

Dorothea Schwung

Maschinelle Lernalgorithmen zur
Selbstopтимierung in verteilten
Produktionssystemen basierend auf
spieltheoretischen Konzepten

Maschinelle Lernalgorithmen zur Selbstoptimierung in verteilten Produktionssystemen basierend auf spieltheoretischen Konzepten

Von der Fakultät für Ingenieurwissenschaften,
Abteilung Elektrotechnik und Informationstechnik
der Universität Duisburg-Essen

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften (Dr.-Ing.)

genehmigte Dissertation

von

Dorothea Schwung

1. Gutachter: Prof. Dr.-Ing. Steven X. Ding
 2. Gutachter: Prof. Dr. Oliver Niggemann
- Tag der mündlichen Prüfung: 26. November 2020

Automatisierungstechnik / Prozessinformatik

Dorothea Schwung

**Maschinelle Lernalgorithmen zur Selbstoptimierung
in verteilten Produktionssystemen basierend auf
spieltheoretischen Konzepten**

Shaker Verlag
Düren 2021

Bibliografische Information der Deutschen Nationalbibliothek

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

Zugl.: Duisburg-Essen, Univ., Diss., 2020

Copyright Shaker Verlag 2021

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

Printed in Germany.

ISBN 978-3-8440-7840-4

ISSN 1614-9874

Shaker Verlag GmbH • Am Langen Graben 15a • 52353 Düren

Telefon: 02421 / 99 0 11 - 0 • Telefax: 02421 / 99 0 11 - 9

Internet: www.shaker.de • E-Mail: info@shaker.de

Vorwort

Die vorliegende Arbeit entstand neben meiner Tätigkeit als Lehrkraft für besondere Aufgaben an dem Fachbereich Elektrische Energietechnik der Fachhochschule Südwestfalen am Standort Soest. Im Gegensatz zu einer Fortsetzung meiner akademischen Laufbahn direkt nach meinem Studienabschluss an der Universität Duisburg-Essen im Jahre 2010, habe ich mich nach über fünf Jahren Industrietätigkeit zu einer Promotion entschlossen. Ausschlaggebend dafür war zum einen mein allgegenwärtiges Interesse an dem Thema “Künstliche Intelligenz” und zum anderen der Wunsch danach, mich mit diesem Themenfeld, speziell mit dem maschinellen Lernen, näher zu befassen.

Umso größer war meine Freude als Herr Prof. Dr.-Ing. Steven X. Ding, Leiter des Instituts Automatisierungstechnik und Komplexe Systeme (AKS) an der Universität Duisburg-Essen, meine Anfrage, diese Arbeit als Doktorvater zu betreuen, auf Anhieb begrüßte. Deshalb gilt Herrn Prof. Dr.-Ing. Steven X. Ding mein besonderer Dank für die Möglichkeit meine Arbeit unter seiner wissenschaftlichen Betreuung angefertigt haben zu dürfen.

Des Weiteren möchte ich mich bei Herrn Prof. Dr. Oliver Niggemann für die Übernahme des Korreferats und das damit verbundene Interesse an dieser Arbeit bedanken.

Ich bedanke mich ebenfalls bei allen Studierenden, die mit Ihren Abschlussarbeiten zum Gelingen dieser Arbeit beigetragen haben. Mein besonderer Dank gilt dabei vor allem Herrn Jan Niclas Reimann, M.Sc., sowie Herrn Vatsal Patel, M.Sc., und Herrn Steve Yuwono, M.Sc..

Frau Sabine Bay vom Institut AKS danke ich vielmals für die Unterstützung bei organisatorischen Fragestellungen.

Meinen herzlichsten Dank richte ich an meinen Mann, Herrn Prof. Dr.-Ing. Andreas Schwung, der mir auch während meiner Promotionszeit stets liebevoll zur Seite stand.

