFAEI & BRUNN 1999

Shafaei, R.; Brunn, P.: Workshop scheduling using practical (inaccurate) data Part 1. The performance of heuristic scheduling rules in a dynamic job

shop environment using a rolling time horizon approach. *International Journal of Production Research* 37 1999 17, S. 3913-3925.

SHAHZAD & MEBARKI 2012

Shahzad, A.; Mebarki, N.: Data mining based job dispatching using hybrid simulation-optimization approach for shop scheduling problem. *Engineering Applications of Artificial Intelligence* 25 2012 6, S. 1173-1181.

SHAHZAD & MEBARKI 2016

Shahzad, A.; Mebarki, N.: Learning Dispatching Rules for Scheduling: A Synergistic View Comprising Decision Trees, Tabu Search and Simulation. *Computers* 5 2016 1, S. 3.

SHAW ET AL. 1992

Shaw, M. J.; Park, S.; Raman, N.: Intelligent Scheduling with Machine Learning Capabilities: The Induction of Scheduling Knowledge. *IIE Transactions* 24 1992 2, S. 156-168.

SHEN ET AL. 2001

Shen, W.; Norrie, D. H.; Barthès, J.-P.: Multi-agent systems for concurrent intelligent design and manufacturing. London: Taylor & Francis 2001. ISBN: 978-0-7484-0882-5.

SHEN ET AL. 2006

Shen, W.; Hao, Q.; Yoon, H. J.; Norrie, D. H.: Applications of agent-based systems in intelligent manufacturing. An updated review. *Advanced Engineering Informatics* 20 2006 4, S. 415-431.

SHEN ET AL. 2014

Shen, X.; Zhang, M.; Fu, J.: Multi-Objective dynamic job shop scheduling: A survey and prospects. *International journal of innovative computing, information & control* 2014 10(6), S. 2113-2126.

SILVER ET AL. 2016

Silver, D.; Huang, A.; Maddison, C. J.; Guez, A.; Sifre, L.; van den Driessche, G.; Schrittwieser, J.; Antonoglou, I.; Panneershelvam, V.; Lanctot, M.; Dieleman, S.; Grewe, D.; Nham, J.; Kalchbrenner, N.; Sutskever, I.; Lillicrap, T.; Leach, M.; Kavukcuoglu, K.; Graepel, T.; Hassabis, D.: Mastering the game of Go with deep neural networks and tree search. *Nature* 529 2016, S. 484-489.

SIMON 2020

Simon, P.: Methodik zur risikoorientierten Bewertung von Energieflexibilität von Produktionssystemen. Dissertation Technische Universität München: Utz 2020.

STERNER & STADLER 2017

Sterner, M.; Stadler, I.: Energiespeicher - Bedarf, Technologien, Integration. 2. korrigierte und ergänzte Auflage. Berlin: Springer Vieweg 2017. ISBN: 978-3-66248-892-8.

STOCK & SELIGER 2015

Stock, T.; Seliger, G.: Multi-objective Shop Floor Scheduling Using Monitored Energy Data. *Procedia CIRP* 26 2015, S. 510-515.

STOLLE ET AL. 2018

Stolle, T.; Hankeln, C.; Blaurock, J.: Die Bedeutung der Energiespeicherbranche für das Energiesystem und die Gesamtwirtschaft in Deutschland. *ENERGIEWIRTSCHAFTLICHE TAGESFRAGEN* 2018 9, S. 54-56.

STROHBÜCKER 2011

Strohbücker, S.: Bepreisen von Preis- und Mengenrisiken der Strombeschaffung unter Berücksichtigung von Portfolioaspekten bei Großkunden im Strommarkt. Zugl.: Essen, Univ., Diss., 2011. 1. Auflage. Wiesbaden: Gabler Verlag / Springer Fachmedien Wiesbaden GmbH Wiesbaden 2011. ISBN: 978-3-8349-3026-2.

