

**Berücksichtigung einer Spannungsanpassung als Handlungsoption im
intelligenten Netzbetrieb**

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

Leschek Kopczynski
aus
Marienburg (Polen)

Gutachter: Prof. Dr.-Ing. Holger Hirsch

Gutachter: Prof. Dr.-Ing. Roland Zeise

Tag der mündlichen Prüfung: 20.02.2024

Vorwort

Der vorliegenden Arbeit liegen die Erfahrungen zugrunde, die ich als Mitarbeiter in der Forschungsgruppe Smart Grid und Virtuelle Kraftwerke des Fachbereichs Elektro- und Informationstechnik an der Hochschule Düsseldorf sammeln durfte. Neben den tiefgreifenden Diskussionen und der guten Arbeitsatmosphäre, welche durch große Hilfsbereitschaft und Freude geprägt war, ist dieser Zeit die Leidenschaft für das Programmieren entsprungen, die einen großen Part in der vorliegenden Arbeit eingenommen hat.

Bei der Realisierung dieser Arbeit haben mich viele Menschen unterstützt, bei denen ich mich an dieser Stelle bedanken möchte. Herrn Prof. Dr.-Ing. Holger Hirsch danke ich auf das Herzlichste für die Bereitschaft mich als Doktoranden aufzunehmen und mir bei der Themenwahl große Freiheiten einzuräumen. Der regelmäßige Austausch, die vielen wertvollen Hinweise und Anregungen sowie der zugesprochene Mut, mein Vorhaben voranzutreiben, haben den erfolgreichen Abschluss der Promotion erst möglich gemacht.

Zu besonderem Dank bin ich Herrn Prof. Dr.-Ing. Roland Zeise nicht nur für die Übernahme des Korreferats verpflichtet, vielmehr bedanke ich mich für die Förderung meiner wissenschaftlichen Ausbildung durch seine persönliche und fachliche Führung sowie sein stetes Interesse an meiner Arbeit.

Den Mitgliedern des Dekanats des Fachbereichs Elektro- und Informationstechnik der Hochschule Düsseldorf, dem ehemaligen Dekan Herrn Prof. Dr.-Ing. Detmar Arlt sowie dem aktuellen Dekan Prof. Dr. rer. nat. Thomas Licht danke ich für die stete Unterstützung in allen Belangen über die vergangenen Jahre.

Meinen ehemaligen Kollegen Michael Schallenburg, Philipp Huppertz und Christian Alteheld danke ich für die Freundschaft, die vielen konstruktiven Diskussionen und die gegenseitige Unterstützung. Darüber hinaus danke ich den Studierenden, die durch ihre Abschlussarbeiten, durch gemeinsame Veröffentlichungen und Datenauswertung zum Erfolg meiner Arbeit beigetragen haben, für ihre Unterstützung.

Bei meinen Eltern Malgorzata und Andreas, meinen Brüdern Bartholomäus und Bogumil sowie allen meinen Freunden bedanke ich mich für ihre unermüdliche moralische Unterstützung sowie ausdauernde Geduld während meines Studiums und der Zeit als Doktorand. Zu guter Letzt danke ich meiner Frau Anna für ihre bedingungslose, liebevolle Unterstützung und motivierenden Worte, wann immer es nötig war.

Düsseldorf, im August 2023
Leschek Kopczynski

Kurzfassung

Ziel der vorliegenden Arbeit ist die Bewertung einer aktiv herbeigeführten Spannungsanpassung als Handlungsoption im intelligenten Netzbetrieb. Hierbei wird unter Anwendung von Algorithmen des maschinellen Lernens eine Weitbereichsregelung im elektr. Niederspannungsnetz adressiert. Im Fokus der Untersuchung steht die aktive Spannungshaltung mit Hilfe eines regelbaren Ortsnetztransformators unter der praxisnahen Prämisse, die Schalthäufigkeit des Stufenschalters gering zu halten. Es wird eine Methodik vorgestellt, die eine Ableitung von Netzführungsstrategien mit Bezug auf das Energieversorgungssystem ermöglicht. Ausgehend von einer Netzsimulation werden bestehende Stamm- und Bewegungsdaten explizit im Zuge der Ableitung der Netzführungsstrategien berücksichtigt. Zudem sieht der simulative Ansatz die Definition unterschiedlicher Szenarien vor, wodurch exogene sowie endogene Unsicherheiten berücksichtigt werden. Dies resultiert in der Ableitung einer robusten Kontrollstrategie.

Der vorgestellte Ansatz sieht den Einsatz einer Simulationsumgebung vor, welche die Prozesse im elektr. Niederspannungsnetz realitätsnah wiedergibt. Hierbei steht die Modellierung spannungsabhängiger Last- und Einspeiseprozesse im Vordergrund. Es werden unterschiedliche Ansätze vorgestellt und in klassische sowie probabilistische Verfahren gruppiert. Insbesondere die Modellierung des individuellen stochastischen Verhaltens einzelner Haushalts-Endkunden mit Hilfe eines probabilistischen Ansatzes wird hervorgehoben. Die erarbeiteten Modelle werden mit Hilfe aufgenommener Messdaten validiert.

Ausgehend von einer detaillierten Modellierung spannungsabhängiger Last- und Einspeiseprozesse wird explizit der Einfluss eines energiekonstanten Verbraucherverhaltens berücksichtigt. Da dieses Verhalten und der hieraus resultierende Kompensationseffekt, der mit fortschreitender Zeit der Spannungsänderung entgegen wirkt, häufig vernachlässigt wird, wird die Notwendigkeit der Berücksichtigung des Kompensationseffekts mit Hilfe von zwei umfangreichen Szenarien belegt. Hierfür werden im Zuge einzelner Optimierungsprobleme die optimalen Schaltzeitpunkte eines regelbaren Ortsnetztransformators für ein passives sowie ein aktives elektr. Niederspannungsnetz bewertet. Ein netzdienlicher Effekt auf die Spitzenlast, die Energieaufnahme und die Verlustenergie untermauern einen signifikanten Einfluss trotz der kompensierenden Wirkung des Kompensationseffekts.

Die vorliegende Arbeit verfolgt einen praxisorientierten Ansatz. Zur Verifizierung und Validierung wird auf reale Stamm- und Bewegungsdaten einzelner elektr. Niederspannungsnetze, aufgenommen im Forschungsvorhaben *Erfassung der niederspannungsseitigen Netzzustandsgrößen in Echtzeit* (ENERGIE), zurückgegriffen.

Abstract

The aim of the present thesis is the evaluation of an actively caused voltage adjustment as an option for action in smart grid operation. For this purpose, a conservation voltage reduction approach in a low-voltage grid is addressed using machine learning algorithms. The focus of the study is on active voltage control with the help of a regulated distribution transformer while reducing the number of tap-operations. The presented method enables the derivation of grid management strategies with reference to the energy supply system. Based on a power system simulation, existing topology and measurement data are explicitly taken into account while deriving grid management strategies. In addition, the simulative approach considers different scenarios so that exogenous and endogenous uncertainties can be considered. This results in the derivation of a robust control strategy.

The presented method uses a simulation environment which reflects the real system conditions in low-voltage grids accurately. Here, the focus is on modelling voltage-dependent load and power injection. Different approaches are presented and grouped into classical as well as probabilistic methods. Especially the modelling of the individual stochastic behaviour of single households using a probabilistic approach is highlighted. The developed models are validated with the help of measurement data.

Based on a detailed modelling of voltage-dependent load and power injection processes, appliances with constant energy behaviour are explicitly taken into account. Since this behaviour and the associated duty cycle rebound effect, which compensates the consequences of the voltage change over time, are often neglected, the need to take the duty cycle rebound effect into account is proven with the help of two extensive studies. For this purpose, optimal switching time tuples of a regulated distribution transformer for a passive as well as an active low-voltage grid are evaluated within individual optimisation tasks. A grid-serving effect on the peak load, energy demand and system loss underpin a significant influence despite the compensating impact of the duty cycle rebound effect.

The present thesis follows a practice-oriented approach. Real topology and measurement data of individual low-voltage grids, which are based on the research project *Acquisition of Low-Voltage Grid States in Real Time* (ENERGIE), are used for verification and validation.

Inhaltsverzeichnis

Verzeichnis verwendeter Abkürzungen und Formelzeichen	III
1. Einleitung	1
1.1. Stand der Technik	2
1.2. Motivation und Ziele der Arbeit	6
2. Modellierung spannungsabhängiger Last- und Einspeiseprozesse	13
2.1. Verbraucherlastmodelle	14
2.1.1. Status quo	15
2.1.1.1. Extremszenarien	15
2.1.1.2. Repräsentative Lastprofile	16
2.1.2. Probabilistischer Ansatz	20
2.1.2.1. <i>Top-Down</i> -Verfahren	22
2.1.2.2. <i>Bottom-Up</i> -Verfahren	26
2.2. Einspeisemodell einer Photovoltaikanlage	39
2.2.1. Wirkungskette des Energiewandlungsprozesses	39
2.2.2. Modellierung der Bestrahlungsstärke auf geneigter Fläche	40
2.2.3. Exogene Unsicherheiten	42
2.2.4. Generierung synthetischer Einspeiseprofile	44
2.3. Modellierung der Spannungsabhängigkeit	48
2.3.1. ZIP-Modell	49
2.3.2. ZIP(E)-Modell zur Berücksichtigung der Energiekonstanz	51
3. Spannungsanpassung als Flexibilitätsoption im elektr. Verteilnetz	57
3.1. Netzsimulation spannungsabhängiger Prozesse	59
3.1.1. Definition der Netzsimulation	59
3.1.2. Schaltzeittpel und Zielfunktionen	60
3.1.3. Verifizierung der Modellannahmen	61
3.2. Untersuchungsszenarien	63
3.2.1. Szenario 1: passives Netz	63
3.2.2. Szenario 2: aktives Netz	67
3.3. Resümee	74
4. Bestärkendes Lernen für die integrierte Energiewende	77
4.1. Grundlagen des bestärkenden Lernens	80
4.1.1. Konzept und Rahmenstruktur	81

4.1.2.	Signale und Zieldefinition	82
4.1.2.1.	Systemzustand	82
4.1.2.2.	Aktion und Aktionsraum	83
4.1.2.3.	Belohnungssignal	83
4.1.2.4.	Zieldefinition des Rahmenwerks	83
4.1.3.	Markow-Entscheidungsprozess	84
4.1.4.	Aufbau der Agenten	85
4.1.4.1.	Optimale Kontrollstrategie	86
4.1.4.2.	Verfahren	89
4.2.	Rahmenwerk für den Einsatz im elektr. Energieversorgungssystem	104
4.2.1.	Lernmodul	104
4.2.2.	Ausführendes Modul	105
4.3.	Weitbereichsregelung im elektr. Niederspannungsnetz	106
4.3.1.	Signale des Markow-Entscheidungsprozesses	106
4.3.2.	Ableiten einer Kontrollstrategie innerhalb des Lernmoduls	109
4.3.3.	Verifizierung der Kontrollstrategie	112
4.3.4.	Resümee	115
5.	Zusammenfassung und Ausblick	117
	Anhang	123
A.	Knotenmodellierung	123
A.1.	Einfluss des Gleichzeitigkeitsgrades auf die anzunehmende Spitzenlast von Haushalts-Endkunden	123
A.2.	Darstellung des Standardlastprofils nach Profiltyp H0	123
A.3.	Definition der Standardabweichung des Top-Down-Verfahrens	124
A.4.	Algorithmus zur Lastmodellierung mittels Top-Down-Verfahren	126
A.5.	Algorithmus zur Lastmodellierung mittels Bottom-Up-Verfahren	127
A.6.	Netztopologie zur Validierung des Bottom-Up-Verfahren	128
A.7.	Niederspannungsnetz mit hoher Durchdringung an Photovoltaikanlagen	129
B.	Komponenten des bestärkenden Lernens	130
B.1.	Monte-Carlo-Vorhersage beim erstmaligen Besuch	130
B.2.	Optimierungsansätze der Monte-Carlo-Methoden	131
B.3.	Algorithmus des <i>Deep Q-Network</i>	132
B.4.	Algorithmus des <i>Deep Deterministic Policy Gradient</i>	133
	Publikationen des Autors	135
	Literatur	137

Verzeichnis verwendeter Abkürzungen und Formelzeichen

Nachfolgend sind die innerhalb dieser Arbeit verwendeten Abkürzungen und Formelzeichen in alphabetischer Reihenfolge aufgeführt.

Abkürzungen

ADAM	adaptiver Momentumschätzer (aus dem engl. adaptive moment estimation)
AI	künstliche Intelligenz (aus dem engl. artificial intelligence)
VDE-AR	Anwendungsregeln des Verbands der Elektrotechnik Elektronik Informationstechnik e. V.
BEWAG	Berliner Städtische Elektrizitätswerke Aktien Gesellschaft
BHKW	Blockheizkraftwerk
BUV	Verbraucherlastmodell nach dem Bottom-Up-Verfahren
BNetzA	Bundesnetzagentur
CVR	conservation voltage reduction
DCR	Kompensationseffekt (aus dem engl. duty cycle rebound)
DPG	deterministischer Policy-Gradient (aus dem engl. deterministic policy gradient)
DDPG	tiefer deterministischer Policy-Gradient (aus dem engl. deep deterministic policy gradient)
DEA	dezentrale Erzeugungsanlage
DNN	tiefe künstliche neuronale Netze (aus dem engl. deep neural networks)
DQN	tiefes Q-Netzwerk (aus dem engl. deep Q-network)
DRL	tiefes bestärkendes Lernen (aus dem engl. deep reinforcement learning)
DOU	Nutzungsdauer (aus dem engl. duration of use)
EE	erneuerbare Energie
EEG	Erneuerbare-Energien-Gesetz
ENERGIE	Erfassung der niederspannungsseitigen Netzzustandsgrößen in Echtzeit
ERB	Erfahrungs-Wiedergabepuffer (aus dem engl. experience replay buffer)
EVK	Evaluierung-Verbesserung-Kreislauf
EVS	Energieversorgung Südsachsen

EVU	Energieversorgungsunternehmen
G	gewerbliche Betriebe
GPI	generalisierte Iteration der Lösungsstrategie (aus dem engl. generalized policy iteration)
IKT	Informations- und Kommunikationstechnik
iMSys	intelligente Messsysteme oder digitale Stromzähler
JEB	Jahresenergiebedarf
H	Haushaltskunden
H0	Haushalt, Privatverbrauch, ggf. geringfügig gewerblicher Bedarf
HAK	Hausanschlusskasten
KI	künstliche Intelligenz
KNN	künstliche neuronale Netze
KSG	Klimaschutzgesetz
KWK	Kraft-Wärme-Kopplung
L	landwirtschaftliche Betriebe
Mio.	Million
ML	maschinelles Lernen (aus dem engl. machine learning)
Mrd.	Milliarde
MRP	Markov-Belohnungsprozess (aus dem engl. Markov reward process)
MSBE	mittlere quadratische Bellman-Abweichung (aus dem engl. mean-squared Bellman error)
MsbG	Messstellenbetriebsgesetz
ONS	Ortsnetzstation
OSE	Oder-Spree Energieversorgung
PG	Policy-Gradient (aus dem engl. policy gradient)
PRNG	Pseudozufallszahlengenerator
PSO	Partikelschwarmoptimierung
PV	Photovoltaik
PVA	Photovoltaikanlage
Ref	Referenzwert
ReLU	gleichgerichtete lineare Einheit (aus dem engl. rectified linear unit)
RL	bestärkendes Lernen (aus dem engl. reinforcement learning)
RLM	registrierte Leistungsmessung
rONT	regelbarer Ortsnetztransformator
RWE	Rheinisch-Westfälisches Elektrizitätswerk
SDLWindV	Verordnung zu Systemdienstleistungen durch Windenergieanlagen
SLP	Standardlastprofil

STC	Standardtestbedingungen (aus dem engl. standard test conditions)
SW	Stadtwerke
TAB	technische Anschlussbedingungen Niederspannung
TAR	technische Anschlussregeln
TD	temporale Differenz (aus dem engl. temporal difference)
TDV	Verbraucherlastmodell nach dem Top-Down-Verfahren
TN	Zielnetzwerk (aus dem engl. target network)
TOU	Nutzungszeitpunkt (aus dem engl. time of use)
VDE	Verband der Elektrotechnik Elektronik Informationstechnik e. V.
VDEW	Verband der Elektrizitätswirtschaft e. V.
VEW	Vereinigte Elektrizitätswerke Westfalen
VNB	Verteilnetzbetreiber
WDF	Wahrscheinlichkeitsdichtefunktion
WKA	Windkraftanlagen
WTS	Waschmaschine, Trockner und Spülmaschine
ZV	Zufallsvariable

Formelzeichen

$\mathcal{A}, \mathcal{A}(s)$	Aktionsraum im Systemzustand s
A	Zufallsvariable der Aktionen
A_{PVA}	Fläche der verbauten Module
A_t, A_{t+1}	Zufallsvariable der Aktion zum Zeitpunkt t bzw. $t + 1$
a, a_i, a_t	Aktionen zum Zeitpunkt t
a', a_{t+1}	Aktionen zum Zeitpunkt $t + 1$
B	Anzahl der gezogenen Übergänge in Form von Tupeln
b	Formparameter der Belohnungsfunktion
C	Schritte zwischen der periodischen Aktualisierung des Zielnetzwerks
CVR_f	Faktor zur Bewertung einer Spannungsanpassung
$CVR_{f,E_{\max}}$	Faktor zur Bewertung einer Spannungsanpassung mit Bezug auf die max. Energieaufnahme
$CVR_{f,E_{\text{Verluste}}}$	Faktor zur Bewertung einer Spannungsanpassung mit Bezug auf die Gesamtverlustenergie
$CVR_{f,P_{\max}}$	Faktor zur Bewertung einer Spannungsanpassung mit Bezug auf die Spitzenlastreduktion
c	Proportionalitätskonstante zur Beschreibung der Montageart
c_1, c_2	konstante Beschleunigungskoeffizienten
$\cos(\varphi)$	Leistungsfaktor
$\cos(\varphi)_{gg,t}$	Leistungsfaktor einer Gerätegruppe
$\cos(\varphi)_{\text{res}}, \cos(\varphi)_{\text{res},t}$	resultierender Leistungsfaktor

\mathcal{D}	finiter Pufferspeicher
d	Tag des jeweiligen Jahres
\mathcal{E}	Umgebung
\mathcal{E}_{real}	reales System
\mathcal{E}_{sim}	Simulationsumgebung
E	elektr. Energie
$E(\mathcal{P}_t), E(X)$	Erwartungswert einer Zufallsvariable
E_{diff}	Anteil der diffusen Bestrahlungsstärke auf geneigter Fläche
E_{dir}	Anteil der direkten Bestrahlungsstärke auf geneigter Fläche
E_g	Bestrahlungsstärke auf geneigter Fläche
E_{max}	Gesamtenergieaufnahme
$E_{max}((t_{SH_1}, t_{SH_2}))$	Gesamtenergieaufnahme beim Schaltzeitpunkt (t_{SH_1}, t_{SH_2})
E_{ref}	Anteil der vom Boden reflektierten Bestrahlungsstärke auf geneigter Fläche
$E_{Verluste}$	Gesamtverlustenergie
$E_{Verluste}((t_{SH_1}, t_{SH_2}))$	Gesamtverlustenergie beim Schaltzeitpunkt (t_{SH_1}, t_{SH_2})
e_t	Beobachtung in Form eines Tupels zum Zeitpunkt t
f	Funktion
f_{BG}	Faktor zur Berücksichtigung des Bedeckungsgrades
f_m	skalierte Steigung
$f_{r_n,t}$	Funktion zur Berücksichtigung der Schalthäufigkeit
$f_{\Delta U_{SH}}$	Funktion einer relativen Spannungsanpassung
\mathcal{G}	Gamma-verteilte Wahrscheinlichkeitsdichtefunktion
G, G_t	Gewinn zum Zeitpunkt t
g, g_∞	Gleichzeitigkeitsgrad
$g_{i,j}$	globaler Bestwert des i -ten Partikels in Richtung der Komponente j für Iterationsschritt n
g_{PVA}	Gleichzeitigkeitsgrad für die Berücksichtigung von Photovoltaikanlagen
gg	Gerätegruppe
h	Faktor zur Skalierung der Steigung
$I_P, I_{P,Q}$	Koeffizient des ZIP-Modells zur Beschreibung des stromkonstanten Anteils der Wirkleistung P sowie Blindleistung Q .
$I_{P,gg}$	Koeffizient des ZIP-Modells zur Beschreibung des stromkonstanten Anteils der Wirkleistung P einer Gerätegruppe
$I_{P,res}$	resultierender Koeffizient des ZIP-Modells zur Beschreibung des stromkonstanten Anteils der Wirkleistung P
I_Q	Koeffizient des ZIP-Modells zur Beschreibung des stromkonstanten Anteils der Blindleistung Q
$I_{Q,gg}$	Koeffizient des ZIP-Modells zur Beschreibung des stromkonstanten Anteils der Blindleistung Q einer Gerätegruppe

$I_{Q,res}$	resultierender Koeffizient des ZIP-Modells zur Beschreibung des stromkonstanten Anteils der Blindleistung Q
i	laufender Index
\mathcal{J}	Zielfunktion
$\mathcal{J}(\theta^\mu)$	Zielfunktion zur Aktualisierung des Parametervektors θ^μ
\mathcal{K}_t	Zeitraum mit diskreten Zeitpunkten zum Zeitpunkt t
K	diskreter Zeitpunkt
k	diskreter Zeitpunkt, Netzknoten
\mathcal{L}	Zielfunktion
$\mathcal{L}(\theta^Q)$	Zielfunktion zur Aktualisierung des Parametervektors θ^Q
L	Anzahl von Zuständen
l	Länge, Leiter
M	Speichergröße des finiten Pufferspeichers \mathcal{D}
m	Gesamtzahl der Partikel
$m_{n,t}$	Steigung der Spannung im definierten Intervall am Netzknoten n zum Zeitpunkt t
\mathcal{N}	Gauß-verteilte Wahrscheinlichkeitsdichtefunktion
N	Anzahl von Haushalts-Endkunden, Summe der gemessenen Knotenpunkte
$N(S_t, A_t)$	Anzahl der Besuche eines Zustand-Aktion-Tupels
N_A	Anzahl der Aktionen
n	Iterationsschritt
\mathcal{P}	Übergangsmatrix
\mathcal{P}_{gg}	Zufallsvariable der Wirkleistung einer Gerätegruppe
$\mathcal{P}_t, \mathcal{P}_{t,ges}$	Zufallsvariable der Wirkleistung zum Zeitpunkt t
P	Wirkleistung
P_{el}	elektr. Einspeiseleistung
\bar{P}_{gg}	mittlere Wirkleistungsaufnahme der Gerätegruppe
$P_{gg}, P_{gg,t}$	Wirkleistungsaufnahme einer Gerätegruppe
\bar{P}_{gg,v_i}	mittlere Wirkleistungsaufnahme eines Verbrauchers der Gerätegruppe
$P_{gg,l}$	Wirkleistungsaufnahme einer Gerätegruppe auf Außenleiter l
$P_{gg, \text{Zeitbereich}}$	Wirkleistungsaufnahme einer Gerätegruppe im definierten Zeitintervall <i>Zeitbereich</i>
P_H	Leistungsbezug eines Haushalts
P_{max}	Leistungsspitze
$P_{max}((t_{SH_1}, t_{SH_2}))$	Leistungsspitze beim Schaltzeitpunkt (t_{SH_1}, t_{SH_2})
P_{min}	Leistungsminimum
P_{ONS}	kumulierte Wirkleistung an der Ortsnetzstation
$P_P, P_{P,Q}$	Koeffizient des ZIP-Modells zur Beschreibung des leistungs-konstanten Anteils der Wirkleistung P sowie Blindleistung Q .