*Für meinen
Lieblingsmenschen.*

Inhaltsverzeichnis

Nomenklatur	VII
1 Einleitung	1
1.1 Stand der Forschung und Technik - Maschinelles Lernen in der Produktion	4
1.2 Motivation und Zielsetzung	8
1.3 Strukturierung der Arbeit	10
2 Problemformulierung	11
2.1 Systembeschreibung	11
2.1.1 Anforderungen an die industrielle Produktion	13
2.1.2 Das Optimierungsproblem	14
3 Verteiltes Maschinelles Lernen	15
3.1 Reinforcement Learning	15
3.1.1 Q-Learning	17
3.1.2 Actor-Critic Reinforcement Learning	18
3.1.3 Künstliche Neuronale Netze als Funktionsapproximatoren	21
3.2 Die Spieltheorie als rationales Werkzeug der Künstlichen Intelligenz	23
3.2.1 Übersicht der wichtigsten Spielformen	25
3.2.2 Optimale Lösungsstrategien	26
3.2.3 Potentialspiele	28
3.2.4 Zustandsbasierte Potentialspiele	29
3.3 Relation zwischen Spieltheorie und Reinforcement Learning	31
3.4 Zusammenfassung	31
4 Verteilte Produktionssysteme als Potentialspiele	33
4.1 Modellierung von modularen Produktionssystemen	33
4.2 Definition der globalen und lokalen Nutzenfunktionen	34
4.3 Lernszenarien	35
4.3.1 Lernen mit zufälliger Initialisierung	36
4.3.2 SPS-informiertes Lernen	38
4.3.3 Meta-Lernen	43
4.4 Zusammenfassung	45
5 Lernverfahren für Potentialspiele	47
5.1 Verteiltes Reinforcement Learning	47
5.2 Spieltheoretische Lernalgorithmen	49
5.2.1 Synchrones und asynchrones Lernen	49
5.2.2 1-Recall Reinforcement Learning	51
5.2.3 Log-Lineares Lernen	52

5.2.4	Interpolierende zustandsbasierte Potentialspiele	53
5.2.5	Kommunikationsbasiertes Lernen	57
5.2.6	Lernen mit Gedächtnis	58
5.3	Zusammenfassung	60
6	Applikationen und Ergebnisse	61
6.1	Modulare Prozess- und Produktionsanlagen: Beispiel Schüttgutanlage . . .	61
6.1.1	Modellierung des Systems	62
6.1.2	Produktionsszenarien	63
6.1.3	Grundsätzlicher Spielbau für die Lernszenarien	64
6.2	Lernen mit zufälliger Initialisierung	68
6.2.1	Ergebnisse für die globale Interpolation	68
6.2.2	Vergleich der Algorithmen	75
6.2.3	Vergleich von globaler mit lokaler Interpolation	78
6.2.4	Vergleich von synchronem mit asynchronem Lernen	80
6.2.5	Vergleich der Kommunikationsformen	80
6.2.6	Lernen mit Gedächtnis	85
6.2.7	Generalisierung	86
6.2.8	Diskussion	88
6.3	SPS-informiertes Lernen	89
6.3.1	Umsetzung	89
6.3.2	Ergebnisse	89
6.4	Meta-Lernen	92
6.4.1	Umsetzung	92
6.4.2	Ergebnisse	94
6.5	Zusammenfassung	97
7	SPS-Implementierung maschineller Lernalgorithmen	99
7.1	Hardware	99
7.2	Software	100
7.3	Implementierung	101
7.4	Ergebnisse	103
7.5	Zusammenfassung	105
8	Zusammenfassung und Ausblick	107
A	Anhang	109
A.1	SPS-Programm	109
A.2	Daten zur Schüttgutanlage	111
A.3	Hyperparameterwerte der Lernalgorithmen	114
	Literaturverzeichnis	117

Nomenklatur

Die folgende Nomenklatur stellt eine Übersicht der in der vorliegenden Arbeit verwendeten Abkürzungen und Formelzeichen in Anlehnung an die DIN 1338 dar. Demnach werden alle skalaren Größen, physikalischen Konstanten und Funktionen kursiv dargestellt. Skalare Variablen und Konstanten als Indizes werden ebenfalls kursiv geschrieben, wohingegen erklärende Variablenindizes gerade abgebildet werden. Mathematische Standardfunktionen und Konstanten werden ebenfalls durch gerade Schriftzeichen ausgedrückt. Vektoren sowie Matrizen werden durch fett gedruckte Buchstaben beschrieben, wobei zwischen Kleinschreibung für Vektoren und Großschreibung für Matrizen unterschieden wird.