SUN & LI 2013

Sun, Z.; Li, L.: Opportunity Estimation for Real-Time Energy Control of Sustainable Manufacturing Systems. *IEEE Transactions on Automation Science and Engineering* 10 2013 1, S. 38-44.

SUN ET AL. 2014

Sun, Z.; Li, L.; Fernandez, M.; Wang, J.: Inventory control for peak electricity demand reduction of manufacturing systems considering the tradeoff between production loss and energy savings. *Journal of Cleaner Production* 82 2014, S. 84-93.

SUN ET AL. 2016

Sun, Z.; Li, L.; Dababneh, F.: Plant-level electricity demand response for combined manufacturing system and heating, venting, and air-conditioning (HVAC) system. *Journal of Cleaner Production* 135 2016, S. 1650-1657.

SUN ET AL. 2018

Sun, Z.; Dababneh, F.; Li, L.: Joint Energy, Maintenance, and Throughput

Modeling for Sustainable Manufacturing Systems. IEEE Transactions on Systems, Man, and Cybernetics: Systems 2018, S. 1-12.

SUTTON & BARTO 1998

Sutton, R. S.; Barto, A.: Reinforcement learning. An introduction. Cambridge, Massachusetts, London: The MIT Press 1998.
ISBN: 9780262193986.

SUTTON & BARTO 2018

Sutton, R. S.; Barto, A.: Reinforcement learning. An introduction. Second edition. Cambridge, MA, London: The MIT Press 2018. ISBN: 978-0-262-19398-6.

TALUKDAR & GELLINGS 1987

Talukdar, S.; Gellings, C. W.: Load management. New York: IEEE Press 1987. ISBN: 0879422149. (IEEE Press selected reprint series).

TAN 1997

Tan, M.: Multi-Agent Reinforcement Learning: Independent versus Cooperative Agents. 1997 Readings in agents, S. 487-494.

TEMPELMEIER 2003

Tempelmeier, H.: Material-Logistik. Modelle und Algorithmen für die Produktionsplanung und -steuerung und das Supply Chain Management. 5. neubearb. Auflage. Berlin: Springer 2003. ISBN: 978-3-540-70906-0.

THEISS 2015

Theiss, S.: Echtzeitfähige Softwareagenten zur Realisierung cyber-physischer Produktionssysteme. Dissertation Technische Universität Dresden 2015.

THIELMANN ET AL. 2015

Thielmann, A.; Sauer, A.; Schnell, M.; Isenmann, R.; Wietschel, M.: Technologie-Roadmap Stationäre Energiespeicher 2030. Karlsruhe: Fraunhofer ISI 2015.

THIELMANN ET AL. 2018

Thielmann, A.; Neef, C.; Fenske, C.; Wietschel, M.: Energiespeicher-Monitoring 2018. Leitmarkt- und Leitanbieterstudie. Lithium-Ionen-Batterien für die Elektromobilität. Fraunhofer ISI. Karlsruhe 2018.

TILLENBURG 2008

Tillenburg, S.: Das Konzept der Ausbalancierten Produktion in Supply

Chains. Dissertation Universität Lüneburg. Wiesbaden: Betriebswirtschaftlicher Verlag Dr. Th. Gabler / GWV Fachverlage GmbH Wiesbaden 2008. ISBN: 978-3-83490-867-4.

ULRICH ET & HILL 1976

Ulrich, P.; Hill, W.: Wissenschaftstheoretische Grundlagen der Betriebswirtschaftslehre. Wirtschaftswissenschaftliches Studium: Zeitschrift für Ausbildung und Hochschulkontakt 1976 5, S. 304-309.

UMWELTBUNDESAMT 2018

Umweltbundesamt: Erneuerbare Energien in Deutschland Daten zur Entwicklung im Jahr 2017. Dessau-Roßlau: UBA 2018.

UN 2015A

United Nations (UN): Paris Agreement. Paris: UN. http://unfccc.int/files/essential_background/convention/application/pdf/english_paris_agreement.pdf. Zuletzt abgerufen am: 27.03.2020.