$P_{P,gg}$	Koeffizient des ZIP-Modells zur Beschreibung des leistungs-konstanten Anteils der Wirkleistung P einer Gerätegruppe
$P_{P,res}$	resultierender Koeffizient des ZIP-Modells zur Beschreibung des leistungs-konstanten Anteils der Wirkleistung P
P_{Peak}	max. Einspeiseleistung
P_{PVA}	Gesamtleistung der installierten Photovoltaikanlagen
P_Q	Koeffizient des ZIP-Modells zur Beschreibung des leistungs-konstanten Anteils der Blindleistung Q
$P_{Q,gg}$	Koeffizient des ZIP-Modells zur Beschreibung des leistungs-konstanten Anteils der Blindleistung Q einer Gerätegruppe
$P_{Q,res}$	resultierender Koeffizient des ZIP-Modells zur Beschreibung des leistungs-konstanten Anteils der Blindleistung Q
P_r	Wahrscheinlichkeit
P_r	Bemessungswirkleistung
P_{res}	resultierender Wirkleistungsbedarf
$P_{U>U_r}$	resultierender Wirkleistungsbedarf bei erhöhter elektr. Span-nung U
$P_{U=U_r}$	resultierender Wirkleistungsbedarf bei Bemessungsspannung U_0
$P_{U<U_r}$	resultierender Wirkleistungsbedarf bei reduzierter elektr. Spannung U
$p_{i,j}(n)$	Bestwert des i -ten Partikels in Richtung der Komponente j für Iterationsschritt n
$p_{s,max}$	Spitzenlast eines Haushalts-Endkunden (vollelektrifiziert, oh-ne Elektroheizung)
$p_s(N)$	Maximallast von N -Haushalts-Endkunden unter Berücksich-tigung der Gleichzeitigkeitsfaktoren
$p_{s,s'}$	Übergangswahrscheinlichkeit von Zustand s in s'
$p(s', r s, a)$	Einzelschritt-Dynamik eines Markow-Entscheidungsprozesses
Q	Blindleistung
Q_{gg}	Blindleistungsbedarf einer Gerätegruppe
Q_{opt}	optimale Wertfunktion des Zustand-Aktion-Tupels
Q_r	Bemessungsblindleistung
Q_{res}	resultierender Blindleistungsbedarf
Q_π	Wertfunktion des Zustand-Aktion-Tupels für Kontrollstrate-gie π
\mathcal{R}	Belohnungsfunktion
R	skalare Belohnung
R_t, R_{t+1}	skalare Belohnung zum Zeitpunkt t bzw. $t + 1$
r, r_i, r_t	skalare Belohnung
r_1, r_2	Zufallszahlen einer stetigen Gleichverteilung
\mathcal{S}	Zustandsraum

S^+	Zustandsraum, inklusive terminalen Systemzustand
S	Zufallsvariable der Zustände
S_t, S_{t+1}	Zufallsvariable der Zustände zum Zeitpunkt t bzw. $t + 1$
$S_{r,\text{rONT}}$	Nennscheinleistung des regelbaren Ortsnetztransformators
S_T	terminaler Systemzustand
s, s_i, s_t	Zustandsvektor zum Zeitpunkt t
s', s_{i+1}, s_{t+1}	Zustandsvektor zum Zeitpunkt $t + 1$
(s, a)	Zustand-Aktion-Tupel
\mathcal{T}	Menge von diskreten Zeitpunkten
$\mathcal{T}_{gg,\text{DOU}}$	stetige Wahrscheinlichkeitsdichtefunktion der Nutzungsdauer einer Gerätegruppe
\mathcal{T}_{TOU}	stetige Wahrscheinlichkeitsdichtefunktion des Nutzungszeitpunktes
$\mathcal{T}_{\text{Zeitbereich}}$	Menge von diskreten Zeitpunkten im definierten <i>Zeitbereich</i>
T	diskreter Zeitpunkt
$T_{\text{Kühleinheit}}$	Zykluszeit der Gerätegruppe Kühleinheit
$T_{\text{aplimit},\pm}$	positive und negative Grenzposition des Stufenschalters
$T_{\text{ap}t}$	Position des Stufenschalters zum Zeitpunkt t
t	diskreter Zeitpunkt, Uhrzeit
$t_{\text{DOU}}, t_{gg,\text{DOU}}$	Nutzungsdauer einer Gerätegruppe
$t_{gg,\text{Zeitbereich}}$	Nutzungszeitpunkt der Gerätegruppe im definierten <i>Zeitbereich</i>
t_{SH}	Zeitpunkt einer Schalthandlung
$(t_{\text{SH}_1}, t_{\text{SH}_2})$	Schaltzeittupel
$(t_{\text{SH}_1}, t_{\text{SH}_2})_{\text{opt}}$	optimales Schaltzeittupel
$(t_{\text{SH}_1}, t_{\text{SH}_2})_{\text{opt},\text{Zeitbereich}}$	optimales Schaltzeittupel innerhalb eines <i>Zeitbereichs</i>
$(t_{\text{SH}_1}, t_{\text{SH}_2})_{\text{Zeitbereich}}$	Schaltzeittupel innerhalb eines <i>Zeitbereichs</i>
t_{TOU}	Nutzungszeitpunkt
\mathcal{U}	gleichverteilte Wahrscheinlichkeitsdichtefunktion
U	elektr. Spannung
U_0	Nennspannung
U_r	Bemessungsspannung
v	Verbraucher
$v_{i,j}(n), v_{i,j}(n + 1)$	Geschwindigkeit des i -ten Partikels in Richtung der Komponente j für Iterationsschritt n respektive $n + 1$
V_{opt}	optimale Wertfunktion des Zustandes
$V_\pi, V_{\pi'}$	Wertfunktion des Zustandes für Kontrollstrategie π
$\mathcal{W}, \mathcal{W}_t$	Rauschprozess
W	elektr. Arbeit
$W_{gg,\text{Tag}}$	Mittlerer Energiebedarf einer Gerätegruppe eines Tages
w	Trägheitskoeffizienten
\mathcal{X}	Zufallsvariable

X	Stapel zufällig gezogener Übergänge
x	korrigierter Profilwert zur Berücksichtigung der Dynamisierung
x_0	Viertelstundenwert des Profils
\mathbf{x}_i	Positionsvektor des i -ten Partikels
$x_{i,j}$	Position des i -ten Partikels in Richtung der Komponente j
$x_{i,j}(n), x_{i,j}(n+1)$	Position des i -ten Partikels in Richtung der Komponente j für Iterationsschritt n respektive $n+1$
$Z_P, Z_{P,Q}$	Koeffizient des ZIP-Modells zur Beschreibung des impedanzkonstanten Anteils der Wirkleistung P sowie Blindleistung Q .
$Z_{P,gg}$	Koeffizient des ZIP-Modells zur Beschreibung des impedanzkonstanten Anteils der Wirkleistung P einer Gerätegruppe
$Z_{P,res}$	resultierender Koeffizient des ZIP-Modells zur Beschreibung des impedanzkonstanten Anteils der Wirkleistung P
Z_Q	Koeffizient des ZIP-Modells zur Beschreibung des impedanzkonstanten Anteils der Blindleistung Q
$Z_{Q,gg}$	Koeffizient des ZIP-Modells zur Beschreibung des impedanzkonstanten Anteils der Blindleistung Q einer Gerätegruppe
$Z_{Q,res}$	resultierender Koeffizient des ZIP-Modells zur Beschreibung des impedanzkonstanten Anteils der Blindleistung Q
<i>Zeitbereich</i>	diskrete Menge von definierten Zeitintervallen innerhalb eines Tages
$\nabla_{\theta^Q} \mathcal{L}_i(\theta^Q)$	Gradient der Zielfunktion $\mathcal{L}_i(\theta^Q)$
$\nabla_{\theta^\mu} \mathcal{J}(\theta^\mu)$	Gradient der Zielfunktion $\mathcal{J}(\theta^\mu)$
$\alpha(t)$	Formparameter einer zeitvarianten Gamma-verteilten Wahrscheinlichkeitsdichtefunktion
α	Hyperparameter zur Definition der Schrittweite
α_F	Ausrichtung der Modulfläche
α_P	Temperaturkoeffizient
α_S	Sonnenazimut
$\beta(t)$	Skalierungsparameter einer zeitvarianten Gamma-verteilten Wahrscheinlichkeitsdichtefunktion
Γ	Gamma-Funktion
γ	Diskontierungsfaktor
γ_F	Neigung der Modulfläche
γ_S	Sonnenhöhe
δ	Fehler der temporalen Differenz
$\Delta E, \Delta E_{\max}$	(max.) Veränderung der Energieaufnahme
$\Delta E_{\text{Verluste}}$	Veränderung der Gesamtverlustenergie
$\Delta P, \Delta P_{\max}$	(max.) Veränderung der elektr. Leistung
$\Delta t, \Delta t_H$	Zeitintervall
Δt_A	Raster der zeitlichen Positionierung einzelner Schaltzeitpunkte

$\Delta t_{\text{Kühleinheit,ein,rel}}$	relative Einschaltdauer der Gerätegruppe <i>Kühleinheit</i>
$\Delta t_{\text{ZIP(E)}}$	Nutzungsdauer unter Berücksichtigung des ZIP(E)-Modells
$\pm \Delta tap_t$	Wechsel der Position des Stufenschalters in positive oder negative Richtung
$\pm \Delta U_{\text{rel}}$	relative Veränderung der Netzversorgungsspannung in positive oder negative Richtung
$\Delta \alpha$	Differenzwinkel von Sonnenazimut α_S und der ausgerichteten Modulfläche α_F
ϵ	Parameter zum Adressieren des Exploration-Exploitation-Dilemmas
η_M	Wirkungsgrad der Photovoltaik-Module
η_{WR}	Wirkungsgrad des Wechselrichters
θ	Parametervektor
$\theta^Q, \theta^{Q'}$	Parametervektor (des Zielnetzwerks) des Q-Netzwerks
$\theta^\mu, \theta^{\mu'}$	Parametervektor (des Zielnetzwerks) des <i>Critic</i>
λ	Hyperparameter der Methode TD(λ)
$\vartheta_M, \vartheta_{M, STC}$	Modultemperatur
ϑ_U	Umgebungstemperatur
μ	deterministische Kontrollstrategie
$\mu_{\cos(\varphi),gg}$	Erwartungswert des Leistungsfaktors einer Gerätegruppe
μ_{DOU}	Erwartungswert der Nutzungsdauer
$\mu_{gg}, \mu_{gg,t}$	Erwartungswert der Wirkleistung einer Gerätegruppe
$\mu_{gg,DOU}, \mu_{gg,DOU, ges}$	Kumulierter Erwartungswert der Nutzungsdauer einer Gerätegruppe
$\mu_{\text{Zeitbereich}}$	Erwartungswert des definierten <i>Zeitbereichs</i>
π, π'	Kontrollstrategie
π_{opt}	optimale Kontrollstrategie
π_θ	stochastische Kontrollstrategie
$\sigma_{\mathcal{P}_t}$	Standardabweichung der Zufallsvariable \mathcal{P}_t
$\sigma_{\cos(\varphi),gg}$	Standardabweichung des Leistungsfaktors einer Gerätegruppe
$\sigma_{gg}, \sigma_{gg,t}$	Standardabweichung der Gerätegruppe
$\sigma_{gg,DOU}$	Standardabweichung der Nutzungsdauer einer Gerätegruppe
σ_{gg,v_i}	Standardabweichung über die Verbraucher der Gerätegruppe
$\sigma_{\text{rel},gg}, \sigma_{\text{rel},gg,t}$	rel. Standardabweichung der Gerätegruppe
$\sigma_{\mathcal{X}}$	Standardabweichung der Zufallsvariable
$\sigma_{\mathcal{X},\text{rel}}$	rel. Standardabweichung der Zufallsvariable
$\sigma_{\text{Zeitbereich}}$	Standardabweichung des definierten <i>Zeitbereichs</i>
τ	Parameter zum Einstellen des weichen Aktualisierens
φ	Phasenwinkel
ω, ω_t	Realisation eines Rauschprozesses zum Zeitpunkt t

Einleitung

Der fortschreitende Paradigmenwechsel des elektr. Energieversorgungssystems sowie die Elektrifizierung weiterer Sektoren erhöht die Komplexität des Gesamtsystems zunehmend. Die daraus resultierenden gesellschaftlichen Kosten finden in der öffentlichen Diskussion um die Themenfelder der Energie-, Wärme- und Verkehrswende immer mehr Beachtung. Auch die notwendigen Netzausbaumaßnahmen und die damit einhergehenden hohen Investitionskosten, welche eine Steigerung der Netzentgelte zur Folge haben, finden Platz im gesellschaftlichen Diskurs. So beziffert die Bundesnetzagentur (BNetzA) in einer Pressemitteilung (vgl. [1]) zum *Jahresbericht 2020 - Märkte im digitalen Wandel* nach [2] einen Gesamtinvestitionsbedarf für die elektr. Übertragungsnetze an Land von ca. 55 Mrd. € bis 2030. Für den Umbau der deutschen Verteilnetze wird der Investitionsbedarf bis 2030 auf 47 Mrd. € geschätzt. Folgt man den Annahmen der Studie nach [3], so machen etwa die Hälfte der 47 Mrd. € Netzausbaumaßnahmen auf der Hochspannungsebene aus. Für die Mittelspannungsebene wird ca. das Dreifache an Investitionskosten im Vergleich zur Niederspannungsebene veranschlagt. Zudem ist der Studie der BNetzA nach [3] zu entnehmen, dass Maßnahmen auf den beiden unteren Spannungsebenen, Mittel- und Niederspannung, nicht in der vorgestellten 10-Jahresplanung der Verteilnetzbetreiber (VNB) enthalten sind, da diese eher kurzfristig geplant und umgesetzt werden. Dies sind insbesondere Maßnahmen, welche die zunehmende Integration von dezentralen Erzeugungsanlagen (DEA), Wärmepumpen oder elektr. betriebenen Personenkraftwagen im elektr. Verteilnetz bei einem zulässigen Netzbetrieb ermöglichen.

Darüber hinaus ist vor dem Hintergrund der für das Jahr 2021 verankerten Änderungen im *Klimaschutzgesetz (KSG)* davon auszugehen, dass die geplanten Vorhaben im KSG die Anzahl und Vielfalt an neuen Betriebsmitteln, Teilnehmern und Akteuren weiter erhöht. So sieht bspw. die Elektrifizierung des Verkehrssektors und der damit einhergehende Umbau der Mobilität in Deutschland einen massiven Ausbau der Ladeinfrastruktur vor. Dies ist notwendig, um die Anzahl von derzeit ca. 1 Mio. elektr. betriebenen Personenkraftwagen auf deutschen Straßen (Stand 1. Januar 2023, rein elektr. betriebene sowie Plug-in-Hybrid-Personenkraftwagen [4]) auf die von Experten empfohlene Anzahl von 14 Mio. elektr. betriebenen Personenkraftwagen bis 2030 anzuheben [5]. Des Weiteren soll der Anteil der erneuerbaren Energien (EE) zur Deckung des Strombedarfs von derzeit ca. 40% bis zum Jahr 2030 auf 80% erhöht werden [6]. Diese ambitionierten Ziele, welche z. B. mit Bezug auf den Ausbau der EE in etwa mit einer Verdoppelung des aktuellen

Anteils am derzeitigen Energiemix gleichzusetzen sind, sind nur durch einen massiven Ausbau der EE und einem rechtlichen und normativen Rahmen, der diesen in dem gesetzten Zeitrahmen zulässt, zu erreichen [7]. Zudem besteht neben der beschriebenen Integration weiterer Teilnehmer und Akteure im Verteilnetz eine wesentliche Herausforderung darin, die volatile Einspeisecharakteristik der DEA im Transformationsprozess des elektr. Energieversorgungssystems sinnvoll zu berücksichtigen [8].

Demzufolge ist davon auszugehen, dass die Verteilnetzebene zunehmend an ihre Leistungsgrenzen gelangt [9]. Dies stellt die VNB auf Mittel- und Niederspannungsebene vor neue Herausforderungen und führt letztlich dazu, dass die eingesetzten Betriebsmittel vermehrt an ihren thermischen Grenzen betrieben werden. Zudem kommt es insbesondere auf den unteren Spannungsebenen des Verteilnetzes vereinzelt zu Spannungsüberhöhungen. Üblicherweise adressieren die VNB diese Problematik mit Hilfe eines konventionellen Netzausbaus [9]. Dieser ist jedoch nicht zwangsläufig die technisch und ökonomisch sinnvollste Lösung. Letztlich erfolgt die Dimensionierung der Netze anhand von Extremszenarien, welche ggf. nur an wenigen Stunden im Jahr auftreten [9].

Um dieser Entwicklung und den daraus resultierenden Auswirkungen auf das Verteilnetz mit einem technisch und ökonomisch sinnvollen Ansatz zu begegnen, bedarf es auch in Zukunft einer zunehmenden Flexibilisierung des Energieversorgungssystems sowie neuer Strategien zur Lösung von Regelungsaufgaben und Entscheidungsprozessen. Die Berücksichtigung regelbarer Betriebsmittel in Betriebsführungsstrategien bietet hierbei das Potenzial, den konventionellen Netzausbau auf ein Minimum zu reduzieren. Mit der vorliegenden Arbeit soll hierzu ein Beitrag geleistet werden.

1.1. Stand der Technik

Für die Flexibilisierung des Energieversorgungssystems bedarf es innovativer, primärtechnischer Betriebsmittel sowie sekundärtechnischer Systeme zur Erschließung weiterer Steuerungsmöglichkeiten, wie bspw. regelbare Ortsnetztransformatoren (rONT), Längsregler oder steuerbare Wechselrichter von Photovoltaikanlagen (PVA). Diese Netztechnologien werden bereits heute aktiv von VNB auf unterschiedlichen Spannungsebenen eingesetzt und erhöhen die Flexibilität des elektr. Verteilprozesses.

Im Allgemeinen werden mit Hilfe von Betriebsführungsstrategien unterschiedliche Ziele der Netzplanung und -führung adressiert. Hierbei ist das vorrangige Ziel Netzengpässe durch Lösung diverser Entscheidungsprozesse zu vermeiden [10]. Dies sind Ansätze, welche die Bestimmung der Netzanschlusskapazität an einzelnen Netzknoten ermöglichen, um bspw. die Position und Einspeiseleistung einzelner DEA zu definieren. Weitere Ansätze legen den Fokus auf die Orchestrierung bestehender steuerbarer Anlagen, wie bspw. DEA oder rONT. Letzteres beschreibt Netzführungsstrategien, welche bspw. der Vermeidung einer (thermischen) Überlastung von Betriebsmitteln oder der Spannungshaltung dienen. Hierbei kann festgehalten werden, dass diese Maßnahmen im Wesentlichen in vier Kategorien der

Regelung zu unterteilen sind [9]:

1. Wirkleistungsregelung
2. Blindleistungsregelung
3. Regelung der Netzspeisespannung
4. eine Kombination der zuerst genannten drei Punkte.

Mit Bezug auf den ersten Punkt wird neben der Anpassung der generatorischen Einspeiseleistung aufseiten von flexiblen Kraftwerken und DEA auch die Anpassung der Verbraucherleistung im Zuge des Lastmanagements diskutiert. Das Lastmanagement, welches in der einschlägigen, englischsprachigen Literatur unter dem Begriff des Demand Side Management bekannt ist, kommt überwiegend bei der energieintensiven Industrie zum Einsatz und bietet bereits heute einen wichtigen Baustein bei der Bereitstellung von Flexibilitätsoptionen auf der Nachfrageseite [11].