Zur besseren Lesbarkeit wird in den Lernalgorithmen auf die Angabe der Zeitabhängigkeit bei den betreffenden mathematischen Ausdrücken verzichtet.

Da im Bereich des maschinellen Lernens die Fachtermini aus dem Englischen stammen, wird oft der englische Fachausdruck verwendet, insbesondere dann, wenn es keine direkte Übersetzung gibt.

Allgemeine Akronyme

Zeichen	Beschreibung
ACRL	Aktor-Kritik Bestärkungslernen (engl. <i>actor-critic reinforcement learning</i>)
AGT	<i>algorithmic game theory</i>
ALK	Aktions-Leistungskennfeld
ÄR	Änderungsrate
BCL	<i>basic control layer</i>
BR	<i>best response</i>
CFC	<i>continuous function chart</i>
CGT	<i>computational game theory</i>
CNN	<i>convolutional neural network</i>
CPS	Cyber-Physische Systeme (engl. <i>cyber-physical systems</i>)
CPU	<i>central processing unit</i>
DB	Datenbaustein
DDPG	<i>deep deterministic policy gradient</i>
DL	tiefes Lernen (engl. <i>deep learning</i>)
DNN	tiefes neuronales Netz (engl. <i>deep neural network</i>)
DP	dynamische Programmierung (engl. <i>dynamic programming</i>)
DPS	dynamisches Potentialspiel
DQN	tiefes Q-Netzwerk (engl. <i>deep Q-network</i>)
DRL	tiefes Bestärkungslernen (engl. <i>deep reinforcement learning</i>)
EGT	<i>engineering game theory</i>
FB	Funktionsbaustein
FBS	Funktionsbausteinsprache
FC	Funktion

Zeichen	Beschreibung
GAN	generatives Netz
HMI	<i>human machine interface</i>
IPC	Industrie-PC
KI	Künstliche Intelligenz
KNN	künstliches neuronales Netz
KOP	Kontaktplan
LLL	Log-lineares Lernen (engl. <i>log-linear learning</i>)
LSTM	<i>long short-term memory</i>
MARL	Multi-Agenten Bestärkungslernen (engl. <i>multi-agent reinforcement learning</i>)
MAS	Multi-Agenten System (engl. <i>multi-agent system</i>)
MC	<i>Monte-Carlo</i>
MEP	Markov Entscheidungsprozess
MG	<i>Markov Games</i>
ML	maschinelles Lernen (engl. <i>machine learning</i>)
MLP	mehrschichtiges Perzeptron (engl. <i>multi-layer perceptron</i>)
MPE	modulare Produktionseinheit
MPC	<i>model predictive control</i>
NGG	Nash-Gleichgewicht
NLK	Nutzen-Leistungskennfeld
OB	Objektbaustein
PBMEP	partiell beobachtbarer Markov Entscheidungsprozess
PC	<i>personel computer</i>
PE	<i>policy evaluation</i>
PI	<i>policy improvement</i>
PIT	<i>policy iteration</i>
PO	Pareto-Optimum
PS	Potentialspiel
ReLU	<i>rectified linear unit</i>
RL	Bestärkungslernen (engl. <i>reinforcement learning</i>)
1-RRL	<i>1-Recall RL</i>
RNN	<i>recurrent neural network</i>
SCL	<i>supervisory control layer</i>
SGG	Stackelberg-Gleichgewicht
SOA	service-orientierte Architektur (engl. <i>service-oriented architectures</i>)
SPS	Speicher-Programmierbare Steuerung
ST	Spieltheorie (engl. <i>game theory</i> (GT))
SuP	Sub-Prozess
SuPS	Sub-Prozess Steuerung
SuSy	Sub-System
SVM	Stützvektormethode (engl. <i>support vector machine</i>)
TD	<i>temporal difference</i>
TIA	<i>totally integrated automation</i>
VIT	<i>value iteration</i>
VP	Vakuumpumpe
ZBPS	zustandsbasiertes Potentialspiel