UN 2015B

United Nations (UN): Transforming our world: The 2030 Agenda for sustainable development. New York: UN 2015.

VAHRENKAMP & SIPERMANN 2008

Vahrenkamp, R.; Siepermann, C.: Produktionsmanagement. 6. überarbeitete Auflage. München: Oldenbourg 2008. ISBN: 978-3-4865-8784-5.

VAN LAARHOVEN & AARTS 1992

van Laarhoven, P. J.M.; Aarts, E. H.: Simulated annealing. Theory and applications. Reprint. Dordrecht: Kluwer 1992. ISBN: 978-94-015-7744-1.

VBW 2012

Vereinigung der Bayerischen Wirtschaft (vbw): Energieeffizienz in der Industrie. München: vbw 2012.

VDI-RICHTLINIE 2653

Verein Deutscher Ingenieure e.V. (VDI): VDI-Richtlinie 22653 Blatt 1: Agentensysteme in der Automatisierungstechnik - Grundlagen. Düsseldorf: Beuth. 2017.

VDI-RICHTLINIE 2815

Verein Deutscher Ingenieure e.V. (VDI): VDI-Richtlinie 2815 Blatt 1: Begriffe für die Produktionsplanung und -steuerung: Begriffsdzusammenhänge und Begriffsdefinition. 4. Auflage: Gesellschaft Produktionstechnik. Düsseldorf: VDI-Verlag 1992.

VDI-Richtlinie 3258

Verein Deutscher Ingenieure e.V. (VDI): VDI-Richtlinie 3258 Blatt 1: Kostenrechnung mit Maschinenstundensätze - Begriffe, Beschreibungen, Zusammenhänge: Blatt 1. Berlin: Beuth 1962.

VDI-Richtlinie 2067

Verein Deutscher Ingenieure e.V. (VDI) 2067: VDI-Richtlinie 2067 Blatt 7: Berechnung der Kosten von Wärmeversorgungsanlagen; Blockheizkraftwerke: Blatt 7. Düsseldorf: Beuth-Verlag 1988.

VDMA 2013

Verband Deutscher Maschinen- und Anlagenbau (VDMA): Der Maschinen- und Anlagenbau kann die Energiewende gestalten. Energiewende-Barometer Maschinenbau. Berlin: VDMA 2013.

VEITH ET AL. 2020

Veith, E.; Balduin, S.; Wenninghoff, N.; Tröschel, M.; Fischer, L.; Nieße, A.; Wolgast, T.; Sethmann, R.; Fraune, B.; Woltjen, T.: Analyzing Power Grid, ICT, and Market Without Domain Knowledge Using Distributed Artificial Intelligence. CYBER 2020, The Fifth International Conference on Cyber-Technologies and Cyber-Systems; 2020.

VETTER ET AL. 2005

Vetter, J.; Novák, P.; Wagner, M. R.; Veit, C.; Möller, K.-C.; Besenhard, J. O.; Winter, M.; Wohlfahrt-Mehrens, M.; Vogler, C.; Hammouche, A.: Ageing mechanisms in lithium-ion batteries. Journal of Power Sources 147 2005 1-2, S. 269-281.

VIEIRA ET AL. 2003

Vieira, G. E.; Herrmann, J. W.; Lin, E.: Rescheduling Manufacturing Systems: A Framework o Strategies, Policies And Methods. Journal of Scheduling 6 2003 1, S. 39-62.

VOGEL-HEUSER ET AL. 2015

Vogel-Heuser, B.; Lee, J.; Leitão, P.: Agents enabling cyber-physical production systems. at-Automatisierungstechnik 2015 (2015) 63(10), S. 777-789.

WAGNER ET AL. 2003

Wagner, T.; Urbano, P.; Göhrner, P.: Softwareagenten-Einführung und Überblick über eine alternative Art der Softwareentwicklung. Teil 1: Agentenorientierte Softwareentwicklung. Automatisierungstechnische Praxis atp 45 (2003).