Die Bereitstellung von Systemdienstleistung in Form von Blindleistung nach Punkt zwei gehört im Übertragungsnetz sowie im Verteilnetz auf Hoch- und Mittelspannungsebene zum Status quo der Spannungshaltung im elektr. Transport- und Verteilprozess (vgl. [12] und [13]). Mit dem fortschreitenden Rückbau von Großkraftwerken (vgl. [14]) und der zunehmenden Anzahl von DEA des integrierten Energiesystems der Zukunft müssen vermehrt die unteren Verteilnetze bei der aktiven Spannungshaltung unterstützen. Im Rahmen einer komplexen Koordinierungsaufgabe sind DEA entsprechend zu orchestrieren und in Einklang mit den Anforderungen des Energieversorgungssystems zu bringen [12]. Hierfür bedarf es der Studie nach [12] eines Blindleistungsmanagements über alle Spannungsebenen des Verteilnetzes hinweg, welches auch Blindleistung im Normalbetrieb und nicht nur bei Auftreten eines Störfalls bereitstellt. Diese Anforderung ist unter anderem für Windkraftanlagen (WKA) bereits heute verpflichtend, welche unter die *Verordnung zu Systemdienstleistungen durch Windenergieanlagen (SDLWindV)* fallen. So ist in der *SDLWindV* bspw. geregelt, dass Anlagenbetreiber ihre WKA dahingehend qualifizieren müssen, dass VNB diese mittels Fernwirktechnik regeln können, um situativ auf Ereignisse im Netz zu reagieren. Diese Maßnahmen stellen sicher, dass WKA einen Beitrag zur Systemstabilität leisten.

Unter den Punkt drei fallen u.a. Maßnahmen zur aktiven Spannungsregelung im elektr. Verteilnetz. So werden bspw. in klassischen Regelungsverfahren der Stufenschalter der Transformatoren in den Umspannanlagen zwischen Hoch- und Mittelspannungsnetz sowie eine lokale Messung genutzt, um gezielt Einfluss auf die Spannung im Mittelspannungsnetz zu nehmen [15]. Neben dem Einsatzzweck der aktiven Spannungshaltung wird dieses Verfahren außerhalb des europäischen Strombinnenmarktes zur Optimierung weiterer elektrischer Größen genutzt. In der einschlägigen, englischsprachigen Literatur wird dieses klassische Regelungsverfahren als Conservation Voltage Reduction (CVR) bezeichnet. Die Grundlage für diese Netzführungsstrategie basiert auf veröffentlichten Tests, welche bereits im Jahr 1973 durchgeführt wurden [16]. Die Autoren konnten dokumentieren, dass eine bewusst herbeigeführte Reduzierung der Versorgungsspannung unter Einhaltung der

Spannungsbandgrenzen zu einer Einsparung des Energiebezugs der Endkunden führt. Darüber hinaus konnten zusätzliche Vorteile, wie eine Spitzenlastkappung und Reduzierung der Systemverluste, festgehalten werden. Auf diesen Erkenntnissen aufbauend wurden in den 1980er bis Mitte der 1990er Jahre besonders im englischsprachigen Raum eine Vielzahl von Untersuchungen zur spannungsabhängigen Verbraucherleistung durchgeführt. Hierbei wurden unterschiedliche statische und dynamische Modelle zur Beschreibung des Verbraucherverhaltens erarbeitet und etabliert. Eine Zusammenfassung der in dem genannten Zeitraum durchgeführten Untersuchungen und der erarbeiteten Modelle zur Beschreibung der spannungsabhängigen Verbraucherleistung ist der Veröffentlichung [17] zu entnehmen. Die Ergebnisse und Erkenntnisse aus diesen Untersuchungen führten in den darauffolgenden Jahren zu einer weltweiten Verbreitung von spannungsoptimierenden Maßnahmen im Rahmen von Netzführungsstrategien bei den Energieversorgungsunternehmen (EVU) [10]. Insbesondere neue regulatorische und normative Vorschriften mit dem Schwerpunkt auf Energieeffizienz, technologische Verbesserungen und der Flexibilisierung des elektr. Energieversorgungsnetzes rücken diese Netzführungsstrategien auch im liberalisierten Strommarktdesign der europäischen Union wieder ins Rampenlicht. Verfahren zur aktiven Spannungshaltung, sowie die bekannten spannungsabhängigen Lastmodelle, wie bspw. das ZIP-Modell, stehen hierbei im Fokus (vgl. [15]).

Ergänzend hierzu findet heutzutage die Weitbereichsregelung Anwendung, welche insbesondere in den Mittelspannungsnetzen eingesetzt wird [15]. Statt einer lokalen Messung werden bei der Weitbereichsregelung im Vergleich zum klassischen Regelungsverfahren Spannungsmessungen an entfernten Knoten im Mittelspannungsnetz als Grundlage zur Regelung der Stufenschalter herangezogen. So konnte im Forschungsvorhaben nach [18] eine durchschnittliche Reduktion der Spannungsschwankung von 30 % mit Hilfe der Weitbereichsregelung im Mittelspannungsnetz festgestellt werden. Insbesondere in ländlich geprägten Netzen und im Kontext *intelligenter Stromnetze* findet diese Netzführungsstrategie Anwendung (vgl. [15] und [18]). Dies ist u.a. dadurch begründet, dass im Vergleich die Maßnahmen zur Regelung der Netzspeisespannung nach Punkt drei, wie bspw. die Weitbereichsregelung oder das CVR, eine besonders kostengünstige Alternative bieten. Im Gegensatz zum zuvor beschriebenen Last- und Blindleistungsmanagement kann die Netzführungsstrategie CVR als sekundär bezeichnet werden [19]. Für die Implementierung kann auf die bestehende Infrastruktur der VNB zurückgegriffen werden. So ist eine Spannungsanpassung in den Ortsnetzstationen (ONS) mittels rONT oder durch Transformatoren in den Umspannanlagen auf Mittelspannungsebene eine gängige Praxis, um die Spannungsqualität bzw. die Spannung in den geltenden Spannungsbandgrenzen der Europäischen Norm *Merkmale der Spannung in öffentlichen Elektrizitätsversorgungsnetzen (DIN EN 50160)* zu halten [20]. Demnach ist im Vergleich mit geringeren Investitionskosten zur Umsetzung der Weitbereichsregelung oder des CVR zu rechnen und sollte insbesondere bei der Umsetzung neuer Netzführungsstrategien im Kontext der *intelligenten Stromnetze* umfassend genutzt werden [21].

Auch die Kombination verschiedener Kategorien nach Punkt vier respektive parallele Anwendung der einzeln beschriebenen Verfahren ist gängige Praxis. Statische Steuerungskonzepte zur Beeinflussung DEA, Anpassung der Stufe des Stufenschalters der Transformatoren in den Umspannanlagen zwischen Hoch- und Mittelspannungsnetz (vgl. [15]) sowie der Einsatz von rONT an einem einzelnen Mittelspannungsring (vgl. [20]) sind bei vielen VNB im Rahmen von Netzführungsstrategien anzutreffen.

Zudem werden weitere Ansätze und Flexibilitätsoptionen, die über die Nutzung der eingangs genannten Netztechnologien hinausgehen, in neuen Betriebsführungsstrategien und Entscheidungsprozessen berücksichtigt. Neben der Einbindung von Speichertechnologien [22] und der Kopplung von leitungsgebundenen Energieträgern, wie bspw. des elektr. Netzes mit dem Erdgasnetz [23], werden in diesem Kontext zunehmend regionale Flexibilitätsmärkte für eine netzdienliche Nutzung dieser Flexibilitätsoptionen diskutiert, wie bspw. in [24] und [25]. Dabei setzt eine Flexibilisierung und Einbindung von notwendigen Akteuren in die Betriebsführungsstrategien und Entscheidungsprozesse der VNB einen gesetzlichen, normativen und technischen Rahmen voraus. Gesetzgeber, Normungsinstitute sowie die technischen Anschluss- (TAR) und Anwendungsregeln (VDE-AR) des Verbands der Elektrotechnik Elektronik Informationstechnik e.V. (VDE) ermöglichen und empfehlen eine solche Berücksichtigung von Akteuren über alle Spannungsebenen des Verteilnetzes hinweg. So wurde bspw. die *Anwendungsregel Erzeugungsanlagen am Niederspannungsnetz (VDE-AR-N 4105:2011-08)* auf Grundlage der Ergebnisse in [26] dahingehend angepasst, dass die Bereitstellung von Systemdienstleistungen (SDL) in Form von Blindleistung durch DEA sowie der Einsatz elektr. Speicher zur Sicherung der Spannungsqualitätsmerkmale (reglementiert in der DIN EN 50160) berücksichtigt werden. Hierfür wurden u. a. zusätzliche Maßnahmen zur statischen Spannungshaltung in die VDE-AR-N 4105:2018-11 aufgenommen, wie bspw. die Q(U)-Regelung (Anwendungsbeginn 01.11.2018). Weitere Maßnahmen zur Schaffung des gesetzlichen und normativen Rahmens sind bspw. im *Energiewirtschaftsgesetz (EnWG)*, welches u.a. die Anbindung steuerbarer Verbrauchseinrichtungen im Niederspannungsnetz regelt oder im *Erneuerbare-Energien-Gesetz (EEG)* zu finden, welches u.a. eine Einspeisereduzierung mittels Fernwirktechnik seitens des VNB ermöglicht.

Erst auf Grundlage der gesetzlichen und normativen Rahmenbedingungen können Flexibilitätsoptionen im Zuge des Transformationsprozesses des elektr. Energieversorgungssystems in neuen Betriebsführungsstrategien und Entscheidungsprozessen der Netzführung und -planung Berücksichtigung finden. So können unter Einbindung weiterer Flexibilitätsoptionen die Planungsgrundsätze ländlicher Mittelspannungsnetze (vgl. [8]) als auch städtischer Verteilnetze (vgl. [27]) um innovative, sektorenübergreifende Planungsmaßnahmen erweitert werden. In der Netzführung wird durch die zunehmende Flexibilisierung eine optimierte Nutzung vorhandener Netzkapazitäten erreicht. Eine Option ist bspw. durch DEA und deren Bereitstellung von SDL gegeben. Hierdurch kann ein sicherer und stabiler Systembetrieb bei schwankenden Erzeugungs-, Last- und Netzzuständen gewährleistet werden. Dies ist jedoch der dena-Studie nach [15] nur unter der Prämisse gegeben, dass ein adäquates

Niveau an Beobachtbarkeit und Steuerbarkeit in den verschiedenen Netzebenen erreicht ist. Die hierbei erhobenen Daten sollten jedoch nicht nur dem datenerfassenden Teilnehmer zur Verfügung stehen, wie bspw. dem VNB oder dem Anlagenbetreiber. Vielmehr sehen die Konzepte der (regionalen) Flexibilitätsmärkte einen Austausch sowohl zwischen Netzbetreibern untereinander (vertikal und horizontal) als auch zwischen Netzbetreibern und weiteren Marktteilnehmern vor [25]. Mit dem Inkrafttreten der Novelle des *Netzausbaubeschleunigungsgesetzes (NABEG 2.0)* im Jahr 2019 wurden neue Vorgaben für ein Redispatch 2.0 eingeführt. Dieses Vorhaben adressiert einen Datenaustausch- und Koordinierungsprozess über eine zentrale Datendrehscheibe zwischen allen Teilnehmern und sollte zum 1. Oktober 2021 von allen Netzbetreibern und Marktteilnehmern umgesetzt worden sein. Ziel des Redispatch 2.0 ist es, EE- und Kraft-Wärme-Kopplungs- (KWK-) Anlagen aus den Verteilnetzen ab einer Nennleistung von 100 kW zukünftig in Redispatch-Maßnahmen einzubeziehen und hierbei die Belange der Verteilnetz- und Übertragungsnetzbetreiber in Einklang zu bringen [12]. Der dena-Studie nach [12] ist davon auszugehen, dass analog zu dem Engpassmanagement des Redispatch 2.0 weitere Maßnahmen zur Bereitstellung von weiteren Systemdienstleistungen in der Betriebsführung bis zum Jahr 2050 etabliert werden müssen. Insbesondere die Bereitstellung von Regelleistung und Blindleistung. Demnach ist davon auszugehen, dass neue Ansätze und Netzführungsstrategien auch in Zukunft einen Beitrag zum Transformationsprozess des elektr. Energieversorgungssystems im Zuge der Energiewende leisten werden (vgl. [28] und [29]).

1.2. Motivation und Ziele der Arbeit

Wie bereits einleitend angemerkt, stehen die VNB zukünftig vor der Herausforderung weitere Betriebsmittel, Teilnehmer und Akteure netzdienlich und ökonomisch sinnvoll im Verteilnetz zu integrieren. Dies kann durch eine zunehmende Flexibilisierung des Energieversorgungssystems sowie die Ableitung neuer Strategien zur Lösung von Regelungsaufgaben und Entscheidungsprozessen adressiert werden. Hierbei verdeutlichen die zuvor beschriebenen Entwicklungen, dass zukünftige Ansätze und Netzführungsstrategien eine zunehmend größere Menge an Daten, Mess- sowie Stammdaten, einbeziehen werden. Darüber hinaus ist davon auszugehen, dass die einzelnen Verteilnetze grundlegende strukturelle Veränderungen in zunehmend kürzeren Zeitabständen erfahren werden. Insbesondere auf den beiden unteren Spannungsebenen. Dies muss in der Definition von Netzführungsstrategien dahingehend Berücksichtigung finden, dass diese flexibel auf Veränderungen im System angepasst werden können.

Der detaillierten Zusammenstellung und Bewertung an Netzführungsstrategien in [30] ist zu entnehmen, dass der Fokus der Forschungstätigkeiten in der Vergangenheit auf der Hoch- und Mittelspannungsebene lag. Jüngste Forschungstätigkeiten hingegen sind zunehmend auf der Niederspannungsebene anzutreffen. Hierbei steht das netzdienliche Einbringen neuer Akteure im Zuge der Flexibilisierung des elektr. Energieversorgungssystems im Schwerpunkt, wie z. B. in [31], [32], [33], [34] und [35]. Diese zu beobachtende Verlagerung der Forschungstätigkeiten auf die unterste Spannungsebene ist durch die zunehmende Ausstat-

tung von neuralgischen Netzknoten mit Informations- und Kommunikationstechnik (IKT) getrieben (vgl. [36]). Darüber hinaus werden neben dem Einsatz von IKT innerhalb des elektr. Netzes zunehmend Messdaten und Lastgänge mit einer Auflösung von 15 Minuten von Anlagenbetreibern, geregelt im EEG, sowie Endkunden, geregelt im *Messstellenbetriebsgesetz* (MsbG), zur Erhöhung der Transparenz im elektr. Verteilnetz erhoben [37]. Dieser Schritt ist für das integrierte Energiesystem der Zukunft essentiell. Erst durch das Erhöhen der Transparenz kann die komplexe Koordinationsaufgabe zur Erfüllung der Systemsicherheit bewältigt werden. Hiervon abhängig ist nicht nur der Informationsaustausch und das Erbringen von Systemdienstleistungen zwischen den Netzbetreibern, vertikal sowie horizontal, sondern auch das Orchestrieren einer großen Anzahl von Betriebsmitteln, Teilnehmern und Akteuren im Einklang der Übertragungs- und Verteilnetzbetreiber sowie weiterer Marktteilnehmer (vgl. [15] und [12]).

Hierbei stellt das zuvor beschriebene Orchestrieren ein Optimierungsproblem dar. Ein möglicher Lösungsansatz hierfür ist der sog. optimale Leistungsfluss. Dieser basiert auf der Veröffentlichung [38] aus dem Jahr 1962 und stellt heute eines der am häufigsten eingesetzten Verfahren zur Lösung von nichtlinearen Optimierungsproblemen bei Fragestellungen im Kontext des Energieversorgungssystems dar (vgl. [39] und [40]). Allgemein wird der optimale Leistungsfluss zur ganzheitlichen Optimierung der generatorischen Einspeiseleistung sowie des elektr. Transport- und Verteilprozesses unter Berücksichtigung von Systembeschränkungen und -grenzen eingesetzt [39]. Allerdings gelten aktuelle Algorithmen als weniger robust und skalieren zudem schlecht (vgl. [39] und [41]). Darüber hinaus gehen die klassischen, modellbasierten Ansätze der Gruppe der Algorithmen zur Bestimmung des optimalen Leistungsflusses von der Prämisse aus, dass die physikalischen Parameter bekannt und fehlerfrei sind [42]. Die Ergebnisse des Forschungsvorhabens *Erfassung der niederspannungsseitigen Netzzustandsgrößen in Echtzeit* (ENERGIE) [36] belegen jedoch, dass dies in der Praxis nicht gegeben ist. Vor dem Hintergrund der integrierten Energiewende mit den daraus resultierenden Herausforderungen eines zunehmend komplexeren elektr. Energieversorgungssystems bedarf es demnach gut skalierender und adaptiver Lösungsansätze, die explizit Unsicherheiten berücksichtigen. Hierzu soll die vorliegende Arbeit einen Beitrag leisten.

Im Hinblick auf die zunehmende Ausstattung der unteren Verteilnetzebenen mit IKT rücken sog. modellfreie, datengetriebene Algorithmen als möglicher Lösungsansatz in den Fokus [43]. Hierbei werden den Methoden der künstlichen Intelligenz (KI) und des maschinellen Lernens (ML), getrieben durch einen enorm ansteigenden Bedarf an Daten- und Informationsaustausch im Zuge der Digitalisierung, ein „grundsätzlich günstiges Umfeld“ für den Einsatz in der Energiewirtschaft bescheinigt und ein zunehmendes Potenzial zugesprochen [44]. Zudem ist aufgrund von aktuellen Entwicklungen und Verbesserungen der bestehenden Ansätze und Algorithmen in den vergangenen Jahren das Interesse der Energieversorgungsbranche an datenbasierten Methoden der KI und des ML weltweit gestiegen (vgl. [28] und [45]).

Im Fokus der vorliegenden Arbeit steht das ML, welches im Allgemeinen anhand der unterschiedlichen *Lernstile* in drei Gruppen unterteilt werden kann [44]:

1. Überwachtes Lernen (aus dem engl. supervised learning)
2. Unüberwachtes Lernen (aus dem engl. unsupervised learning)
3. Bestärkendes Lernen (aus dem engl. reinforcement learning; RL)

Speziell die modellfreien Ansätze des RL wurden zunehmend in der jüngsten Vergangenheit in wiss. Publikationen zur Lösung von Entscheidungsprozessen im Kontext des Energieversorgungssystems genutzt, wie z.B. in [42], [45] sowie [46]. Den Autoren in [46] zufolge sind insbesondere die Ergebnisse und Verbesserungen hinsichtlich der Stabilität der *Lernphase* des Algorithmus in [47] und der damit einhergehende Beginn des sog. tiefen bestärkenden Lernens (aus dem engl. deep reinforcement learning; DRL) für die Renaissance des RL ausschlaggebend.

Mit Blick auf die zuvor beschriebenen Herausforderungen im Zuge der integrierten Energiewende soll die vorliegende Arbeit eine aktiv herbeigeführte Spannungsanpassung als Handlungsoption, welche im Kontext eines intelligenten Netzbetriebs im Niederspannungsnetz Berücksichtigung findet, adressieren. Der Fokus wird auf Netzführungsstrategien mit dem Ziel der Einhaltung der Spannungsbänder, vorgegeben von der Europäischen Norm DIN EN 50160, gelegt. Konkret stehen Maßnahmen im Vordergrund, welche unter dem Themenfeld der Spannungshaltung zusammengefasst werden. Mit Bezug auf die zuvor abgeleiteten Anforderungen an Netzführungsstrategien wird mit Hilfe des DRL ein zentrales, übergeordnetes Regelkonzept zur Steuerung aktiv regelbarer Betriebsmittel erarbeitet. Im Zuge der Untersuchungsszenarien wird auf reale Mess- sowie Stammdaten des Forschungsvorhabens ENERGIE [36] zurückgegriffen. Zusammengefasst adressiert die vorliegende Arbeit eine Weitbereichsregelung auf Niederspannungsebene unter Zuhilfenahme des RL und liefert einen Beitrag zum Forschungsschwerpunkt der autonom gesteuerten Lösungsansätze im Niederspannungsnetz.

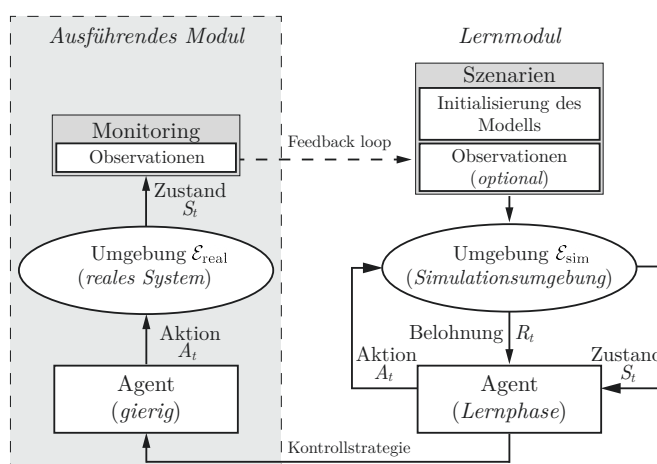


Abbildung 1.1.: Ein bewährtes Rahmenwerk für Regelungsaufgaben und Entscheidungsprozesse, welches Fragestellungen mit Bezug auf das Energieversorgungssystem adressiert (eigene Darstellung nach [28]).