Akronyme der Lernalgorithmen

Zeichen	Beschreibung
dSPSi	diskretes SPS-informiertes Lernen
GlobiZBPS-ÄR	Global interpolierendes ZBPS mit Änderungsrate
GlobiZBPS-ÄR-asyn	Global interpolierendes ZBPS mit Änderungsrate u. asynchronem Lernen
GlobiZBPS-ÄR-syn	Global interpolierendes ZBPS mit Änderungsrate u. synchronem Lernen (entspricht GlobiZBPS-ÄR)
GlobiZBPS-BR	Global interpolierendes ZBPS mit Best Response
GlobiZBPS-BR-koma	Global interpolierendes ZBPS mit Best Response u. Kommunikation der Agenten
GlobiZBPS-BR-koms	Global interpolierendes ZBPS mit Best Response u. Kommunikation der Zustände (entspricht GlobiZBPS-BR)
GlobiZBPS-BR-komsa	Global interpolierendes ZBPS mit Best Response u. Kommunikation der Agenten u. Zustände
GlobiZBPS-MS	Global interpolierendes ZBPS mit Varianzlernen
kSPSi	kontinuierliches SPS-informiertes Lernen
LokiZBPS-BR	Lokal interpolierendes ZBPS mit Best Response
LokiZBPS-BR-mem	Lokal interpolierendes ZBPS mit Best Response u. Gedächtnis
LokiZBPS-BR-omem	Lokal interpolierendes ZBPS mit Best Response ohne Gedächtnis (entspricht LokiZBPS-BR)

Mathematische Formelzeichen

Zeichen	Beschreibung
$a_i, a_{i,d}, a_{i,c}$	beliebige, diskrete, kontinuierliche Aktion des i -ten Agenten
\mathbf{a}_{-i}	Aktionsvektor aller Agenten ohne der Aktion des i -ten Agenten
\mathbf{a}_i^e	Eingangsvektor der externen Größen des i -ten Agenten
\mathbf{a}_i^s	Eingangsvektor der Steuerungsgrößen des i -ten Agenten
\mathbf{a}_{\max}^i	bisher beste Aktion des ZBPS im l -ten Stützvektor
\mathbf{a}_{SPS}	Aktionen des Steuerungsgesetzes der SPS
$\hat{\mathbf{a}}^{\text{PS}}$	Schätzwerte der Aktionen beim Push-Sum Algorithmus
A	Vorteilsfunktion (engl. <i>advantage function</i>)
A_i	i -ter Agent
$\mathcal{A}, \mathcal{A}_i$	Aktionsmenge aller Agenten, des i -ten Agenten
$\mathcal{A}_{i,j}$	Aktionsmenge des j -ten Aktuators des i -ten Agenten
$\mathcal{A}_{\text{act}}^{A_i}$	Menge der benachbarten Aktionen des i -ten Agenten
b	Bias des Neurons, Biasvektor der Schicht
\mathcal{B}_i	<i>best response</i> Menge des i -ten Agenten
c_{ij}, c_0	lokale Zielfunktion der Produktionsbeschränkungen
d_{loc}	Abstandsschwellwert der lokalen Interpolation
$q_{i,k}^{\text{PS}}$	Zahl der benachbarten Aktionen beim Push-Sum Algorithmus
d_θ	Zahl der Parameter des Actor-KNN
d_ψ	Zahl der Parameter des Critic-KNN
d_{ij}, d_0	lokale Zielfunktion des Durchsatzes
d_e	Abfall der Explorationsrate beim Meta-Lernen
d_α	Abfall der Aktionänderung beim Meta-Lernen