WANG & LUO 2016

Wang, D.; Luo, H.: Simultaneous Lot-Sizing and Scheduling for Single-Stage Multi-product Production-Distribution System with Due Date Considerations to Minimize Total Logistics Cost. 2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2016, S. 200-203

WANG & LI 2013

Wang, Y.; Li, L.: Time-of-use based electricity demand response for sustainable manufacturing systems. *Energy* 63 2013, S. 233-244.

WANG ET AL. 2017

Wang, W.; Yang, H.; Zhang, Y.; Xu, J.: IoT-enabled real-time energy efficiency optimisation method for energy-intensive manufacturing enterprises. *International Journal of Computer Integrated Manufacturing* 2017, S. 1-18.

WANG ET AL. 2019A

Wang, J.; Zhang, Y.; Liu, Y.; Wu, N.: Multiagent and Bargaining-Game-Based Real-Time Scheduling for Internet of Things-Enabled Flexible Job Shop. *IEEE Internet of Things Journal* 6 2019 2, S. 2518-2531.

WANG ET AL. 2019B

Wang, T.; Bao, X.; Clavera, I.; Hoang, J.; Wen, Y.; Langlois, E.; Zhang, S.; Zhang, G.; Abbeel, P.; Ba, J.: Benchmarking Model-Based Reinforcement Learning. 2019.

WANG ET AL. 2019C

Wang, Y.; Liu, H.; Zheng, W.; Xia, Y.; Li, Y.; Chen, P.; Guo, K.; Xie, H.: Multi-Objective Workflow Scheduling With Deep-Q-Network-Based Multi-Agent Reinforcement Learning. *IEEE Access* 7 2019, S. 39974-39982.

WANNAGAT & VOGEL-HEUSER 2008

Wannagat, A.; Vogel-Heuser, B.: Agent oriented software-development for networked embedded systems with real time and dependability requirements the domain of automation. 2008 17th World Congress of International Federation of Automation Control (IFAC). Seoul, Korea, S. 4144-4149.

WASCHNECK ET AL. 2018

Waschneck, B.; Reichstaller, A.; Belzner, L.; Altenmuller, T.; Bauernhansl, T.; Knapp, A.; Kyek, A.: Deep reinforcement learning for semiconductor production scheduling, 2018 29th Annual SEMI Advanced Semiconductor

Manufacturing Conference (ASMC). Saratoga Springs, NY, USA, S. 301-306.

WEBER 2012

Weber, J.: Logistikkostenrechnung. Kosten-, Leistungs- und Erlösinformationen zur erfolgsorientierten Steuerung der Logistik. 3. Auflage. Berlin, Heidelberg: Springer Vieweg 2012. ISBN: 978-3-6422-5172-6.

WECKMANN ET AL. 2017

Weckmann, S.; Kuhlmann, T.; Sauer, A.: Decentral Energy Control in a Flexible Production to Balance Energy Supply and Demand. Procedia CIRP 61 2017, S. 428-433.

WEINERT 2010

Weinert, N.: Vorgehensweise für Planung und Betrieb energieeffizienter Produktionssysteme. Dissertation Technische Universität Berlin. Stuttgart: Fraunhofer Verlag 2010. ISBN: 978-3-83960-173-0.

WEINERT & MOSE 2016

Weinert, N.; Mose, C.: Electrical Load Management for Production Equipment Applying a Decentralized Optimization Approach. Procedia CIRP 48 2016, S. 259-264.

WEIß 2010

Weiß, M.: Datenauswertung von Energiemanagementsystemen. Datenerfassung, Messwertdarstellung und -interpretation, Kennwerte zur Energieverteilung, Praxisbeispiele. Erlangen: Publicis Publ 2010. ISBN: 978-3-8957-8347-0.

WEITZEL & GLOCK 2018

Weitzel, T.; Glock, C. H.: Scheduling a storage-augmented discrete production facility under incentive-based demand response. International Journal of Production Research 2018, S. 1-21.