Im Allgemeinen kann beim Adressieren von Fragestellungen im Kontext des elektr. Energieversorgungssystems mit dem RL auf das in [28] skizzierte Rahmenwerk zurückgegriffen werden. Dem Rahmenwerk, illustriert in Abbildung 1.1, sind zwei unterschiedliche Module, das *Lernmodul* sowie das *ausführende Modul*, zu entnehmen. Darüber hinaus sind ein Agent sowie zwei unterschiedliche Umgebungen \mathcal{E} , das *reale* System $\mathcal{E}_{\text{real}}$ und eine Simulationsumgebung \mathcal{E}_{sim} , für das *ausführende Modul* respektive *Lernmodul*, dargestellt. Das *ausführende* sowie das *Lernmodul* stellen hierbei verschiedene Phasen dar und demnach verschiedene Arbeitsweisen des Agenten. Diese Aufteilung in zwei separate Module ist notwendig, um kritische Systemzustände im *realen* (Energieversorgungs-) System zu vermeiden.

Zur Lösung des definierten Anwendungsfalls mit Hilfe des DRL wird zunächst die Simulationsumgebung \mathcal{E}_{sim} in Kapitel 2 beschrieben. Der Schwerpunkt des Kapitels liegt bei der Beschreibung der individuellen statistischen Systematik des elektr. Niederspannungsnetzes. Zudem wird eine detaillierte Modellierung der Spannungsabhängigkeit der Last- und Einspeiseprozesse thematisiert, welche im Kontext spannungsoptimierender Maßnahmen unerlässlich ist. Hierbei erfolgt mit Hilfe von Modellen zur Beschreibung der spannungsabhängigen Prozesse im Zuge einer Netzsimulation die Bewertung der Güte der vorgestellten Maßnahmen und Regelungskonzepte (vgl. [31], [32] und [34]). Dabei wird das energiekonstante Verhalten (mit $E = \text{konstant}$) einzelner Verbraucher i.d.R. nicht berücksichtigt. Pinney vergleicht im Rahmen der Studie in [21] ein Modell eines haushaltsüblichen Warmwasserspeichers, auch als Untertischgerät bezeichnet, welches explizit als energiekonstant modelliert ist, mit dem ZIP-Modell, das zur Modellierung des spannungsabhängigen Verhaltens von Verbraucherlasten genutzt wird (vgl. [17]). Der Autor vergleicht die beiden Modelle und das Lastverhalten während einer aktiv herbeigeführten Spannungsänderung über einen Zeitraum von vier Stunden und stellt eine zunehmende Abweichung beider Modelle mit Fortschreiten der Zeit fest. Diesen Effekt, den Pinney in [21] dokumentiert, beschreibt Willis in [48] als Kompensationseffekt (aus dem engl. duty cycle rebound; DCR), der entgegen der Spannungsänderung wirkt. Beide Autoren belegen eine direkte Korrelation sowie einen signifikanten Einfluss auf die Güte der Netzführungsstrategien unter Berücksichtigung des DCR. So dokumentiert Pinney bspw. eine Reduzierung der Energieeinsparung von 14 % ohne Berücksichtigung des DCR, auf 4 % mit Berücksichtigung des DCR. Demnach ist bei der Bewertung von Netzführungsstrategien in Abhängigkeit der zeitlichen Auflösung der Maßnahme das energiekonstante Verhalten einzelner Verbraucher explizit zu berücksichtigen. Infolgedessen thematisiert das Kapitel 2 die Notwendigkeit der Berücksichtigung energiekonstanter Lasten. Hierfür wird die Beschreibung der Spannungsabhängigkeit um den Bereich der energiekonstanten Prozesse erweitert und das ZIP(E)-Modell vorgestellt.

Da eine umfangreiche Bewertung energiekonstanter Prozesse im Niederspannungsnetz der einschlägigen Literatur nicht zu entnehmen ist, soll diese Lücke im Rahmen der vorliegenden Arbeit adressiert werden. Insbesondere der zuvor skizzierte Einfluss des energiekonstanten Verbraucherlastverhaltens und der daraus resultierende DCR wird hierbei im Zuge potenzieller Flexibilitätsoptionen erstmals in der vorliegenden Arbeit im Kapitel 3 im De-

tail untersucht. Hierfür werden einzelne Prozesse mit einer konstanten Energiebedingung modelliert und im Kontext spannungsoptimierender Maßnahmen innerhalb von Betriebsführungsstrategien berücksichtigt. Mit Hilfe einer Netzsimulation wird der Einfluss auf ein passives sowie aktives Niederspannungsnetze untersucht. Dabei wird der Fokus der Untersuchung im Kapitel 3 auf die Identifizierung versteckter bzw. ungenutzter Flexibilität im Netz gelegt. Abschließend wird eine potenzielle Berücksichtigung des intrinsischen Betriebsverhaltens der Netzteilnehmer bei aktiver Änderung der Versorgungsspannung im Rahmen des intelligenten Netzbetriebs diskutiert.

In Kapitel 4 wird das in Abbildung 1.1 dargestellte Rahmenwerk des RL und dessen Beitrag zur integrierten Energiewende im Kontext von Netzführungsstrategien vorgestellt. Die einzelnen Module, der damit einhergehende Arbeitsmodus des Agenten sowie die Definition der Signale Systemzustand S_t , Aktion A_t sowie Belohnung R_t des Markow-Entscheidungsprozesses (aus dem engl. Markov decision process) sind beschrieben. Der Algorithmus, welcher den Agenten beschreibt und die damit einhergehende Ableitung einer Kontrollstrategie werden darüber hinaus thematisiert. Anschließend wird ein generalisierter Ansatz zur Spannungshaltung vorgestellt. Hierbei wird die Weitbereichsregelung im Niederspannungsnetz fokussiert. Unter Verwendung eines tiefen deterministischen Policy-Gradienten (aus dem engl. deep deterministic policy gradient; DDPG) nach [49] wird ein DRL-basierter Ansatz zur Ableitung einer (optimalen) Kontrollstrategie vorgestellt und mithilfe des ML ein zentrales, übergeordnetes Regelkonzept zur Steuerung aktiv regelbarer Betriebsmittel abgeleitet. Hierfür sieht der vorgestellte Ansatz vor, mit Hilfe von Lernszenarien die Ableitung einer Kontrollstrategie innerhalb der *Lernphase* zu ermöglichen. Es werden gezielt Szenarien definiert, welche zum einen auf Grundlage stochastischer Last- und Einspeisemodelle im Rahmen einer Monte-Carlo-Simulation (möglichst) alle potenziellen Netzzustände im normalen Netzbetrieb erzeugen, und zum anderen sowohl exogene als auch endogene Einflüsse berücksichtigen.

Im Fokus des vorgestellten datenbasierten Ansatzes steht eine Reduzierung von Schalt-handlungen des Transformatorstufenschalters. Dies ist insbesondere vor dem praktischen Hintergrund interessant, dass häufige Schalthandlungen zu einer Verkürzung des Revisionsintervalls führen. Hierfür wird erstmals in der vorliegenden Arbeit eine lineare Funktion sowie ein Intervall zurückliegender Beobachtungen zur Berechnung des Trends verwendet und als zusätzliche Information im Rahmen des Markow-Entscheidungsprozesses berücksichtigt. Im Vergleich zu Beiträgen der einschlägigen Literatur wird hierdurch eine Reduzierung der Komplexität erzielt. So sehen bspw. die Autoren der Publikation nach [42] eine Vorhersage des kommenden Systemzustandes mit Hilfe von rekurrenten neuronalen Netzen vor.

Im Zuge eines Untersuchungsszenarios folgt eine Verifizierung der gemachten Annahmen. Hierfür werden zwei unterschiedliche Kontrollstrategien erarbeitet. Zum einen mit Hilfe der zusätzlichen Steigung, zum anderen ohne Berücksichtigung dieser. Im anschließenden

Resümee erfolgt eine Bewertung der gemachten Prämisse sowie der Weitbereichsregelung im elektr. Niederspannungsnetz mit Hilfe des RL im Allgemeinen.

Abschließend folgt in Kapitel 5 eine Zusammenfassung und Diskussion der wesentlichen Erkenntnisse der vorliegenden Arbeit.

Modellierung spannungsabhängiger Last- und Einspeiseprozesse

Die Realisierung der Energiewende findet überwiegend auf den unteren Spannungsebenen statt und stellt VNB vor neue Herausforderungen. Dabei müssen insbesondere Methoden und Erkenntnisse auf Grundlage erhobener Informationen zurück in die Entscheidungsprozesse der Betriebsführung fließen, um eine ganzheitliche Modellierung der Systemcharakteristika zu gewährleisten [50]. Diese Zunahme der Komplexität basiert auf einer erhöhten resultierenden Dynamik und Diversität im Niederspannungsnetz, die durch die Ergebnisse des Forschungsvorhabens *Erfassung der niederspannungsseitigen Netzzustandsgrößen in Echtzeit (ENERGIE)* bestätigt wurden [36]. Zur Verdeutlichung dienen die in Abbildung 2.1 dargestellten Mess- und Analysedaten einzelner Niederspannungsabgänge eines ländlichen Ortsnetzes der Gemeinde Wachtendonk. Dem linken Diagramm sind 1-Minuten Mittelwerte des Betriebsstroms an vier gemessenen Netzabgängen eines Tages zu entnehmen. Das rechte Diagramm visualisiert die statistische Analyse über einen Zeitraum von Dezember 2013 bis März 2014. Den Diagrammen der Abbildung 2.1 ist das diverse Verhalten der einzelnen Netzabgänge des Ortsnetzes zu entnehmen. Zudem ist der Ausprägung der unsymmetrischen Strombelastung eine individuelle statistische Systematik zu unterstellen (vgl. Abbildung 2.1, rechte Seite) [51].

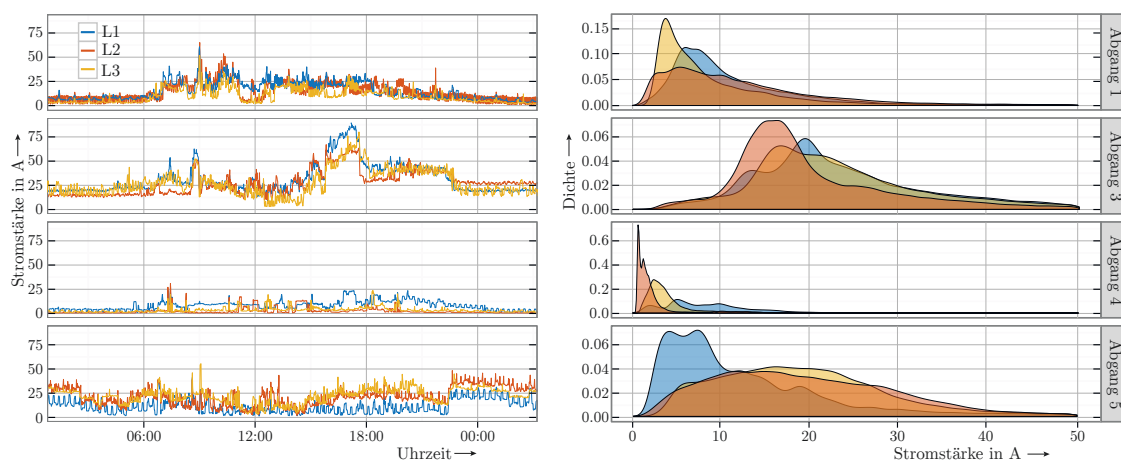


Abbildung 2.1.: Niederspannungsseitige Messung des 1-Minuten Mittelwertes des Betriebsstroms eines ländlichen Ortsnetzes der Gemeinde Wachtendonk über 24 h am 4. Dezember 2013 (links) sowie die statistische Analyse (rechts) für einen Zeitraum von Dezember 2013 bis März 2014. Gemessen wurden vier Netzabgänge (eigene Darstellung nach [36]).

Zur Berücksichtigung des beschriebenen individuellen statistischen Verhaltens sind neue Modellierungsansätze, insbesondere für die Knotenmodellierung von Last- und Einspeiseprozessen vorzusehen. Darüber hinaus sind vorhandene Unsicherheiten im Rahmen der Modellbildung einzubeziehen. Letztlich gehört die Abschätzung des Leistungsbedarfs der Endkunden zum Alltag der EVU sowie VNB und ist essentiell für eine ökonomische Bewirtschaftung der elektr. Netze. Mit Fortschreiten der integrierten Energiewende, welche eine Zusammenführung und Bewirtschaftung weiterer Sektoren mit EE in einem intelligenten Energiesystem vorsieht, bedarf es dieser neuen Ansätze zur Ermittlung des Leistungsbedarfs. Zudem muss eine effiziente Einbindung neuer Akteure, wie PVA [52] oder Wärmepumpen [53], in unterschiedliche interne Betriebs- und Planungsprozesse vorgesehen werden [8] [27].

Nachfolgend werden Modellierungsansätze für Last- und Einspeiseprozesse vorgestellt. Der Status quo sowie neue probabilistische Ansätze werden thematisiert. Hierbei werden die erstmals in [54] erarbeiteten und in der vorliegenden Arbeit erweiterten probabilistischen Ansätze zur Ableitung synthetischer Last- und Einspeiseprozesse im elektr. Niederspannungsnetz dahingehend erweitert, dass eine Spannungsabhängigkeit und das energiekonstante Verhalten einzelner Prozesse explizit berücksichtigt werden. Die erarbeiteten Modelle werden anschließend verifiziert und validiert. Eine reduzierte Datenbasis, die auf Informationen beruht, welche dem VNB vorliegenden, liegt im Fokus der Modellbildung.

2.1. Verbraucherlastmodelle

Im Zusammenhang mit einer Abschätzung der Vorgänge im elektr. Verteilnetz stellt insbesondere eine realitätsnahe Modellierung der Endverbraucher eine Herausforderung dar. Speziell das Einzelverhalten von Individuen erschwert das Ableiten guter Modelle. Hierbei ist das primäre Ziel, ein Verbraucherlastmodell abzuleiten, welches mit den bereits heute zur Verfügung stehenden Informationen aufgebaut wird. Die Erzeugung möglichst realitätsnaher Lastgänge soll eine Untersuchung der Prozesse im Detail ermöglichen.

Der einschlägigen Literatur sind eine Vielzahl an unterschiedlichen Ansätzen mit unterschiedlicher Modellierungstiefe zu entnehmen. Im Rahmen der vorliegenden Arbeit finden folgende vier Ansätze Berücksichtigung:

- Status quo:
 - Extremszenarien sowie
 - Repräsentative Lastprofile.
- Probabilistischer Ansatz:
 - *Top-Down*-Verfahren sowie
 - *Bottom-Up*-Verfahren.

Das Betrachten von Extremszenarien sowie die Nutzung repräsentativer Lastprofile stellen gängige Ansätze dar, die auf dem Jahresenergiebedarf (JEB) und den Standardlastprofilen

(SLP) beruhen. Diese beiden Verfahren, welche fester Bestandteil der täglichen Planungs- und Bilanzierungsprozesse der VNB sind und demnach den Status quo darstellen, sollen im nachfolgenden Abschnitt näher betrachtet werden. Im Gegensatz hierzu werden im direkten Anschluss an Abschnitt 2.1.1 das *Top-Down*- und *Bottom-Up*-Verfahren vorgestellt. Diese stellen ein probabilistisches Verfahren dar, welches die stochastischen Vorgänge im Netz unter Berücksichtigung von Unsicherheiten modelliert.

2.1.1. Status quo

In der Praxis erfolgt die Abschätzung des Leistungsbedarfs von Endkunden für eine Dimensionierung elektr. Betriebsmittel anhand weniger Informationen. Dabei zeichnet insbesondere die reduzierte Datenmenge diese Verfahren aus, welche u.a. auf eine Tarifreform der 80er Jahre zurückzuführen ist [55]. Aus der Reform ging hervor, dass zur Tarifaufnahme ausschließlich elektr. Kriterien, wie der jeweilige JEB des Kunden Berücksichtigung finden und nicht etwa die Anzahl der Räume eines Haushaltes oder die Personenanzahl festgehalten und nachgehalten werden müssen [55]. Davon ausgenommen sind Kunden mit registrierten Leistungsmessungen, der sog. RLM. Dies erscheint unter Berücksichtigung der Rechnerleistung und Speicherkapazitäten der damaligen Zeit sowie der damit einhergehenden Kosten der Systeme als ökonomisch sinnvolle Prämisse.

2.1.1.1. Extremszenarien

Ein weitverbreiteter Ansatz zur Planung und Auslegung von Netzkapazitäten stellt die Betrachtung von Extremszenarien dar [56]. Hierbei werden im Rahmen der Netzdimensionierung folgende Szenarien durchgespielt:

- a) Max. Leistungsbedarf bei min. Rückspeisung sowie
- b) min. Leistungsbedarf bei max. Rückspeisung.

Szenario a

Für das erste Szenario ist lediglich eine Abschätzung des max. Leistungsbedarfs notwendig. Dies erfolgt mit Hilfe des Gleichzeitigkeitsgrades nach

$$p_s(N) = p_{s, \max} \cdot [g + (1 - g) \cdot N^{-\frac{3}{4}}] , \quad (2.1)$$

mit $p_{s, \max}$ für die Spitzenlast eines Haushalts-Endkunden, g für den Gleichzeitigkeitsgrad und N für die Anzahl der Haushalts-Endkunden [57]. Die Spitzenlast $p_{s, \max}$ wird am Beispiel eines vollelektrifizierten Haushaltes (ohne Berücksichtigung einer Elektroheizung) typischerweise mit $p_{s, \max} = 30 \text{ kW}$ und einem Gleichzeitigkeitsgrad im Bereich von $g = 0,06 \dots 0,07$ angegeben [58]. Exemplarisch ist der Verlauf der Spitzenlast $p_s(N)$ eines vollelektrifizierten Haushaltes der Abbildung A.1 des Anhangs A.1 zu entnehmen. Der abnehmende Einfluss $p_s(N)$ eines einzelnen Haushaltes bedingt durch den Gleichzeitigkeitsgrad ist deutlich erkennbar.

Wieland et al. vergleicht in [59] den konventionellen mit einem probabilistischen Ansatz. Die Autoren begründen die Notwendigkeit neuer Ansätze mit fehlenden Erfahrungswerten und Gleichzeitigkeitsfaktoren zur Berücksichtigung neuer Akteure im konventionellen Ansatz und empfehlen eine Ergänzung mit Hilfe synthetisch erstellter Leistungsprofile.

Szenario b

In diesem Fall wird unterstellt, dass alle installierten Anlagen (sofern vorhanden), wie bspw. PVA, mit ihrer projektierten Leistung P bzw. der max. Einspeiseleistung P_{Peak} ins Netz einspeisen und diesem nur der Schwachlastfall der Verbraucher gegenübersteht. Dabei liegen die projektierten Leistungen dem VNB i.d.R. durch die im EEG verankerte Meldepflicht für Anlagenbetreiber und Eigenversorger vor. Als Schwachlastfall kann der Anteil des Stand-by-Betriebs der Verbraucher mit 11 % des JEB angenommen und gleichmäßig auf einen Tag verteilt werden [60].

Die angenommene max. Rückspeisung des beschriebenen Szenarios b unterstellt einen Gleichzeitigkeitsgrad $g = 1$. Am Beispiel der PVA bedeutet diese Herangehensweise, dass die Anlage mit den unter Standardtestbedingungen (aus dem engl. standard test conditions; STC) theoretisch erreichbaren Werte einspeist [61]. Darüber hinaus entspricht die Annahme des Gleichzeitigkeitsgrads für PVA $g_{\text{PVA}} = 1$ der VDE-AR-N 4105 folgend einer identischen Ausrichtung aller im Netzgebiet installierten PVA. Dies entspricht aufgrund der Diversität von Standort, Bauform und Ausrichtung einzelner Gebäude kaum der Wirklichkeit. Eine Reduzierung der abgeschätzten Einspeiseleistung auf $g_{\text{PVA}} = 0.85$ wie von den Autoren in [62] vorgeschlagen kann dies kompensieren. Die gesamte Betrachtung liefert einfache Leistungswerte und berücksichtigt keine zeitliche Verschiebung (vgl. [52]).

Zudem weisen Wiest und Finkel in [63] darauf hin, dass die Wahrscheinlichkeit des Auftretens eines der beiden Szenarien sehr gering ist und das Netz nur zu sehr wenigen Zeitpunkten im Jahr, an die durch die Szenarien beschriebenen Belastungsgrenzen kommt. Demzufolge tendiert die Betrachtung von Extremszenarien im Allgemeinen zu einer Überdimensionierung der Netze. Dabei bieten besonders der Rollout der Messsysteme und die heutigen Rechenkapazitäten das Potenzial komplexe Methoden zur Kapazitätsbestimmung und Netzdimensionierung mit einer höherwertigen Qualität der Aussage einzusetzen.

2.1.1.2. Repräsentative Lastprofile

Die repräsentativen Lastprofile, heutzutage bekannt als SLP, finden nach wie vor Anwendung in der Praxis. So ist es üblich, dass ein EVU mit Hilfe von [64] SLP für das kommende Abrechnungsjahr bestimmt und jeder in der *Stromlieferkette* Zugriff hierauf bekommt. Hierdurch konnte im Rahmen des sog. analytischen Lastprofilverfahrens (beschrieben in [65]) der bis dahin vorherrschende Ansatz zur Berücksichtigung von Endkunden mit geringem Verbrauch ersetzt werden, welcher auf empirischen Größen und Gleichzeitigkeitsgraden beruhte [55].