Zeichen	Beschreibung
D_i	Durchsatz des Moduls/Aktuators i
$\mathcal{D}, \mathcal{D}_B$	<i>replay buffer</i>
e	Kante im Produktionsgraph
$e_{\text{current}}, e_{\text{max}}$	aktuelle, max. Zahl an Episoden
E	Erwartungswert
\mathcal{E}	Menge der Kanten des Produktionsgraphs
\mathbf{f}_i	Zustandsfunktion des i -ten Agenten
\mathbf{f}^{PS}	Hilfsvektor des Push-Sum Algorithmus
$g_{1\text{RRL}}$	Funktion im 1-RRL
\mathbf{g}_i	Ausgangsfunktion des i -ten Agenten
$G(\cdot)$	n -Schritt Prädiktion der Belohnungen
$h_{1\text{RRL}}$	Funktion im 1RRL
H_{SPS}	Gütefunktion des SPS-informierten Lernens
H_0	Grenzwert für die Überblendung
i	Laufindex der Agenten
I	Zahl der Schichten des KNN
j	Laufindex
J, J_ψ	durchschnittlicher <i>reward</i> , Gütefunktion der Critic
k	Laufindex der Iterationen
$k_{i,j}$	Steuerungs- und Regelungsparameter der SPS
\mathbf{k}_i	Steuerungsfunktion des i -ten Agenten
k_{max}	max. Zahl an Iterationen
$K_{1\text{RRL}}$	Parameter des 1-RRL
l	Laufindex der Stützvektoren
$L_{1\text{RRL}}$	Parameter des 1-RRL
L	Zahl der Stützvektoren des ZBPS
\mathcal{L}	Menge der Stützvektoren des ZBPS
m	Laufindex des Gedächtnisses
m_i^e	Dimension des Eingangsvektors der externen Eingangsgrößen des i -ten Agenten
$m_{i,c}^e, m_{i,d}^e$	Dimension der kontinuierlichen, diskreten Größen in m_i^e
m_i^s	Dimension des Eingangsvektors der Steuerungsgrößen des i -ten Agenten
$m_{i,c}^s, m_{i,d}^s$	Dimension der kontinuierlichen, diskreten Größen in m_i^s
M	Gedächtnislänge
\mathcal{M}	Gedächtnismenge
\mathcal{M}_u	Updatemenge des Gedächtnisses
n, n_i	Dimension des Gesamtzustandsvektors, des Zustandsvektors des i -ten Agenten
N	Zahl der Lernagenten
$\mathcal{N}(\cdot)$	Normalverteilung
\mathcal{N}_A	Menge der Spieler/Lernagenten/Aktuatoren
o_i	Dimension des Ausgangsvektors des i -ten Agenten
$p(\cdot), p(\cdot \cdot)$	Wahrscheinlichkeit, bedingte Wahrscheinlichkeit
p_0	Eingangswahrscheinlichkeit