WEIZSÄCKER 1997

Weizsäcker, E. U.: Grenzen-los? Jedes System braucht Grenzen - aber wie durchlässig müssen diese sein?. Birkhäuser Basel 1997. ISBN: 978-3-03486-106-9.

WENIGER ET AL. 2014

Weniger, J.; Bergner, J.; Tjaden, T.; Quaschnig, V.: Bedeutung von prognosebasierten Betriebsstrategien für die Netzintegration von PV-Speichersystemen. Kloster Banz, Bad Staffelstein: 2014.

WIENDAHL 2010

Wiendahl, H.-P.: Betriebsorganisation für Ingenieure. 7. aktualisierte Auflage. München: Hanser 2010. ISBN: 978-3-446-42288-9.

WIENDAHL & WEDEMEYER 1993

Wiendahl, H.-P.; Wedemeyer, H.-G. v.: Das Dilemma der Fertigungssteuerung. In: Industrielles Management. Beschaffung - Produktion - Qualität - Innovation - Umwelt Reader zur Industriebetriebslehre. Wiesbaden, s.l.: Gabler Verlag 1993, S. 162-177. ISBN: 978-3-663-02131-5.

WILLEKE 2018

Willeke, S.: Integration volatiler Energiepreise in die Fertigungssteuerung. Dissertation Leibniz Universität Hannover. Hannover: TEWISS - Technik und Wissen GmbH 2018. ISBN: 9783959002059.

WILLEKE ET AL. 2015A

Willeke, S.; Wesebaum, S.; Ullmann, G.: Integration zeitvariabler Energiekosten in die Belastungsorientierte Auftragsfreigabe. Logistics Journal 2015. <https://www.logistics-journal.de/not-reviewed/2015/07/4200>. Zuletzt abgerufen am: 16.09.2019

WILLEKE ET AL. 2015B

Willeke, S.; Wesebaum, S.; Ullmann, G.; Nyhuis, P.: Energiekostenorientierte Reihenfolgebildung. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb 110 2015 9, S. 495-499.

WILLEKE ET AL. 2015C

Willeke, S.; Wesebaum, S.; Prinzhorn, H.; Ullmann, G.; Nyhuis, P.: Energiekostenorientierte Rückstandsregelung. IT&Production 16 2015. http://www.it-production.com/index.php?seite=einzel_artikel_an_sicht&id=62954. Zuletzt abgerufen am: 16.09.2019.

WILLEKE ET AL. 2018

Willeke, S.; Prinzhorn, H.; Stonis, M.; Nyhuis, P.: Preconditions for applying an energy price-oriented sequencing rule. Production Engineering 12 2018 1, S. 73-81.

WILLEKE ET AL. 2016

Willeke, S.; Ullmann, G.; Nyhuis, P.: Method for an Energy-Cost-Oriented Manufacturing Control to Reduce Energy Costs: Energy Cost Reduction by Using a New Sequencing Method, 2016 International Conference on Industrial Engineering, Management Science and Application (ICIMSA). Jeju, South Korea, 5/23/2016 - 5/26/2016, S. 1-5. ISBN: 978-1-5090-1671-6.

WIRTH 2018

Wirth, H.: Aktuelle Fakten zur Photovoltaik in Deutschland. Freiburg: Fraunhofer ISE 2018. <https://www.ise.fraunhofer.de/content/dam/ise/de/documents/publications/studies/aktuelle-fakten-zur-photovoltaik-in-deutschland.pdf>. Zuletzt abgerufen am: 19.05.2019.

WÜNSCH ET AL. 2014

Wünsch, M.; Eikmeier, B.; Eberhard, J.; Gailfuß, M.: Potenzial- und Kosten-Nutzen-Analyse zu den Einsatzmöglichkeiten von Kraft-Wärme-Kopplung (Umsetzung der EU-Energieeffizienzrichtlinie) sowie Evaluierung des KWKG im Jahr 2014. BMWi 2014.