Die Einführung repräsentativer Lastprofile ist auf die Harmonisierung und Liberalisierung des Energiebinnenmarkts zurückzuführen. Hierfür wurde die erste Richtlinie 1996 von der Europäischen Union verabschiedet und 1998 in Deutschland in nationales Recht umgesetzt. Für die leitungsggebundene Stromversorgung beinhaltete dies u.a. Maßnahmen, die dem Endkunden eine freie Wahl des Energielieferanten einräumte, um die Märkte schrittweise für den Wettbewerb zu öffnen. Ziel des Vorhabens war, die in den 1990er-Jahren vorherrschenden Monopole im Strom- und Erdgassektor zu liberalisieren [66]. Um dies zu ermöglichen, war es zwingend erforderlich die Netzbetreiber zu verpflichten die Netzinfrastruktur diskriminierungsfrei für jeden Lieferanten zur Verfügung zu stellen. Für diese Bereitstellung und die damit verbundenen Kosten sah der Rechtsrahmen eine Zahlung eines Netznutzungsentgelts vor (u.a. im EnWG geregelt).

Die Autoren in [55] verweisen darauf, dass die Definition der Netznutzungsentgelte jedoch nicht ausreichend sei, um die gesamte *Stromlieferungskette* zu beschreiben. Dabei legen die Autoren ihren Fokus auf Endkunden mit geringem Verbrauch, bei welchen eine RLM aus ökonomischen Gründen keine Anwendung findet. Dem entgegen spricht die Tatsache, dass die Lieferanten verpflichtet sind den Endkunden stets zeitgerecht zu beliefern. Dies muss über die gesamte *Stromlieferkette* gegeben sein, um eine möglichst synchrone Einspeisung und Abnahme zu gewährleisten. Hierfür ist jedoch eine genaue Kenntnis über die Kundenstruktur und die zu erwartende Netzlast erforderlich.

Dieses beschriebene Spannungsfeld zwischen guter Vorhersage der Netzlast und kostengünstiger Abrechnung von Endkunden mit geringem Verbrauch adressieren die Autoren in [55] mit dem Ziel,

„(...) aus früheren Messungen (...) Lastprofile zu ermitteln, die dem wechselwilligen Einzelkunden in der Praxis einfach zuzuordnen sind und dennoch die Charakteristika bestimmter Kundengruppen so hinreichend genau widerspiegeln, daß sie von Anbietern, Versorgungsunternehmen und Netzbetreibern gleichermaßen anerkannt und als Grundlage für die erforderlichen Transaktionen dienen können.“ [55]

Als Grundlage der SLP wurden Lastgänge aus verschiedenen Quellen zusammengetragen. In Summe wurden 1209 Einzelkunden-Lastgänge berücksichtigt, davon 332 Haushaltskunden (H), 260 landwirtschaftliche Betriebe (L) und die übrigen stellen einen gewerblichen Bedarf (G) dar [55]. Die Einzelkunden-Lastgänge stammen dabei aus Messreihen verschiedener EVU sowie aus breit angelegten Untersuchungen, wie bspw. in [67], welche in Auftrag durch den Verband der Elektrizitätswirtschaft e. V. (VDEW) durchgeführt wurden. Die Aufteilung der Lastgänge auf die einzelnen Kundengruppen ist der Abbildung 2.2 zu entnehmen. Die verwendeten Datensätze sind mit Angabe des Messzeitraums sowie den zugehörigen Profiltypen in Tabelle 2.1 aufgelistet.

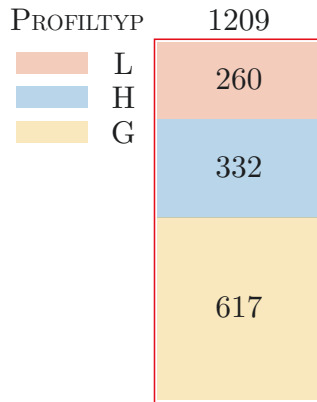


Abbildung 2.2.: Aufteilung der 1209 Datensätze in Kundengruppen zur Ableitung der SLP [55].

Tabelle 2.1.: In [55] verwendete Datensätze zur Ableitung der SLP.

DATENSATZ	ZEITRAUM	PROFILTYP
VDEW	81/82	H, G, L
VDEW	86/87	H, G, L
BEWAG	82-85	G
RWE	92/93	L
Neckarwerke	94-96	G
EVS Chemnitz	95	H, G
OSE Fürstenwalde	96/97	H
RWE	80-97	H
SW Brühl	96-98	G
SW Finsterwalde	98	H
VEW	96/97	H, G, L

Nach Auswertung der Datensätze wurden im Rahmen der Untersuchung nach [55] erstmals die SLP einzelner Kundengruppen nach Tabelle 2.2 definiert. Zudem verweisen die Autoren auf grundsätzliche Charakteristiken der Lastgänge. Insbesondere auf eine Unterscheidung der Zeiträume in

- Wochentag:
 - *Werktag*,
 - *Samstag* sowie
 - *Sonntag*.
- Jahreszone:
 - *Winter*: 01.11. bis 20.03.,
 - *Sommer*: 15.05. bis 14.09. sowie
 - *Übergang*: 21.03. bis 14.05. und 15.09. bis 31.10. des Jahres.

Demnach wurden für jeden Profiltyp neun Lastprofile (jeweils drei Wochentage in drei Jahreszonen) mit 96 Viertelstunden-Leistungswerten abgeleitet, welche auf einen JEB von 1000 kWh/a normiert wurden. In Anhang A.2 sind exemplarisch die SLP des Profiltypen H0 für die Jahreszone *Winter* aller Wochentage dargestellt.

Zudem empfehlen die Autoren, dass bei der Berücksichtigung der Jahreszonen für den Profiltyp H0 nach Tabelle 2.2 eine über das Jahr hinweg kontinuierliche Anpassung Anwendung finden soll. Dies wurde dadurch begründet, dass die Messdaten der H0-Profiltypen im Gegensatz zu den Übrigen, eine stetige Änderung der Last auch innerhalb der Jahreszone, aufweisen. Um dies in den SLP zu berücksichtigen wurde ein Polynom 4. Ordnung, der sog. Dynamisierungsfaktor

$$x = x_0(-3,92 \cdot 10^{-10}d^4 + 3,20 \cdot 10^{-7}d^3 - 7,02 \cdot 10^{-5}d^2 + 2,10 \cdot 10^{-3}d + 1,24) \quad (2.2)$$

eingeführt [55]. Die SLP vom Typ H0 sind demnach vor der Anwendung mit dem Dynamisierungsfaktor zu multiplizieren, wobei für Gleichung 2.2 gilt

- x der resultierende Viertelstundenwert,
- x_0 der Viertelstundenwert des Profils sowie
- d Tag des jeweiligen Jahres, beginnend mit 1 am ersten Januar.

Zudem lässt Tabelle 2.2 den Schluss zu, dass die Ableitung eines generalisierten Lastprofils für gewerbliche Betriebe (Profiltyp G) nicht zu der erwünschten Approximation führte. Dies ist bei der Diversität, die die unterschiedlichen gewerblichen Betriebe aufweisen können, wenig verwunderlich. Um eine mögliche Problematik bei der Zuordnung der Endkunden mit geringem Verbrauch zu den Profiltypen vorwegzunehmen, wurde eine Empfehlung in [68] ausgearbeitet, welche als Unterstützung und Hilfestellung für die Praxis gedacht ist.

Das SLP findet nach wie vor Anwendung in den betrieblichen Prozessen der VNB und dient der Abschätzung des Verbraucherverhaltens [57]. Jedoch wird in [69] daraufhin gewiesen, dass die SLP nur eine globale Aussagekraft haben. Es sind Abweichungen zu erwarten, welche durch geographische, sozio-ökonomische und -demographische Unterschiede zu erklären sind. Diesen Punkt adressieren die Autoren in [70]. Sie empfehlen Daten intelligenter Messsysteme oder digitaler Stromzähler (iMSys) zu nutzen, um aktuelle und regionale Profile nach gleichem Muster abzuleiten und die Aussagekraft der Profile im regionalen Kontext zu verbessern.

Tabelle 2.2.: Darstellung resultierender Profiltypen der SLP mit der zugehörigen Kundengruppe nach Auswertung von 1209 Einzelkunden-Lastgängen [55].

PROFILTYP	KUNDENGRUPPE
H0	Haushalt, Privatverbrauch, ggf. geringfügig gewerblicher Bedarf
G0	Gewerbe allgemein, Mittelwert der Gesamtgruppe
G1	Gewerbe, werktags 8-18 Uhr (z. B. Büros, Arztpraxen, Werkstätten, Verwaltungseinrichtungen, ...)
G2	Gewerbe, Überwiegender Verbrauch in den Abendstunden (z. B. Abendgaststätten, Freizeiteinrichtungen, Sportvereine, Fitnessstudios, Solarien, ...)
G3	Gewerbe durchlaufend (Kühlhäuser, Pumpen, Gemeinschaftsanlagen, ...)
G4	Gewerbe, Läden aller Art, Friseur
G5	Gewerbe, Bäckerei mit Backstube
G6	Gewerbe, Wochenendbetrieb (Schwerpunkt) (z. B. Gaststätten, Ausflugslokale, Kinos, Sporteinrichtungen, ...)
L0	Landwirtschaft allgemein, Mittelwert der Gesamtgruppe
L1	Landwirtschaftsbetriebe mit Milchwirtschaft/ Tierzucht
L2	Übrige Landwirtschaftsbetriebe

2.1.2. Probabilistischer Ansatz

Zur Analyse und Berücksichtigung der charakteristischen Netzdynamik der zukünftigen Energieversorgung bedarf es Ansätze, welche insbesondere die innertäglichen Prozesse im Verteilnetz abbilden. Zwar bringt das SLP im Vergleich zu den Gleichzeitigkeitsgraden eine zeitliche Abhängigkeit mit sich, jedoch entspricht die Lastcharakteristik eines einzelnen Endkunden in keiner Weise den geglätteten Verläufen der SLP. Erst die Akkumulation von ca. 150 Endkunden zeigt (annähernd) einen solchen Verlauf [71]. So ist es wenig verwunderlich, dass im Rahmen der Netzdimensionierung und -planung sowie der Bewertung neuartiger Netzführungsstrategien ein Trend hinsichtlich der probabilistischen Modellierung der Prozesse auf Niederspannungsebene zu erkennen ist (vgl. [57], [72] sowie [73]). Dabei werden u.a. die zuvor beleuchteten Verfahren in Abschnitt 2.1.1 ergänzt. So wird bspw. in [59] ein Ansatz vorgestellt, mit welchem probabilistische Lastgänge mit dem Ziel erzeugt werden, Spitzenleistungen bzw. Gleichzeitigkeitsfaktoren abzuleiten. Dies erfolgt unter Berücksichtigung von iMSys und soll insbesondere bei der Dimensionierung der Netze unterstützen. Die Autoren weisen darauf hin, dass im Rahmen des konventionellen Planungsprozesses der VNB neben den Gleichzeitigkeitsfaktoren auch Flächenlasten auf Grundlage von praktischen Erfahrungen einfließen. Diese fehlen letztlich bei der Berücksichtigung und Anbindung neuer Akteure, wie Ladesäulen für Elektroautos oder dezentrale erneuerbare Energiequellen. Diese Lücke füllen die Autoren mit dem vorgestellten probabilistischen Ansatz.

Insbesondere mit dem Einzug von rechnergestützten Systemen und der Verknüpfung von Datenbanken sowie Geoinformationssystemen stehen den EVU und VNB eine Vielzahl von unterschiedlichsten Informationen zur Verfügung, um im Rahmen von Netzplanung und -führung höherwertige Aussagen treffen zu können. Ergänzt durch Messdaten von iMSys sowie dem Einbeziehen weiterer sozio-ökonomischer und -demografischer Daten stellen probabilistische Ansätze eine Alternative zur Abschätzung des Verbraucherverhaltens dar. Dies wird jedoch nur durch eine erheblich größere Informationsdichte und dem damit einhergehenden Aufwand möglich.

Die probabilistische Verbraucherlastmodellierung zielt im Allgemeinen auf die Berücksichtigung der individuellen statistischen Systematik der Endkunden ab und wird i.d.R. zur Modellierung synthetischer Lastprofile genutzt [73]. Der Grundgedanke des Ansatzes ist die Modellierung der Verbraucherlast durch Zufallsvariablen (ZV) mit beliebiger Wahrscheinlichkeitsdichtefunktion (WDF) [51]. Die eigentliche Lastapproximation erfolgt letztlich mit Hilfe einer Ziehung der ZV $\mathcal{P}_t(P)$ der WDF $f : \mathbb{R} \rightarrow [0|\infty)$ folgend, stellvertretend für die Wirkleistungsaufnahme P zum Zeitpunkt $t \in \mathcal{T}$ mit $\mathcal{T} = \{t_1, t_2, \dots, t_n\}$. Die Wahrscheinlichkeit \Pr für das Auftreten der Wirkleistung P im Intervall $[P_{\min}, P_{\max}] := \{P \in \mathbb{R} \mid P_{\min} \leq P \leq P_{\max}\}$ ist durch

$$\Pr(P_{\min} \leq \mathcal{P}_t \leq P_{\max}) = \int_{P_{\min}}^{P_{\max}} f_{\mathcal{P}_t}(P) dP \quad (2.3)$$

beschrieben [51]. Hierdurch ist die Generierung synthetischer Lastprofile einzelner Verbraucher möglich, welche in einem nachfolgenden Prozess, bspw. im Rahmen einer Lastflussberechnung, ausgewertet werden. Dieser Schritt ist insbesondere für die Entwicklung neuartiger Netzführungsstrategien notwendig, um diese mit Hilfe einer Netzsimulation zu erproben [73].

Hierbei handelt es sich keineswegs um einen neuen Ansatz. So wurde bereits 1979 in [74] die Modellierung des Leistungsbedarfs individueller Haushalte mit Wahrscheinlichkeitsdichten und -funktionen thematisiert. Im Zuge der Publikation zog Piller die Messdaten einer großangelegten messtechnischen Untersuchung der Technischen Universität München heran, welche u.a. die Mess- und Analysedaten eines Wohnblocks mit 44 Wohneinheiten über einen längeren Zeitraum berücksichtigt. Der Abbildung 2.3 ist eine Auswertung der Messdaten nach [74] zu entnehmen. Visualisiert sind zwei Zeitreihen sowie eine stetige Verteilungsfunktion. Die Zeitreihe in blau stellt den Leistungsmittelwert je Viertelstunde dar, während die Zeitreihe in gelb die Konfidenzgrenze darstellt, unterhalb derer 97,7% aller im statistischen Mittel aufgetretenen Leistungen liegen [74]. In rot ist exemplarisch das Leistungswahrscheinlichkeitsprofil zum Zeitpunkt $t = 14:00$ Uhr dargestellt, welches die Unsicherheiten des Prozesses mit Hilfe der Auftrittswahrscheinlichkeit einzelner Leistungswerte zum jeweiligen Zeitpunkt berücksichtigt. Pillers Überlegungen dienen im Rahmen der vorliegenden Arbeit als Basis der probabilistischen Verbraucherlastmodellierung.

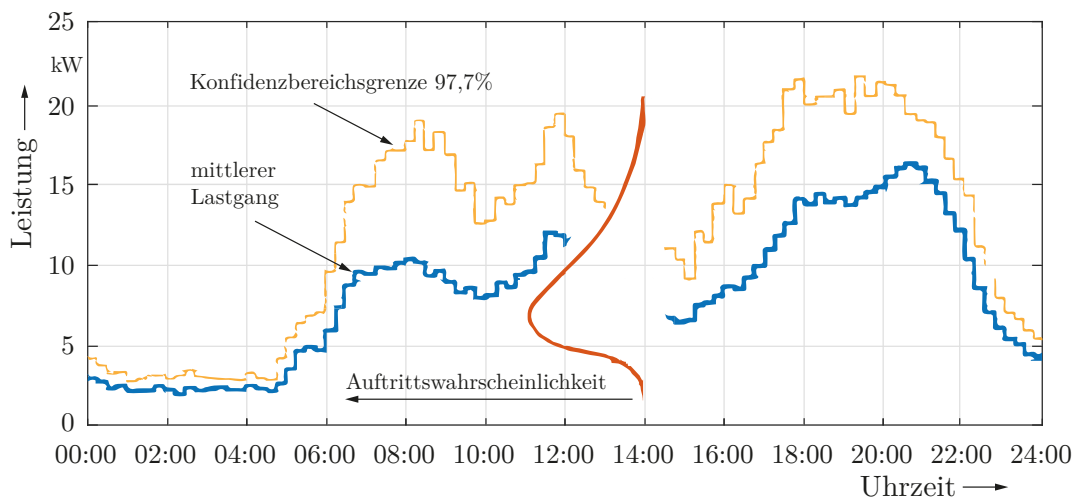


Abbildung 2.3.: Auswertung des Leistungsbedarfs eines Wohnblocks mit 44 Wohneinheiten. Dargestellt ist der Leistungsmittelwert (blau), die Konfidenzbereichsgrenze (gelb) sowie das Leistungswahrscheinlichkeitsprofil eines Modellfalls zum Zeitpunkt $t = 14:00$ Uhr (eigene Darstellung nach [74]).

In der Literatur sind eine Vielzahl an unterschiedlichen Ansätzen mit unterschiedlicher Modellierungstiefe zu finden, welche dabei helfen sollen die Netzdynamik zu modellieren. Diese Ansätze können im Wesentlichen in zwei Gruppen unterteilt werden [73]:

1. *Top-Down*- Verfahren sowie
2. *Bottom-Up*- Verfahren.

Piller beschreibt in [74] eine Betrachtung der Mess- und Analysedaten nach dem *Top-*

Down- Verfahren (TDV), weist jedoch zugleich auf die Möglichkeit hin eine Modellierung einzelner elektr. Geräte oder Gerätegruppen zur besseren Approximation der Vorgänge in Betracht zu ziehen, was wiederum dem *Bottom-Up*-Verfahren (BUV) gleicht. Abbildung 2.4 soll die wesentlichen Unterschiede beider Verfahren verdeutlichen. Der schematischen Darstellung kann entnommen werden, dass beim BUV die Modellierung aus Sicht des Netzes am *untersten Ende* (engl. bottom) beginnt, bei der Modellierung einzelner elektr. Geräte, um die Prozesse *darüberliegender* Ebenen (engl. up) zu untersuchen. Dies steht im Gegensatz zum TDV, welches mit Hilfe aggregierter Lastverläufe Rückschlüsse von einer *höheren* Netzebene (engl. top) auf die Gesetzmäßigkeit einer *tieferen* Ebene (engl. down) zulässt [51].

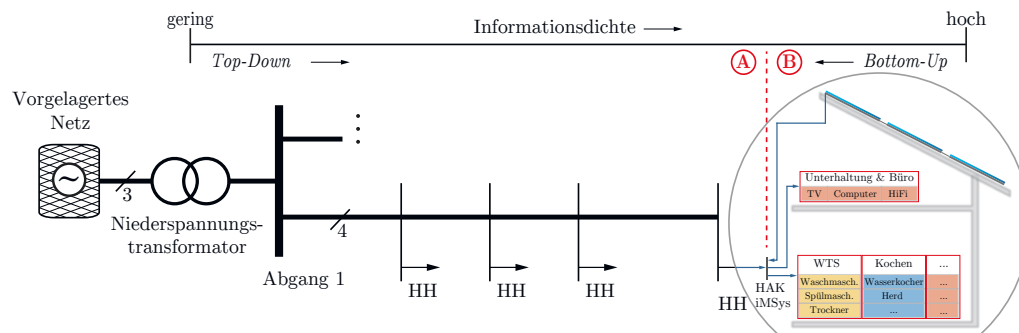


Abbildung 2.4.: Geltungsbereich des BUV und TDV. Zudem ist die ansteigende Informationsdichte vom TDV zum BUV (qualitativ) zu entnehmen.

Der Übergang beider Verfahren ist dabei nicht eindeutig. So wird in [51] sowie in [75] ein Ansatz zur Modellierung eines Haushaltes am Hausanschlusskasten (HAK) vorgestellt, jedoch stuft letzterer den Ansatz als BUV ein, während in [51] dieser unter dem TDV verortet wird (vgl. Markierung A und B in Abbildung 2.4). Festzuhalten bleibt, dass die Ansätze nach [51] und [75] sich dahingehend gleichen, dass am HAK statistisch aufbereitete Lastprofile für die nachfolgende Leistungsflussberechnung genutzt werden.

Folglich gilt es im Rahmen der Modellierung stochastischer Vorgänge im Netz geeignete WDF sowie deren Parameter zu identifizieren. Erst hierdurch kann im nachfolgenden Prozess ein synthetisches Lastprofil zur weiteren Analyse abgeleitet werden. Im Rahmen der vorliegenden Arbeit werden zwei neuartige Ansätze zur Generierung synthetischer Verbraucherlastprofile nach dem TDV sowie dem BUV unter den Prämissen, die notwendige Informationsdichte zu minimieren sowie einer *energietreuen* Modellierung, vorgestellt. Die einzelnen Modellierungsschritte sowie deren Rahmenbedingungen zur Generierung synthetischer Verbraucherlastprofile werden nachfolgend im Kontext beider Verfahren kurz vorgestellt und anhand eines Beispiels verdeutlicht.

2.1.2.1. *Top-Down*-Verfahren

Das TDV ist an Pillers Überlegungen in [74] angelehnt. Mit Bezug auf Abbildung 2.3 ist für die Leistungsmodellierung der Verbraucherlast ein *mittlerer Lastgang* sowie ein Streumaß vonnöten. Die eigentliche Nachbildung des elektr. Leistungsverhaltens erfolgt auf Basis

probabilistischer Gesetzmäßigkeiten und deren (zeitvarianten) Parametern. Die Auftrittswahrscheinlichkeit einzelner Ereignisse zum jeweiligen Zeitpunkt t wird mit Hilfe einer WDF modelliert. Die Bestimmung der (zeitvarianten) Parameter erfolgt bspw. mit Hilfe der Momentenmethode [76] oder der Maximum-Likelihood-Methode [75]. Eine ausführliche Zusammenfassung und Zuordnung hypothetischer Verteilungsfunktionen zur statistischen Nachbildung des elektr. Leistungsbedarfs kann der Arbeit nach [75] und [77] entnommen werden.