Zeichen	Beschreibung
$p_{LLL,i}$	Wahrscheinlichkeiten des LLL
P	Transitionsmodell des MEP
\mathcal{P}	Strategiemenge
Q	Q-Funktion/Zustands-Wertfunktion
R	Belohnungsfunktion
$s_i, \mathbf{s}, \mathbf{s}^{A_i}$	i -ter Zustand, Gesamtzustandsvektor, Zustandsvektor des i -ten Agenten
\mathbf{s}^l	Zustandsvektor des ZBPS im l -ten Stützpunkt
$\mathcal{S}, \mathcal{S}_i$	interne Zustandsmenge aller Agenten, des i -ten Agenten
$\mathcal{S}_e^{A_i}$	lokale externe Zustandsmenge des i -ten Agenten
\mathcal{S}^{A_i}	lokaler Gesamtzustand des i -ten Agenten
\mathcal{S}^g	Menge globaler Zustände in Produktionssystemen
$\mathcal{S}_{\text{prior}}^{A_i}$	Vorgänger Zustandsmenge des i -ten Agenten
$\mathcal{S}_{\text{next}}^{A_i}$	Nachfolger Zustandsmenge des i -ten Agenten
$\mathcal{S}^{\text{shared}}$	geteilte Zustandsmenge zweier Agenten
\mathcal{S}_{SPS}	Zustandsmenge der Schaltfläche der SPS
$U_i (\bar{U}_i)$	Nutzenfunktion des i -ten Agenten (beim SPS-informierten Lernen)
U_{max}^l	bisher bester Nutzenfunktionswert des ZBPS im l -ten Stützvektor
\bar{U}	Mittelwert der Nutzenfunktionswerte beim Lernen mit Gedächtnis
\mathcal{U}	Gleichverteilung
V	Wertefunktion
w^l	Gewichte der Interpolation
W	Sequenz von gleichverteilten Zufallsvariablen
\mathbf{w}, \mathbf{W}	Gewichtsvektor eines Neurons, Gewichtsmatrix der KNN-Schicht
\mathbf{w}^{PS}	Hilfsvektor des Push-Sum Algorithmus
x	Hilfsvariable 1-RRL
\mathbf{x}^{PS}	Hilfsvektor des Push-Sum Algorithmus
y^{PS}	Hilfsvariable des Push-Sum Algorithmus
\mathbf{y}, \mathbf{y}_i	Gesamtausgangsvektor, Ausgangsvektor des i -ten Agenten
z_{BM}	Zufallszahl des Box-Muller Verfahrens
z, \mathbf{z}	Ausgangswert eines Neurons, Ausgangsvektor einer KNN-Schicht
$\alpha_{1\text{RRL}}$	Parameter des 1-RRL
$\alpha_e, \alpha_{mc}, \alpha_{pc}$	ÄR, max. ÄR, rel. ÄR beim ZBPS-ÄR
α_θ	Lernrate im ACRL für den Parameter θ
α_ψ	Lernrate im ACRL für den Parameter ψ
β_1, β_2	Gewichtungsfaktoren des SPS-informierten Lernens
$\beta_{1\text{RRL}}$	Parameter des 1-RRL
γ	Diskontierungsfaktor
$\gamma_{1\text{RRL}}$	Parameter des 1-RRL
γ_{IP}	Parameter der Interpolation
γ^{PS}	Sequenz von Zeitschritten beim Push-Sum Algorithmus
Γ	Spiel
$\delta_{\bar{U}}$	Parameter des ZBPS-ÄR
δ_{MS}	Parameter des ZBPS-MS
ϵ	Explorationswahrscheinlichkeit des ϵ -greedy Algorithmus

Zeichen	Beschreibung
ϵ_2	Parameter für Exploration beim Meta-Lernen
ϵ^{PS}	Sequenz von unabhängigen und gleichverteilten Zufallsvektoren
ζ	Schaltparameter der SPS
θ, θ_z	Parameter des Actor-KNN, des Ziel-KNN
l	Schicht des KNN
κ, κ_0	Parameter des ZBPS-MS
λ_{1RRL}	Parameter des 1-RRL
μ	Mittelwert der Verteilung
μ_{1RRL}	Mittelwert beim 1-RRL
μ^l	Mittelwert im l -ten Stützvektor des ZBPS-MS
ξ	Aktivierungsfunktion des Neurons
π	Policy/Steuerungsstrategie
ρ	Belohnung
ρ_{ABR}	Abklingrate für die Explorationswahrscheinlichkeit
ρ_{exp}	Explorationswahrscheinlichkeit bei der Interpolation
$\rho_z, \theta, \rho_z, \psi$	Faktor gleitender Mittelwert Ziel-KNN
σ	Standardabweichung der Normalverteilung
σ_{1RRL}	Standardabweichung beim 1-RRL
σ^l	Standardabweichung im l -ten Stützvektor des ZBPS-MS
τ	Temperatur beim LLL
ϕ	globale Nutzenfunktion, Potentialfunktion
χ	Parameter des ZBPS-ÄR
ψ, ψ_z	Parameter des Critic-KNN, des Ziel-KNN