XU ET AL. 2018

Xu, B.; Oudalov, A.; Ulbig, A.; Andersson, G.; Kirschen, D. S.: Modeling of Lithium-Ion Battery Degradation for Cell Life Assessment. IEEE Transactions on Smart Grid 9 2018 2, S. 1131-1140.

YU ET AL. 2016

Yu, M.; Lu, R.; Hong, S. H.: A real-time decision model for industrial load management in a smart grid. Applied Energy 183 2016, S. 1488-1497.

ZAPF 2017

Zapf, M.: Stromspeicher und Power-to-Gas im deutschen Energiesystem. Wiesbaden: Springer Fachmedien Wiesbaden 2017. ISBN: 978-3-658-15072-3.

ZÄPFEL 2001

Zäpfel, G.: Grundzüge des Produktions- und Logistikmanagement. 2. unwesentlich veränd. Auflage. München: Oldenbourg 2001. ISBN: 9783486256185.

ZHAI ET AL. 2017A

Zhai, Y.; Biel, K.; Zhao, F.; Sutherland, J. W.: Dynamic scheduling of a flow shop with on-site wind generation for energy cost reduction under real time electricity pricing. CIRP Annals - Manufacturing Technology 66 2017 1, S. 41-44.

ZHAI ET AL. 2017B

Zhai, Y.; Wang, H.; Zhao, F.; Sutherland, J. W.: Dynamic Manufacturing Scheduling Under Real-Time Electricity Pricing Based on MILP and ARIMA, ASME 2017 12th International Manufacturing Science and Engi-

neering Conference collocated with the JSME/ASME 2017 6th International Conference on Materials and Processing. Los Angeles, California, USA, 04.-08.06.2017. ISBN: 978-0-7918-5075-6.

ZHANG & WHITE 2008

Zhang, Q.; White, R. E.: Capacity fade analysis of a lithium ion cell. *Journal of Power Sources* 179 2008 2, S. 793-798.

ZHANG ET AL. 2015

Zhang, T.; Zheng, Q. P.; Fang, Y.; Zhang, Y.: Multi-level inventory matching and order planning under the hybrid Make-To-Order/Make-To-Stock production environment for steel plants via Particle Swarm Optimization. *Computers & Industrial Engineering* 87 2015, S. 238-249.

ZHANG ET AL. 2017A

Zhang, H.; Cai, J.; Fang, K.; Zhao, F.; Sutherland, J. W.: Operational optimization of a grid-connected factory with onsite photovoltaic and battery storage systems. *Applied Energy* 205 2017, S. 1538-1547.

ZHANG ET AL. 2017B

Zhang, H.; Zhao, F.; Sutherland, J. W.: Scheduling of a Single Flow Shop for Minimal Energy Cost Under Real-Time Electricity Pricing. *Journal of Manufacturing Science and Engineering* 139 2017 1, S. 14502.

ZHANG ET AL. 2017C

Zhang, X.; Hug, G.; Harjunoski, I.: Cost-Effective Scheduling of Steel Plants With Flexible EAFs. *IEEE Transactions on Smart Grid* 8 2017 1, S. 239-249.

ZHANG ET AL. 2019

Zhang, K.; Yang, Z.; Başar, T.: *Multi-Agent Reinforcement Learning: A Selective Overview of Theories and Algorithms*. 2019.

ZHOU ET & LI 2013

Zhou, Z.; Li, L.: Real time electricity demand response for sustainable manufacturing systems considering throughput bottleneck detection. *IEEE International Conference on Automation Science and Engineering (CASE)* 2013, S. 640-644.

ZHOU ET AL. 2018

Zhou, S.; Li, X.; Du, N.; Pang, Y.; Chen, H.: A multi-objective differential evolution algorithm for parallel batch processing machine scheduling considering electricity consumption cost. *Computers & Operations Research* 96 2018, S. 55-68.