Zur synthetischen Generierung der Verbraucherlast eines einzelnen Endkunden wird im Rahmen der vorliegenden Arbeit der *mittlere Lastgang* mittels SLP approximiert. Dieses stellt den definierten Erwartungswert $E(\mathcal{X})$ dar und kann über den elektr. JEB sowie den Dynamisierungsfaktor im betrachteten Zeitraum beliebig skaliert werden (vgl. Abschnitt 2.1.1). Hierdurch wird gewährleistet, dass die einzelnen, synthetisch generierten Lastprofile gegen den vorgegebenen Erwartungswert respektive SLP tendieren (vgl. *das Gesetz der großen Zahlen* der Stochastik [78]).

Die Bestimmung einer sinnvollen Größenordnung des Streumaßes respektive der Standardabweichung $\sigma_{\mathcal{X}}$ wird hingegen in [76] als problematisch beschrieben und erfolgt im Rahmen der vorliegenden Arbeit auf Basis statistischer Erhebungen. In [76] und [77] werden ländlich gelegene Haushalts-Endkunden im Zuge einer Messstudie statistisch ausgewertet. Dabei wird die Standardabweichung $\sigma_{\mathcal{X}}$ relativ zum Erwartungswert $E(\mathcal{X})$ mit $\sigma_{\mathcal{X},rel}$ gebildet und ist abhängig vom jeweiligen Zeitpunkt t und der Anzahl der betrachteten Haushalte N . Zur abschließenden Bestimmung der notwendigen Standardabweichung des TDV werden die Daten der statistischen Auswertung der Messdatenerhebung nach [76] und [77] im Querschnitt bewertet und extrapoliert (vgl. Anhang A.3). Das beschriebene Verfahren resultiert in einer Generierung synthetischer Lastprofile einzelner Haushalts-Endkunden.

Generierung synthetischer Haushaltslastprofile

Nachfolgend wird der durch ein Netz mit überwiegend ländlichen Haushalts-Endkunden verursachte tägliche Leistungsbedarf exemplarisch modelliert. Hierfür wird die probabilistische Gesetzmäßigkeit über eine Gamma-WDF $\mathcal{G} : \mathbb{R}^+ \rightarrow [0|\infty)$ approximiert (vgl. [77]). Dabei gilt für die ZV $\mathcal{X}(x) \sim \mathcal{G}(\alpha(t), \beta(t))$ mit den Form- und Skalierungsparametern [76]

$$\begin{aligned} \alpha(t) &= E(\mathcal{X}(t))^2 / \sigma_{\mathcal{X}}(t)^2 \in \mathbb{R}^+ \\ \beta(t) &= \sigma_{\mathcal{X}}(t)^2 / E(\mathcal{X}(t)) \in \mathbb{R}^+ \end{aligned} \tag{2.4}$$

sowie der Gamma-Funktion Γ die WDF [51]

$$f_{\mathcal{X}}(x; \alpha(t), \beta(t)) = \frac{x^{\alpha(t)-1}}{\beta(t)^{\alpha(t)} \cdot \Gamma(\alpha(t))} \cdot \exp\left(-\frac{x}{\beta(t)}\right) \tag{2.5}$$

mit $x \geq 0$.

Ausgehend von einem *mittleren Lastgang* der Wirkleistung unter Anwendung des SLP der Kundengruppe *Haushalt* (Profiltyp H0) wird für jeden Simulationszeitpunkt t ein Parametersatz bestimmt. Hierfür wird zunächst der Erwartungswert $E(\mathcal{P}_t)$ der zu modellierenden WDF auf Basis des vorgegebenen SLP ermittelt und als Gesamtknotenleistung für N Haushalts-Endkunden angenommen. Die Fluktuation einzelner Realisierungen der Haushalts-Endkunden wird über die Angabe der tatsächlichen Standardabweichung mit $\sigma_{\mathcal{P}_t} = \sigma_{\mathcal{P}_t, \text{rel}} \cdot E(\mathcal{P}_t)$ berücksichtigt. Damit liegen die nötigen Informationen zur Bestimmung der Parameter aus den ersten beiden statistischen Momenten vor, sodass unter Anwendung der hieraus resultierenden zeitvarianten WDF nach Gleichung 2.5 die eigentliche Lastapproximation der Wirkleistung $P(t)$ mit $\mathcal{P}_t(P) \sim \mathcal{G}(\alpha(t), \beta(t))$ zum Zeitpunkt t erfolgt.

Per Definition wird zur Berechnung der Blindleistung zufällig ein zugehöriger Leistungsfaktor einer Gleichverteilung \mathcal{U} folgend gezogen. In Anlehnung an technische Anschlussbedingungen „Niederspannung“ (TAB) diverser VNB im deutschen Versorgungsgebiet erfolgt im Rahmen der vorliegenden Arbeit die Definition des Leistungsfaktors mit $\cos(\varphi) > 0,9$ (kapazitiv oder induktiv). Laut TAB sind kleinere Leistungsfaktoren zu vermeiden bzw. sehen die VNB i.d.R. Einrichtungen zur Blindleistungskompensation vor. Zudem wird ein max. Leistungsfaktor von 0,98 angenommen (vgl. [58]), sodass der resultierende Leistungsfaktor $\cos(\varphi) \in \mathbb{R}$ im Intervall $\cos(\varphi) = [0,9; 0,98]$ modelliert wird. Darüber hinaus bietet das vorgestellte Verfahren eine optionale Verteilung der Leistungswerte auf drei Außenleiter $l = \{L1, L2, L3\}$. Eine Übersicht des beschriebenen Prozesses ist dem Algorithmus im Anhang A.4 zu entnehmen.

Verifizierung eines synthetisch erzeugten Lastprofils

Zur Verifizierung des vorgestellten Verfahrens soll Abbildung 2.5 herangezogen werden. Dem oberen Diagramm sind drei Realisationen zu entnehmen. Es ist eine deutliche Fluktuation der Leistung der einzelnen Lastprofile zu erkennen. Zudem sind den Realisationen unterschiedliche Bandbreiten im Tagesverlauf zu entnehmen, welche den Beobachtungen nach [74] folgen. So ist während der Zeiten mit individuell stark unterschiedlichem Verbraucherverhalten wie morgens, mittags und abends mit deutlich höheren Bandbreiten zu rechnen. Die unterschiedlichen Verhaltensmuster und Gewohnheiten der Haushalts-Endkunden korrelieren in der Modellierung mit einer weit größeren Bandbreite respektive Unsicherheit. Dem gegenüber stehen die Nachtstunden mit weitgehend gleichem Verhalten der Verbraucher.

Dem unteren Diagramm der Abbildung 2.5 ist die zugrundeliegende Modellprämisse einer *energie-treuen* Modellierung zu entnehmen. Hierfür soll beispielhaft das SLP des Profiltyps H0 verwendet und im Zuge der Modellierung als Erwartungswert interpretiert werden. Dem *Gesetz der großen Zahlen* der Stochastik folgend muss der Mittelwert einzelner Realisationen mit steigender Anzahl an Ziehungen im statistischen Mittel gegen den Erwartungswert streben [78]. Gleiches gilt für die aufgespannte Fläche eines synthetisch erzeugten Lastpro-

fils nach dem TDV, welches dem modellierten elektr. Energiebedarf eines Tages d im Jahr entspricht. Folglich muss der Mittelwert des elektr. Energiebedarfs eines Tages d einzelner, synthetisch erzeugter Lastprofile (Realisationen) gegen den elektr. Energiebedarf des Tages d des vorgegebenen SLP (Erwartungswert) streben. Dieser Sachverhalt ist dem Histogramm der Abbildung 2.5 zu entnehmen. Dargestellt ist der tägliche Energiebedarf in kWh von 300 synthetisch erzeugten Verbraucherlastprofile einzelner Haushalts-Endkunden. In rot markiert ist der Erwartungswert sowie der Mittelwert der 300 Realisationen in grün. Die Abweichung zum Erwartungswert liegt bei $< 1\%$. Dies unterstreicht die *Energietreue* des vorgestellten Verfahrens.

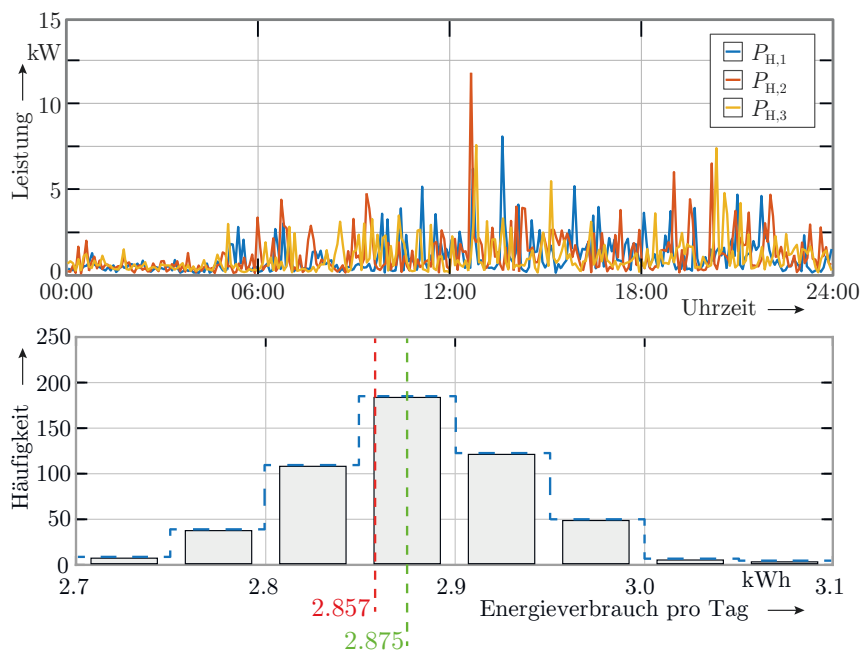


Abbildung 2.5.: (oben) Exemplarisch dargestellt sind drei Ziehungen der ZV, welche im Rahmen des TDV dem Leistungsbedarf eines einzelnen Haushalts-Endkunden entspricht. (unten) Dem Histogramm ist der tägliche Energiebedarf in kWh von 300 Realisationen synthetisch erzeugter Verbraucherlastprofile einzelner Haushalts-Endkunden zu entnehmen. Hervorgehoben ist der Erwartungswert (rot) und der Mittelwert der 300 Realisationen (grün). Die Abweichung zum Erwartungswert liegt bei $< 1\%$.

Resümee

Ein wesentlicher Vorteil des hier vorgestellten TDV zu seinem Pendant dem BUV ist die im Vergleich geringe Anzahl an Informationen. Zudem bietet der Ansatz die Möglichkeit Messdaten iMSys zur Berücksichtigung regionaler Einflüsse direkt in die Modellierung zu überführen [70]. Lediglich eine Extrahierung der notwendigen Informationen mit Hilfe der Ansätze in [75] oder [77] gehen dem voraus.

Die Verifizierung synthetisch erzeugter Lastprofile belegt die *Energietreue* des vorgestellten Verfahrens. Zudem können der Abbildung 2.5 im Allgemeinen realitätsnahe Größenordnung entnommen werden (vgl. [77]).

2.1.2.2. *Bottom-Up-Verfahren*

Piller weist in [74] auf die Möglichkeit hin, den individuellen Leistungsbedarf einzelner Wohneinheiten durch die Modellierung einzelner elektr. Geräte oder Gerätegruppen innerhalb der Wohneinheiten zu approximieren, und unterstellt dieser Herangehensweise eine höhere Güte der Modellierung. Heutzutage ist dieser Ansatz in der einschlägigen Literatur unter *Bottom-Up* klassifiziert und wird bspw. in den Arbeiten nach [57], [58], [79] sowie [80] verfolgt. Ein wesentlicher Beitrag zur Verbraucherlastmodellierung nach dem BUV kann der Arbeit nach [80] zugesprochen werden. Aufbauend auf den Erkenntnissen zur Berücksichtigung der Korrelation psychologischer Faktoren einzelner Individuen auf den Energieverbrauch von Haushalts-Endkunden mit Hilfe von WDF in [81] erarbeiten Capasso et al. ein Simulationsmodell, welches den Einfluss einer dynamischen Tarifstruktur im Kontext des Lastmanagements analysiert. Das von den Autoren erarbeitete BUV berücksichtigt sozio-demografische Faktoren und projiziert das Verhalten einzelner Individuen auf den elektr. Energiebedarf auf Ebene einzelner elektr. Geräte.

Dem vorgestellten komplexen Modell kann der Einfluss nicht-elekt. Faktoren sowie die Vielzahl und das Spektrum an notwendigen Informationen zur Verbraucherlastmodellierung nach dem BUV entnommen werden. Eine Zusammenfassung wiss. Arbeiten im Kontext der Lastapproximation mit Hilfe des BUV kann [54], [73], sowie [82] entnommen werden.

Im Allgemeinen können die BUV dahingehend voneinander differenziert werden, ob

- der resultierende Leistungsbezug einer Gerätegruppe durch Überlagerung und den Grad der Durchdringung einzelner elektr. Geräte bestimmt wird (vgl. [79] sowie [80]) oder
- die Gruppe ganzheitlich unter Berücksichtigung der Gruppen zugehörigen elektr. Geräte und deren Durchdringung im Haushalt modelliert wird (vgl. [57] sowie [58]).

Letzteres stellt eine höhere Abstraktionsstufe und demnach ein geringeres Informationslevel sowie einen geringeren Modellierungsaufwand dar und findet daher in der vorliegenden Arbeit Anwendung.

Basierend auf den Annahmen und Erkenntnissen des zuvor beschriebenen TDV nach Abschnitt 2.1.2.1 sollen mit Hilfe des BUV Prozesse *hinter* dem HAK beschrieben werden (vgl. Markierung B in Abbildung 2.4). Hierbei liegen dem BUV die Überlegungen nach [57] sowie [83] zugrunde. Das Verfahren basiert auf drei Modellierungsschritten und beschreibt je Gerätegruppe den

1. elektr. Leistungsbezug,
2. den Nutzungszeitpunkt, sowie die
3. Nutzungsdauer.

Hierbei wird in Anlehnung an die Publikation [79] und [84] den Gerätegruppen eine der folgenden Nutzungsarten zugeordnet:

1. Grundlast: Dauerbetrieb.
2. Thermostat betrieben: Automatikbetrieb.
3. Regelmäßig betrieben: benutzerabhängiger Betrieb.

Im Allgemeinen liegen dem vorgestellten BUV spezifische statistische Informationen zugrunde. Die Reduzierung der Informationsdichte erfolgt in Anlehnung an die Überlegungen in [57] mit Hilfe von Gerätegruppen. Die Definition und Datenlage der einzelnen Schritte werden nachfolgend jeweils im Kontext der drei Nutzungsarten vorgestellt. Anschließend wird das Verfahren verifiziert und im Vergleich mit realen Messdaten validiert.

Gerätegruppen und deren elektr. Leistungsbezug

Das hier vorgestellte BUV ist von Ansätzen der einschlägigen Literatur dahingehend zu unterscheiden, dass bei der Modellierung ein *energetreuer* Ansatz verfolgt wird. Im Zentrum des gewählten Modellierungsansatzes steht dem zu Folge der JEB. Dieser liegt dem VNB i.d.R. vor. Sollte dies nicht der Fall sein, kann auf eine statistische Referenz wie bspw. in Tabelle 2.3 Teil (a) zurückgegriffen werden. Zur Berücksichtigung meteorologischer Größen, hierbei primär der Temperatur, kann darüber hinaus der Dynamisierungsfaktor nach Gleichung 2.2 zur Skalierung des JEB hinzugezogen werden.

Tabelle 2.3.: (a) Mittlerer JEB deutscher Haushalte (vollelektrisch) [85]. (b) Definition der Gerätegruppen und Aufteilung des Stromverbrauchs privater Haushalte [85].

(a)		(b)	
ANZAHL PERSONEN ^a	JEB ^b IN kWh	GERÄTEGRUPPE	ANTEIL IN %
1	2050	Kühleinheit	16.7
2	3440	Unterhaltung & Büro	25.7
3	4050	Warmwasseraufbereitung	14.8
4	4750	WTS ^a	12.4
≥ 5	5370	Kochen	9.8
		Beleuchtung	8.1
		Sonstige Elektrogeräte	12.5

^aAnzahl der im Haushalt lebenden Personen

^bJahresenergiebedarf

^aWaschmaschine, Trockner und Spülmaschine

Grundlegend erfolgt die Modellierung des elektr. Leistungsbezugs mit Hilfe einer stetigen Verteilungsfunktion

$$\mathcal{P}_{gg} \sim \mathcal{N}(\mu_{gg}, \sigma_{gg}^2) \tag{2.6}$$

je Gerätegruppe gg . Der Erwartungswert μ_{gg} wird mit Hilfe des JEB modelliert. Die Gruppierung elektr. Verbraucher in Gerätegruppen folgt der Einteilung in [85], dargestellt in Tabelle 2.3 Teil (b).

Darüber hinaus wird im Rahmen des hier vorgestellten BUV die Gerätegruppe Stand-By berücksichtigt, welche eine Grundlast gleichverteilt über den Tag darstellt. Dabei wird der Grundlast eines Haushalts-Endkunden ein Anteil von 11 % bezogen auf den gesamt JEB

zugesprochen [58], welcher zu gleichen Teilen den Gerätegruppen der Tabelle 2.3 Teil (b) zugeordnet wird. Charakteristisch für diese Gerätegruppe ist ein konstanter elektr. Leistungsbezug, welcher unabhängig von der aktiven Nutzung des Endkunden ist. Daher wird dieser dauerhafte Betrieb der Nutzungsart 1 zugeordnet. Unter Berücksichtigung des Dynamisierungsfaktors nach Gleichung 2.2 kann abschließend der Erwartungswert $\mu_{\text{Stand-By}}$ für einen beliebigen Tag d des Jahres bestimmt werden.

Das Lastprofil der Gerätegruppe Kühleinheit, beschrieben durch Nutzungsart 2, zeigt einen getakteten Verlauf auf, welcher einen vom Benutzer unabhängigen, dauerhaften Leistungsbezug darstellt. Basierend auf dem JEB des jeweiligen Haushalts-Endkunden kann nach Tabelle 2.3 Teil (b) zunächst ein mittlerer, konstanter Leistungsbezug bestimmt werden. Der Arbeit [58] ist zu entnehmen, dass Kühl- und Gefriergeräte im getakteten Betrieb Zykluszeiten von $T_{\text{Kühleinheit}} = [60, 100]$ min aufweisen, wovon $\Delta t_{\text{Kühleinheit,ein,rel}} = [25, 40]$ % auf die Einschaltdauer zurückfallen. Unter Einbeziehung dieser Daten kann abschließend ein getaktetes Lastprofil abgeleitet und der Gerätegruppe ein Erwartungswert $\mu_{\text{Kühleinheit},t}$ zugeordnet werden. Sowohl die Ziehung der Zykluszeiten $T_{\text{Kühleinheit}} \sim \mathcal{U}$ als auch die der relativen Einschaltdauer $\Delta t_{\text{Kühleinheit,ein,rel}} \sim \mathcal{U}$ folgen einer Gleichverteilung \mathcal{U} .

Mit Bezug auf die Nutzungsart 3 korreliert die Positionierung des elektr. Leistungsbezugs auf der Zeitachse mit einer bestimmten Tätigkeit. Als Beispiel sei an dieser Stelle Kochen nach Tabelle 2.3 Teil (b) mit 9.8 % des JEB genannt. Hierbei ist es belanglos welches elektr. Gerät der Gerätegruppe genutzt wird. Die korrespondierende statistische Referenz einer Gerätegruppe wird im Zuge des Verfahrens in einzelne *Energieblöcke* zerlegt und in Abhängigkeit des Nutzungszeitpunktes sowie der Nutzungsdauer über den Tag verteilt. Um die *Energietreue* des Ansatzes sicherzustellen, wird in Abhängigkeit der mittleren Anzahl der *Energieblöcke*, bestimmt durch die Nutzungszeitpunkte (siehe unten), sowie dem Erwartungswert der Nutzungsdauer (siehe unten) ein *Energieblock* als Repräsentant für die Tätigkeit Kochen mit einem mittleren elektr. Leistungsbezug von $\mu_{\text{Kochen},t}$ bestimmt.

Neben der Definition des Erwartungswertes μ_{gg} einer Gerätegruppe gg sieht die Modellierung nach Gleichung 2.6 eine Varianz σ_{gg}^2 vor. Die Grundlage für die Beschreibung des Streumaßes stellen die in [72] erhobenen Daten dar. Diese werden analog zum TDV umgerechnet, sodass die Beschreibung über eine relative Standardabweichung $\sigma_{\text{rel},gg}$ je Gerätegruppe gg erfolgt. Hierfür wird zunächst die Standardabweichung σ_{gg} über die Summe der gewichteten Varianzen aller elektr. Verbraucher v der Gerätegruppe gg nach

$$\sigma_{gg} = \sqrt{\sum_i \frac{\bar{P}_{gg,v_i}}{\bar{P}_{gg}} \cdot \sigma_{gg,v_i}^2}, \quad (2.7)$$

mit der Summe der mittleren Leistung $\bar{P}_{gg} = \sum_i \bar{P}_{gg,v_i}$ aller Verbraucher v der Gerätegruppe gg , bestimmt. Abschließend kann die relative Standardabweichung durch

$$\sigma_{\text{rel},gg} = \frac{\sigma_{gg}}{\bar{P}_{gg}} \quad (2.8)$$

berechnet werden. Für die Gerätegruppe Stand-By wird $\sigma_{\text{rel, Stand-By}} = 0.1$ festgelegt. Alle weiteren abgeleiteten Werte je Gerätegruppe können der Tabelle 2.4 entnommen werden.