Formelzeichen Schüttgutanlage

Zeichen	Einheit	Beschreibung
A^*	[mm ²]	mittlerer Flächeninhalt des Behälters/Trichters *
C_i	[s]	Produktionsbeschränkungen des Moduls/Aktuators i
D	[l/s]	momentane Abweichung beim Durchsatz des letzten Moduls
E_i	[kW]	Energieverbrauch des Moduls/Aktuators i
E_{max}	[kW]	max. Energieverbrauch
f_{norm}	[-]	normalisierte Frequenz [0/1] des Vibrationsförderers
$f_{\text{vak}}^*(t_{\text{akt},*})$	[g/s]	Massenstrom der Vakuumpumpe *
f_{vf}	[1/s]	Frequenz des Vibrationsförderers
g^*	[kW]	Energieverbrauchsfunktion des Aktuators *
h_i, h^*	[mm]	Füllstand des i -ten Behälters/Trichters *
$h_{\text{lower},i}^{\text{next}}, h_{\text{upper},i}^{\text{next}}$	[mm]	Schwellwerte für Füllstände
$h_i^{\cdot,*}$	[-]	Hilfsgrößen zur Berechnung der Überläufe
$h_{i,\text{max}}$	[mm]	max. Füllstand des i -ten Behälters/Trichters
$h_{i,c}, h_w$	[mm]	Hilfsgrößen zur Initialisierung der Füllstände der Behälter
$\dot{m}_{\text{in},*}$	[g/s]	Massenstrom in den Behälter/Trichter *
$\dot{m}_{\text{out},*}$	[g/s]	Massenstrom aus dem Behälter/Trichter *
\dot{m}_{out}	[g/s]	Ausgangsmassenstrom der Anlage

Zeichen	Einheit	Beschreibung
n_{fb}	[1/min]	Drehzahl des Förderbands
n_{rot}	[1/min]	Drehzahl der Zellenradschleuse
$r_{C,i}$	[-]	Überlauf-Belohnung beim Meta-Lernen
$r_{E,i}$	[-]	Energie-Belohnung beim Meta-Lernen
$r_{T,i}$	[-]	Transport-Belohnung beim Meta-Lernen
$t_{start,*}$	[s]	Startzeit der Vakuumpumpe *
$t_{akt,*}$	[s]	Aktivierungszeit der Vakuumpumpe *
$t_{stop,*}$	[s]	Stoppzeit der Vakuumpumpe *
T	[s]	Zeithorizont
T_e	[s]	Episodenlänge
T_I	[s]	Iterationslänge
T_{max}	[s]	max. Trainingszeit
V_B	[l]	vorgegebener Bedarf an geförderter Menge in einer Iteration
\dot{V}_B	[l/s]	vorgegebener Bedarf an gefördertem Durchsatz
\dot{V}_{out}	[l/s]	Ausgangsvolumenstrom des letzten Moduls
\dot{V}_i	[l/s]	Volumenstrom des i -ten Behälters/Trichters
α_1	[1/s]	Gewichtungsparameter des Überlaufs
α_2	[1/l]	Gewichtungsparameter des Bedarfs
α_3	[1/kWs]	Gewichtungsparameter des Energieverbrauchs
α_{fb}	[min · g/s]	Massenstromkoeffizient des Förderbands
α_{vf}	[g]	Massenstromkoeffizient des Vibrationsförderers
α_{rot}	[min · g/s]	Massenstromkoeffizient der Zellenradschleuse
η_C	[1/s]	Gewichtsfaktor für Überlauf beim Meta-Lernen
η_E	[1/kWs]	Gewichtsfaktor für Energieverbrauch beim Meta-Lernen
η_T	[1/l]	Gewichtsfaktor für transportierte Masse beim Meta-Lernen
ρ_S	[g/l]	Dichte des Schüttguts