Der Blindleistungsbezug wird über die Vorgabe eines Intervalls des Leistungsfaktors $\cos(\varphi)$ modelliert und mit der Realisation der Leistungsaufnahme P_{gg} einer Gerätegruppe gg berechnet. Die vorgegebenen Werte des Leistungsfaktors $\cos(\varphi)$ der Tabelle 2.4 entsprechen den Daten nach [58]. Darüber hinaus kann den Gerätegruppen optional ein Außenleiter $l \in \{L1, L2, L3\}$ zugeordnet werden. Hierfür wird der elektr. Leistungsbezug $P_{gg,l}$ der ZV $\mathcal{P}_{gg} \sim \mathcal{N}(\mu_{gg}, \sigma_{gg}^2)$ einer Gerätegruppe gg folgend entweder gleichmäßig auf alle drei Außenleiter l verteilt oder einem einzelnen Außenleiter zugeordnet. Letzteres ist dabei nicht auf die Zuordnung auf einen einzelnen Außenleiter beschränkt. Vielmehr kann vorgesehen werden, dass der zugeordnete Außenleiter l einer Gerätegruppe bei der Generierung eines neuen *Energieblocks* als Repräsentant für eine Tätigkeit alterniert. Die jeweilige Zuordnung der Außenleiter je Gerätegruppe kann der Zeile *Anzahl Phasen* der Tabelle 2.4 entnommen werden. Die beschriebene Option einer neuen Zuordnung des Außenleiters ist hervorgehoben.

Tabelle 2.4.: Datengrundlage des Verbraucherlastmodells nach dem BUV.

KLASSIFIKATION	JAHRESARBEIT ANTEIL (%)	ZEITBEREICH			ANZAHL PHASEN	$\cos(\varphi)$	σ_{rel}	μ_{DOU}	σ_{DOU}
		MORGEN	MITTAG	ABEND					
Kühltheit	16.7		Automatik		1	0.85...0.95	0.3	∞	
Unterhaltung & Büro	25.7	-	✓	✓	1 ^a	0.5...0.85	0.45	75	43
Warmwasseraufbereitung	14.8	✓	✓	✓	3	>0.95	0.21	Referenzprofil ^b	
WTS ^c	12.4	✓	✓	✓	1	0.7...0.95	0.1	98	43
Kochen	9.8	✓	✓	✓	3	>0.95	0.1	43	13
Beleuchtung	8.1	✓	-	✓	1 ^a	>0.95	1.1	150	90
Sonstiges	12.5	✓	✓	✓	1 ^a	>0.7	0.17	59	27
Stand-By ^d	11.0		Dauerbetrieb		3	0.2...0.85	0.1	∞	

^aDie verwendete Phase alterniert

^bnach Richtlinie VDI 4655

^cWaschmaschine, Trockner und Spülmaschine

^dDer Anteil bezieht sich auf die gesamte Jahresarbeit

Nutzungszeitpunkt

Die Erkenntnisse aus [74] zeigen deutlich, dass nicht nur elektr. Größen bei der Bewertung elektr. Lastprofile Anwendung finden sollten. Der Autor verweist dabei insbesondere auf ein stark individuelles Verbrauchsverhalten der einzelnen Individuen am Morgen, Mittag und Abend. Dieses resultiert in einer Variation des Energiebedarfs und letztlich zu einem ausgeprägten Lastprofil (vgl. Abbildung 2.3). Auch in [80] und [81] wird darauf verwiesen, dass sozio-ökonomische und -demographische Faktoren einen direkten Einfluss auf die Ausprägung des Leistungsbezuges haben. Dieses individuelle Verbrauchsverhalten der einzelnen Individuen trifft für die erste als auch die zweite Nutzungsart nicht zu, da beide unabhängig vom Benutzer betrieben werden. Somit müssen diese keinem bestimmten Zeitbereich zugeordnet werden und sind in Tabelle 2.4 mit Dauerbetrieb für Nutzungsart eins respektive Automatik für Nutzungsart zwei gekennzeichnet. Lediglich die Nutzungsart drei ist direkt vom Benutzer abhängig und unterliegt einer adäquaten Modellierung des Nutzungszeitpunkts.

Dem im Rahmen der vorliegenden Arbeit vorgestellten BUV ist ein Ansatz zur Modellierung des Nutzungszeitpunkts zu entnehmen, welcher mit drei unabhängigen Normalverteilungen

$$\mathcal{T}_{\text{Zeitbereich}} \sim \mathcal{N}(\mu_{\text{Zeitbereich}}, \sigma_{\text{Zeitbereich}}^2) \quad (2.9)$$

für $\text{Zeitbereich} \in \{\text{Morgen}, \text{Mittag}, \text{Abend}\}$ entlang der Zeitachse eine *Zeitschablone* definiert. Die je *Zeitbereich* resultierende Realisation $t_{\text{Zeitbereich}}$ folgt der ZV $\mathcal{T}_{\text{Zeitbereich}}$ und repräsentiert den Nutzungszeitpunkt (aus dem engl. time of use; TOU). In der Abbildung 2.6 sind die stetigen WDF \mathcal{T}_{TOU} für $\text{Zeitbereich} \in \{\text{Morgen}, \text{Mittag}, \text{Abend}\}$ des Tages d dargestellt. Der Abbildung ist eine Überlappung der einzelnen WDF zu entnehmen. Des Weiteren ist für den Zeitbereich $\mathcal{T}_{\text{Abend}}$ eine nicht unerhebliche Wahrscheinlichkeit für eine Realisation im darauffolgenden Tag $d + 1$ des jeweiligen Jahres zu entnehmen.

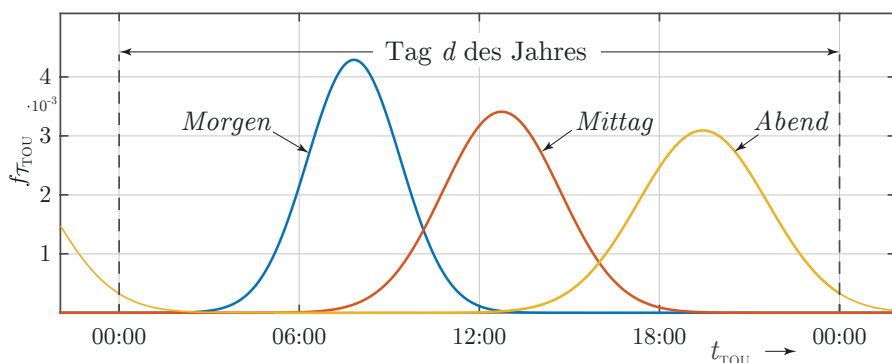


Abbildung 2.6.: Beispielhafte Zeitschablone eines Tages d im Jahr, modelliert mit Hilfe von drei unabhängigen Normalverteilungen $\mathcal{T}_{\text{Zeitbereich}} \sim \mathcal{N}(\mu_{\text{Zeitbereich}}, \sigma_{\text{Zeitbereich}}^2)$ für $\text{Zeitbereich} \in \{\text{Morgen}, \text{Mittag}, \text{Abend}\}$.

Die zuvor definierten Gerätegruppen können allgemein in Anlehnung an [86] dem zugehörigen $\text{Zeitbereich} \in \{\text{Morgen}, \text{Mittag}, \text{Abend}\}$ zugeordnet werden. Für die Gerätegruppen WTS und Beleuchtung wurden die Informationen nach [87] berücksichtigt. Die Gerätegruppe Beleuchtung wird dabei gemäß den Ergebnissen in [87] für die Zeitbereiche *Morgen* und *Abend* modelliert. Die Autoren verweisen darauf, dass die Beleuchtung einer erheblichen saisonalen Schwankung unterliegt und nutzen im Rahmen der Modellierung die globale Beleuchtungsstärke, welche mit einer minimalen Beleuchtungsstärke je Haushalt verglichen wird. Ein ähnlicher Ansatz wird u.a. in [72] verfolgt.

Für die Warmwasseraufbereitung erfolgt die Einordnung der Europäischen Norm *Heizungsanlagen in Gebäuden – Planung von Heizungsanlagen mit Wärmepumpen* (DIN EN 15450) nach, mit 35 % am Morgen, 20 % am Mittag und der Mehrheit von 45 % in den Abendstunden. Demnach ist eine Mehrfachbelegung einer Gerätegruppe über mehr als einen *Zeitbereich* möglich und wird bei der Bestimmung des elektr. Leistungsbedarfs berücksichtigt. Der Tabelle 2.4 ist eine Zusammenfassung der beschriebenen Nutzungszeitpunkte je Gerätegruppe zu entnehmen.

Nutzungsdauer

Sowohl die erste als auch die zweite Nutzungsart ist unabhängig von einer individuellen Nutzungsdauer, da beide unabhängig vom Benutzer betrieben werden. Die Nutzungsdauer beider Gerätegruppen wird in Tabelle 2.4 plakativ mit ∞ angegeben.

Um die zuvor beschriebenen Unsicherheiten des Verhaltens eines einzelnen Individuums zu modellieren, welche auch in einer variablen Nutzungsdauer einzelner Gerätegruppen resultieren, wird in Anlehnung an [83] die Nutzungsdauer der Gerätegruppen der Nutzungsart drei mit Hilfe einer Normalverteilung abstrahiert. Demnach wird die Nutzungsdauer (aus dem engl. duration of use; DOU) über die Realisation $t_{gg,DOU} \sim \mathcal{T}_{gg,DOU}$ einer Gerätegruppe gg in Minuten angegeben und folgt der ZV

$$\mathcal{T}_{gg,DOU} \sim \mathcal{N}(\mu_{gg,DOU}, \sigma_{gg,DOU}^2) . \quad (2.10)$$

Die Grundlage zur Ableitung der Parameter $\mu_{gg,DOU}$ und $\sigma_{gg,DOU}^2$ der Normalverteilungen $\mathcal{T}_{gg,DOU}$ der jeweiligen Gerätegruppe basieren auf spezifischen statistischen Daten. Der statistischen Erhebung nach [88] können die Daten zur Ableitung der DOU der Gerätegruppen Unterhaltung & Büro, Kochen sowie Sonstiges entnommen werden. Die Gerätegruppe Beleuchtung wird gemäß den Ergebnissen in [87] modelliert und für die Definition der Nutzungsdauer von WTS wurde eine eigene, nicht repräsentative statistische Erhebung von 70 Spül-, Reinigungs- und Trocknerprogrammen verschiedener Geräte und Hersteller im Rahmen [54] erstellt. Zum Vergleich wurde die Arbeit nach [72] herangezogen. Darüber hinaus wurde, abweichend zu der oben beschriebenen Modellierung für die Gerätegruppe Warmwasseraufbereitung, das *Referenzlastprofile von Ein- und Mehrfamilienhäusern für den Einsatz von KWK-Anlagen* der Richtlinie VDI 4655 herangezogen. Die aufgeführten Nutzungsdauern sind in der Tabelle 2.4 zur Beschreibung des hier vorgestellten BUV zusammengefasst.

Generierung synthetischer Haushaltslastprofile

Wie oben beschrieben basiert das Verfahren im Allgemeinen auf drei Modellierungsschritten. Mit Bezug auf die Nutzungsart drei soll zunächst schematisch das Erzeugen eines einzelnen *Energieblocks* anhand der Gerätegruppe $gg = \text{Unterhaltung} \ \& \ \text{Büro}$ unter Einbindung der Modellierungsschritte vorgestellt werden. Der Abbildung 2.7 sind drei stetige Verteilungsfunktionen zu entnehmen, welche die Wirkleistung nach Gleichung 2.6 (vgl. Abbildung 2.7 Punkt ①), den Nutzungszeitpunkt nach Gleichung 2.9 (vgl. Abbildung 2.7 Punkt ②) sowie die Nutzungsdauer nach Gleichung 2.10 (vgl. Abbildung 2.7 Punkt ③) der Gerätegruppe *Unterhaltung & Büro* für *Zeitbereich = Mittag* modellieren. Durch Ziehen der Realisationen des Nutzungszeitpunktes $t_{\text{Mittag}, TOU} \sim \mathcal{T}_{\text{Mittag}}$ sowie der Nutzungsdauer $t_{\text{Unterhaltung} \ \& \ \text{Büro}, DOU} \sim \mathcal{T}_{\text{Unterhaltung} \ \& \ \text{Büro}, DOU}$ wird unter Berücksichtigung des Erwartungswertes $\mu_{\text{Unterhaltung} \ \& \ \text{Büro}}$ nach Gleichung 2.6 ein *Energieblock* als Repräsentant für eine Tätigkeit der Gerätegruppe Unterhaltung & Büro erzeugt (vgl. Abbildung 2.7, gestrichelte Linie in rot). Der resultierende elektr. Leistungsbezug entspricht der Realisationen $P_{\text{Unterhaltung} \ \& \ \text{Büro}} \sim \mathcal{P}_{\text{Unterhaltung} \ \& \ \text{Büro}}$, exemplarisch in Abbildung 2.7 in grün darge-

stellt. Die abschließende Generierung eines resultierenden synthetischen Lastprofils elektr. Haushalts-Endkunden erfolgt auf Basis einer Überlagerung der stetigen Verteilungsfunktion. Hierbei werden die zuvor definierten unabhängig normalverteilten ZV $\mathcal{P}_{gg,t}$ aller Gerätegruppen gg unter Berücksichtigung des Zeitpunktes t aufsummiert.

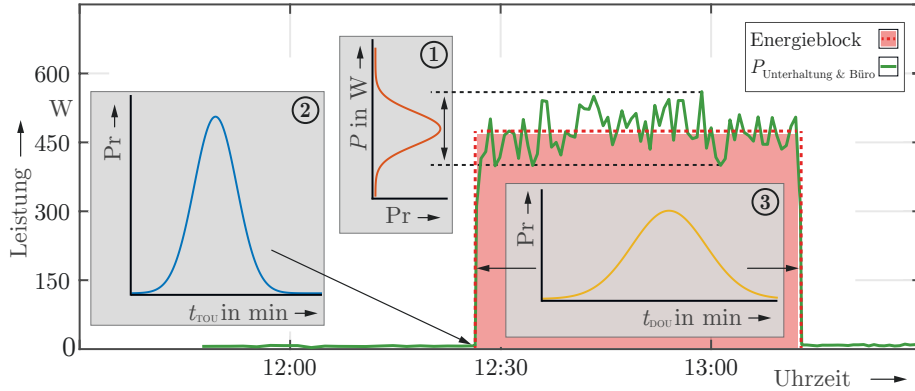


Abbildung 2.7.: Darstellung eines *Energieblocks* der Gerätegruppe *Unterhaltung & Büro* (Nutzungsart drei), hervorgehoben in Rot. Die Modellierung nach dem BUV erfolgt mit Hilfe von drei WDF je Gerätegruppe, welche die Wirkleistung ①, den Nutzungszeitpunkt (aus dem engl. time of use; TOU) ② sowie die Nutzungsdauer (aus dem engl. duration of use; DOU) ③ modellieren. Die grüne Kennlinie entspricht dem hieraus resultierenden elektr. Leistungsbezug $P_{\text{Unterhaltung \& Büro}} \sim \mathcal{P}_{\text{Unterhaltung \& Büro}}$.

Mit den in Tabelle 2.4 definierten Parametern ist das hier vorgestellte BUV ganzheitlich modelliert und soll nachfolgend genutzt werden, um exemplarisch ein synthetisches Haushaltslastprofil für einen *Werktag* im *Winter* zu modellieren. Hierfür bedarf es zunächst der Parameterfestlegung der drei unabhängigen Normalverteilungen nach Gleichung 2.9 zur Definition der *Zeitzbereiche* $\in \{\text{Morgen, Mittag, Abend}\}$ entlang der Zeitachse nach Abbildung 2.6. Zur Modellierung der *Zeitschablone* eines Tages d des Jahres wird das SLP des Profiltyps H0 des *Wochentags = Werktag* und *Jahreszone = Winter* nach Abschnitt 2.1.1 herangezogen. Das SLP ist bis heute von Anbietern, Versorgungsunternehmen und Netzbetreibern anerkannt und spiegelt spezifische Charakteristika bestimmter Kundengruppen im deutschlandweiten statistischen Mittel hinreichend genau wider [55]. Die resultierenden Parameter der drei WDF \mathcal{T}_{TOU} nach Abbildung 2.6 sind wie folgt festgelegt

$$\begin{aligned} \text{Morgen:} & \quad \mu_{\text{Morgen}} = 470 \text{ min}, & \sigma_{\text{Morgen}} &= 93 \text{ min} \\ \text{Mittag:} & \quad \mu_{\text{Mittag}} = 765 \text{ min}, & \sigma_{\text{Mittag}} &= 117 \text{ min} \\ \text{Abend:} & \quad \mu_{\text{Abend}} = 1167 \text{ min}, & \sigma_{\text{Abend}} &= 129 \text{ min} \end{aligned}$$

und definieren die Nutzungszeitpunkte der Gerätegruppen. Die Parameter $\mu_{\text{Zeitzbereich}}$ für *Zeitzbereich* $\in \{\text{Morgen, Mittag, Abend}\}$ werden zur besseren numerischen Handhabung als Minuten des Tages $\mathcal{T} \in \{1, 2, \dots, 1440\}$ min angegeben und $\sigma_{\text{Zeitzbereich}}$ als Minutenintervall.

Nach Definition der *Zeitschablone* mit Hilfe der drei WDF \mathcal{T}_{TOU} können entsprechend der Tabelle 2.4 die Realisationen der Nutzungszeitpunkte $t_{gg, \text{Zeitzbereich}}$ je Gerätegruppe gg im jeweiligen *Zeitzbereich* ermittelt werden, welche der jeweiligen ZV $\mathcal{T}_{\text{Zeitzbereich}}$ folgen.

Analog zum ersten Schritt, der Definition der TOU, wird im zweiten Schritt die DOU der Gerätegruppen der Nutzungsart drei *Regelmäßig betrieben: benutzerabhängiger Betrieb* bestimmt. Hierfür wird auf Grundlage der ZV $\mathcal{T}_{gg,DOU} \sim \mathcal{N}(\mu_{gg,DOU}, \sigma_{gg,DOU}^2)$ eine Realisation $t_{gg,DOU}$ einer Gerätegruppe gg in Minuten bestimmt. Dies erfolgt je Gerätegruppe und je TOU, in welchem diese aktiv sind.

Nach Abbildung 2.7 beschreibt der dritte und letzte notwendige Schritt, neben der Positionierung (TOU) und Dauer (DOU) des *Energieblocks* entlang der Zeitachse, die aufgenommene Leistung der Gerätegruppe. Wie oben beschrieben wird diese im Allgemeinen durch die ZV $\mathcal{P}_{gg,t} \sim \mathcal{N}(\mu_{gg,t}, \sigma_{gg,t}^2)$ repräsentiert, mit

$$\mu_{gg,t} = \frac{W_{gg,Tag}}{\mu_{gg,DOU, ges}} \quad (2.11)$$

für den Erwartungswert der stetigen Normalverteilung $\mathcal{P}_{gg,t}$. Der Erwartungswert basiert auf dem mittleren Energiebedarf $W_{gg,Tag}$ eines Tages und der kumulierten DOU $\mu_{gg,DOU, ges} = \sum_{\text{Zeitbereich}} E(\mathcal{T}_{gg,DOU})$ im Tagesverlauf. Die Summe der Erwartungswerte $E(\mathcal{T}_{gg,DOU})$ wird dabei in Abhängigkeit der Anzahl an *Zeitbereichen*, in welchen die Gruppe aktiv ist, bestimmt. Als erklärendes Beispiel soll die Gerätegruppe Beleuchtung herangezogen werden. Für die Gerätegruppe würde demnach eine Gesamtnutzungsdauer $\mu_{\text{Beleuchtung, DOU, ges}} = 300$ min bestimmt werden, zumal die Gerätegruppe Beleuchtung in zwei Zeitbereichen mit einem Erwartungswert von je $\mu_{\text{Beleuchtung, DOU, Zeitbereich}} = 150$ min aktiv ist.

Die Varianz $\sigma_{gg,t}^2$ wird mit Hilfe der in Gleichung 2.8 bestimmten relativen Standardabweichung $\sigma_{\text{rel},gg,t}$ und dem Erwartungswert $\mu_{gg,t}$ nach Gleichung 2.11 berechnet, durch

$$\sigma_{gg,t} = \sigma_{\text{rel},gg,t} \cdot \mu_{gg,t} . \quad (2.12)$$

Zur Ableitung eines resultierenden Lastprofils entlang der Zeitachse erfolgt im Zuge des hier vorgestellten Verfahrens zunächst eine Summation der ZV $\mathcal{P}_{gg} \sim \mathcal{N}(\mu_{gg}, \sigma_{gg}^2)$ über alle zum Zeitpunkt t aktiven Gerätegruppen gg für alle Minuten des Tages $\mathcal{T} \in \{1, 2, \dots, 1440\}$ min mit

$$\mathcal{P}_{t,ges} = \sum_{gg} \mathcal{P}_{gg,t} . \quad (2.13)$$

Dabei gilt für die unabhängig normalverteilten ZV $\mathcal{P}_{gg,t}$ der Gerätegruppen gg , dass auch deren Summe normalverteilt bleibt [78]. Demzufolge gilt für die resultierende ZV

$$\mathcal{P}_{t,ges} \sim \mathcal{N} \left(\sum_{gg} \mu_{gg,t}, \sum_{gg} \sigma_{gg,t}^2 \right) . \quad (2.14)$$

Auf Basis der resultierenden zeitvarianten WDF nach Gleichung 2.14 wird abschließend ein synthetisches Lastprofil $P_t \sim \mathcal{P}_{t,ges}$ elektr. Haushalts-Endkunden über alle Zeitpunkte t des Tages erzeugt.

Zur Berechnung der Blindleistung werden die in Tabelle 2.4 definierten Intervalle der Leistungsfaktoren $\cos(\varphi)$ mit Hilfe einer stetigen Normalverteilung $\mathcal{N}(\mu_{\cos(\varphi),gg}, \sigma_{\cos(\varphi),gg}^2)$ modelliert. Der Erwartungswert $\mu_{\cos(\varphi),gg}$ wird als Mitte des definierten Intervalls der jeweiligen Gerätegruppe nach Tabelle 2.4 definiert. Die Standardabweichung $\sigma_{\cos(\varphi),gg}$ ist per Definition so gewählt, dass ca. 99.73% der Realisationen im jeweiligen Intervall liegen. Nicht valide Realisationen werden auf $\cos(\varphi)_{gg,t} = 0.85$ gesetzt. Der resultierende Leistungsfaktor $\cos(\varphi)_{res,t}$ wird gewichtet über alle zum Zeitpunkt t aktiven Gerätegruppen, analog zur Gleichung 2.14, bestimmt.

Die einzelnen beschriebenen Schritte zur synthetischen Generierung eines Lastprofils nach dem hier vorgestellten BUV können dem Algorithmus des Anhangs A.5 entnommen werden.

Verifizierung und Validierung eines synthetisch erzeugten Lastprofils

Nachfolgend soll das synthetisch generierte Haushaltslastprofil anhand von spezifischen Merkmalen der Modellierung bewertet und verifiziert werden. Anschließend wird das vorgestellte BUV mit realen Messdaten verglichen und die Gültigkeit validiert. Im Zuge der Verifizierung wird zur besseren Darstellung und Vergleichbarkeit der Ergebnisse auf die Option des 3-phasigen Profils verzichtet und eine 1-phasige Generierung der synthetischen Lastprofile durchgeführt. Für die Validierung mit Hilfe realer Messdaten werden abschließend 3-phasige Haushaltsprofile herangezogen.

Die Grundlage des BUV sind statistische Referenzdaten, die dahingehend aufgearbeitet wurden, dass das resultierende Lastprofil:

1. einer einzelnen Realisierung eine hohe Diversität zu weiteren Realisierungen aufweist.
2. kumuliert gegen ein vorgegebenes globales Profil tendiert, wie z. B. dem SLP.
3. im Mittel den vorgegebenen Erwartungswert der Energie trifft (*energietreuer Ansatz*).

Zur Verifizierung der *ersten* und *zweiten* Modellannahmen soll Abbildung 2.8 herangezogen werden. Visualisiert sind zwei Diagramme mit jeweils drei Zeitreihen, die mit Hilfe des oben beschriebenen Beispiels erstellt wurden. Das obere Diagramm der Abbildung zeigt drei Lastprofile $P_{H,1}$, $P_{H,2}$ und $P_{H,3}$, welche stellvertretend für drei identisch modellierte Haushalte stehen. Die Diversität der einzelnen Zeitreihen zueinander ist deutlich zu erkennen.

Dem gegenüber steht die *zweite* Annahme, welche mit dem unteren Diagramm verdeutlicht werden soll. Dargestellt sind jeweils 110 an einem Netzknoten kumulierte Haushaltslastprofile. Wird die Form der Zeitreihen mit der des SLP des Profiltyps H0 an einem *Werktag* im *Winter* nach Anhang A.2 verglichen, so ist eine klare Tendenz zum vorgegebenen globalen Profil zu entnehmen. Zudem nimmt die Diversität der Profile zueinander mit steigender Akkumulierung ab. Demzufolge konvergieren die synthetisch generierten Lastprofile hin zum vorgegebenen Profil der *Zeitschablone*.

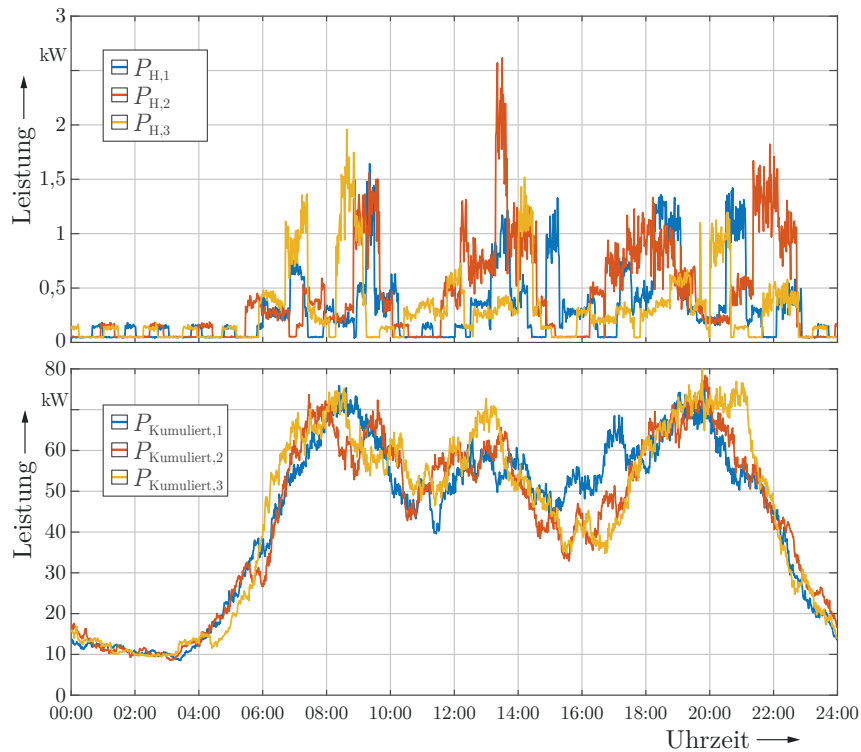


Abbildung 2.8.: Drei Realisationen einzelner Lastprofile auf Basis des BUV (oben) sowie der über 110 Haushaltslastprofile kumulierte Leistungsbedarf (unten).

Zur Verifizierung der *Energietreue* nach Annahme *drei*, werden 1000 Realisationen einzelner Haushalte mit einem JEB von 1000 kWh erzeugt und deren täglicher Energiebedarf berechnet. Die Ergebnisse sind in Form eines Histogramms in Abbildung 2.9 für teilelektrifizierte Haushalte (oben) und vollelektrifizierte Haushalte (unten) dargestellt. Zudem sind die Erwartungswerte des täglichen Energiebedarfs eines durchschnittlichen *Werktages* im *Winter* in rot hervorgehoben. Demnach beträgt der tägliche Energiebedarf an einem Wintertag im Durchschnitt ca. 2,77 kWh / Tag (teilelektrisch) und ca. 3,25 kWh / Tag (vollelektrisch) unter Verwendung des Dynamisierungsfaktors nach Gleichung 2.2 mit $x = 1.187$. In grün hervorgehoben ist jeweils der mittlere Energiebedarf von 1000 Realisationen einzelner Haushalte. Der Analyse ist eine Abweichung von ca. 1% vom Erwartungswert zu entnehmen.

Wie dargestellt folgt das hier vorgestellte BUV den Modellannahmen. Im nächsten Schritt soll die Gültigkeit des Modells mit Hilfe realer Messdaten festgestellt werden. Hierzu wird mit Hilfe der Simulationsumgebung nach [51] eine Leistungsflussberechnung mit minütiger Auflösung durchgeführt. Die Grundlage des Szenarios ist ein Netzgebiet, welches im Rahmen des Forschungsprojektes ENERGIE [36] mit Messtechnik ausgestattet wurde. Das Netzgebiet liegt am Niederrhein und ist von einer ländlichen Struktur geprägt. Die Netzstruktur im Allgemeinen kann der Abbildung A.5 des Anhangs A.6 entnommen werden. Hinzuzufügen ist, dass das Netz über PVA an den Abgängen 1, 3 und 5 verfügt, dessen Einfluss durch die Wahl eines stark bewölkten Tages im Winter minimiert wurde. Um den reduzierten Einfluss der PVA im Zuge der Validierung dennoch zu berücksichtigen, wurden diese mit Hilfe des Modells des folgenden Abschnitts (Abschnitt 2.2) projiziert und

berechnet. Ein Auszug der Datenbank mit den korrespondierenden JEB der Endkunden je HAK sowie der installierten Leistungen der PVA ist der Tabelle A.1 des Anhangs A.6 zu entnehmen.

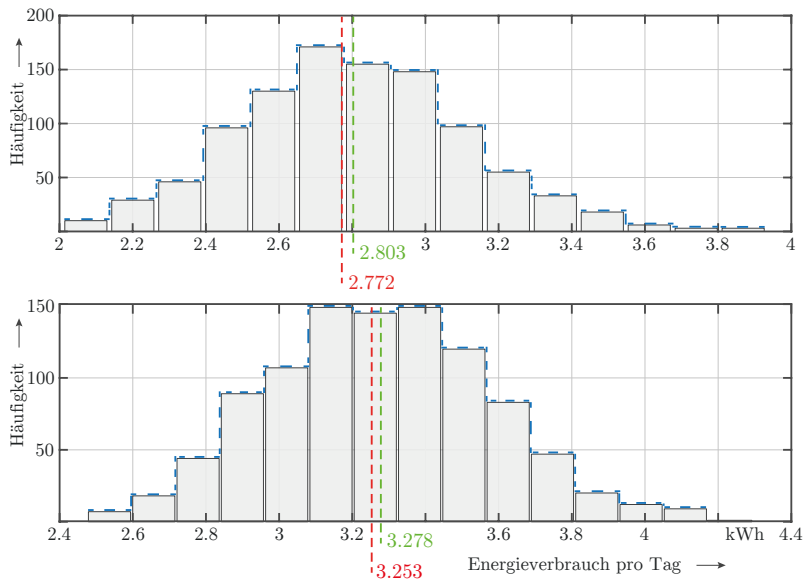


Abbildung 2.9.: Evaluierung des täglichen Energiebedarfs von 1000 simulierten Haushalten nach dem BUV. Das obere Histogramm zeigt die statistische Auswertung teilelektrifizierter Haushalte und das untere Histogramm vollelektrifizierter Haushalte. Jeweils hervorgehoben ist der Erwartungswert des täglichen Energiebedarfs eines durchschnittlichen *Werktages* im *Winter* in rot sowie der resultierende Mittelwert in grün. Beiden Histogrammen ist eine Abweichung von ca. 1% zu entnehmen.

Abbildung 2.11 ist das Ergebnis der Lastapproximation nach dem BUV in der Gegenüberstellung mit realen Messwerten der Stromstärke zwecks Validierung zu entnehmen. Auf der linken Seite der Abbildung sind die hochaufgelösten Messungen der drei Außenleiter $l \in \{L1, L2, L3\}$ der Abgänge 1, 3, 4 & 5 der Netztopologie nach Anhang A.6 am 4. Dezember visualisiert. Dem gegenüber stehen auf der rechten Seite die simulierten Stromstärken der drei Außenleiter. Den Verläufen sind ähnliche Größenordnungen der Stromstärke zu entnehmen sowie eine unsymmetrische Auslastung der Außenleiter. Darüber hinaus sind zwei wesentliche Unterschiede in der direkten Gegenüberstellung ersichtlich. Die reale Messung der linken Seiten verfügt über die folgenden spezifischen Merkmale:

1. Eine zusätzliche Taktung des Stromes der Außenleiter L2 des Abgangs 1 sowie der Außenleiter L1, L2 des Abgangs 5.
2. Eine signifikant höhere Grundlast aller Außenleiter des Abgangs 3 und 5. Insbesondere in den Morgen- und Abendstunden.

Beide Abweichungen könnten durch ein Heizsystem begründet sein. Aufgrund fehlender Informationen wurde ein elektr. Heizsystem, z. B. in Form einer Wärmepumpe [53], im Rahmen der Simulation nicht berücksichtigt.

Neben der Gegenüberstellung der Stromstärken soll zwecks eines weiterführenden Vergleichs der Abgang 4 herangezogen werden, da dieser über keine PVA verfügt. Ziel des

Modellierungsansatzes ist es, ähnliche Prozesse respektive Größenordnungen der elektr. Größen im Rahmen der Generierung des synthetischen Lastprofils zu erzeugen. Hierfür wurden für einen weiteren Vergleich die gemessenen und simulierten Stromstärken in einem Histogramm gegenübergestellt und in Abbildung 2.10 visualisiert. Der Abbildung ist eine gute Übereinstimmung zu entnehmen.

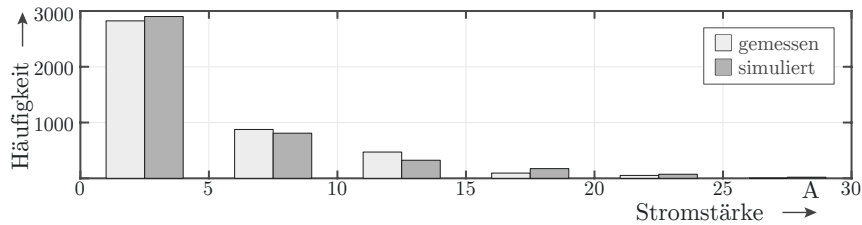


Abbildung 2.10.: Häufigkeitsverteilung der aufgetretenen Stromstärke des Abgangs 4 für gemessene und simulierte Werte [82].

Resümee

Im Zuge der Verifizierung und Validierung konnte gezeigt werden, dass das erarbeitete BUV im Vergleich zu realen Messwerten ähnliche Ereignisse im elektr. Netz produziert. Per Definition erfolgt die Modellierung auf Ebene der Gerätegruppen mit dem Ziel, das notwendige Informationslevel zur Modellierung möglichst gering zu halten. In Abgrenzung zu Arbeiten der einschlägigen Literatur stehen insbesondere der JEB sowie das SLP im Fokus des erarbeiteten Modellierungsansatzes. Darüber hinaus bietet der erarbeitete Ansatz der *Zeitschablone* in Kombination mit einem *energietreuen* Ansatz die Möglichkeit regionale Gegebenheiten schnell im Rahmen der Modellierung zu berücksichtigen, um die Dynamik im Niederspannungsnetz weitergehend zu untersuchen. Dies konnte in [89] bereits erfolgreich umgesetzt werden. Mit Hilfe detaillierter statistischer Daten konnten neue Nutzungszeitpunkte sowie Nutzungsdauern definiert und zur synthetischen Lastapproximation genutzt werden.

Zudem bleibt festzuhalten, dass im Vergleich zum vorgestellten TDV mit einer höherer Informationsdichte für die Lastapproximation mittels BUV zu rechnen ist [73] (vgl. Abbildung 2.4). Auch wenn die Autoren in [55] darauf hinweisen, dass bspw. die Information über die Anzahl der im Haushalt lebenden Personen keinen nennenswerten Einfluss auf das aggregierte Lastprofil im Zuge der Klassifizierung nimmt, kann die Berücksichtigung dieser Informationen einen tieferen Einblick in die Netzstruktur und -dynamik des Niederspannungsnetzes liefern. Im Kontext der integrierten Energiewende sind diese detaillierten Untersuchungen zwingend notwendig und erfordern eine komplexere Modellierung der Vorgänge auf den unteren Spannungsebenen. Folglich findet im Rahmen der vorliegenden Arbeit das BUV innerhalb der Simulationsumgebung \mathcal{E}_{sim} Anwendung (vgl. Abbildung 1.1). Hierdurch werden für den Agenten des RL im Kontext des *Lernmoduls* möglichst realitätsnahe Vorgänge des Niederspannungsnetzes erzeugt, sodass dieser eine adäquate Kontrollstrategie zur Lösung der Problemstellung ableiten kann.

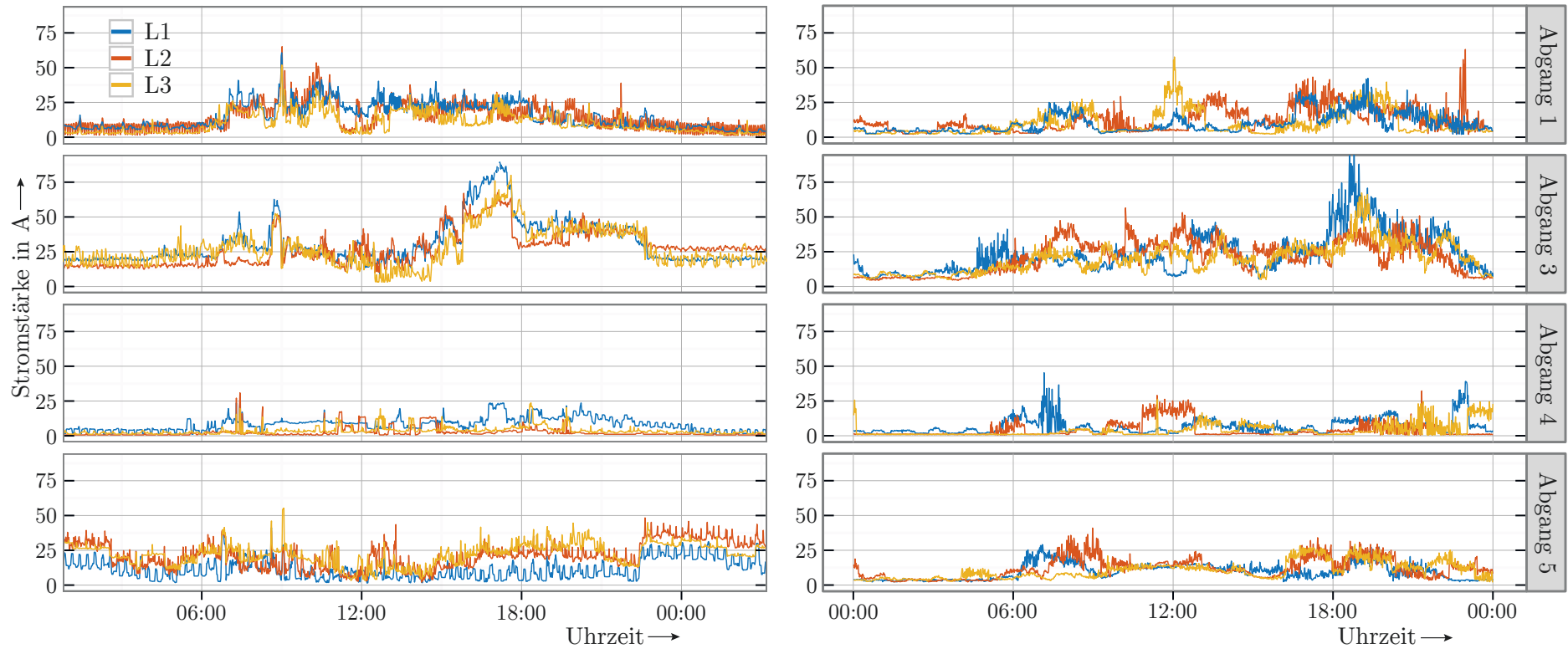


Abbildung 2.11.: Vergleich von hochauflösenden Messungen der Stromstärke (links) sowie den Simulationsergebnissen (rechts) je Abgang des Niederspannungsnetzes in minütiger Auflösung. Da das Niederspannungsnetz über PVA verfügt, wurde für einen besseren Vergleich ein stark bewölkter Werktag im Winter gewählt, da andernfalls die Messungen durch die hohe Durchdringung der PVA verzerrt wird. Lediglich Abgang 4 verfügt über keine PVA. Im Allgemeinen sind den Verläufen ähnliche Größenordnungen der Stromstärke zu entnehmen.

2.2. Einspeisemodell einer Photovoltaikanlage

Im Zuge der integrierten Energiewende stehen VNB vor der Herausforderung neue Akteure, wie u.a. PVA, Wärmepumpen, (Mikro)-Blockheizkraftwerke (BHKW), (Klein-) WKA oder elektr. betriebene Personenkraftwagen, effizient in Betriebs- und Planungsprozesse einzubinden. Mit Hilfe von Modellen für Last- und Einspeiseprozesse können Netzführungsstrategien vorab erarbeitet und bewertet werden. Basierend auf einer Netzberechnung können Untersuchungsszenarien definiert werden, welche die Einbindung neuer Akteure mit unterschiedlicher Durchdringung und unter Berücksichtigung vorhandener Unsicherheiten ermöglichen. Aufbauend hierauf kann eine zukünftige Entwicklung des jeweiligen Netzes im Zuge der Netzplanung bestimmt werden [90].

Im Fokus der vorliegenden Arbeit liegen die Gegebenheiten und Ergebnisse des Forschungsvorhabens ENERGIE [36]. Die Netztopologien der betrachteten Niederspannungsnetze weisen überwiegend PVA auf, weshalb nachfolgend die Modellierung des Energiewandlungsprozesses einer PVA vorgestellt wird. Demnach finden in der vorliegenden Arbeit jene Niederspannungsnetze des Forschungsvorhabens keine Berücksichtigung, welche ein Mikro-BHKW, Wärmepumpen und Klein-WKA aufweisen. Dies ist durch die vergleichsweise geringe Durchdringung der Akteure respektive der geringen im Niederspannungsnetz wirksamen elektr. Leistung im Vergleich zum Modellierungsaufwand begründet.

2.2.1. Wirkungskette des Energiewandlungsprozesses

Das Modell einer PVA berücksichtigt explizit die zeitliche Entwicklung der Einspeiseleistung im Tagesverlauf. Hierfür sind neben der globalen Bestrahlungsstärke und den variierenden atmosphärischen Bedingungen über dem Netzgebiet insbesondere die Anlagenparameter ausschlaggebend. Im Allgemeinen versetzt die im EEG verankerte Meldepflicht für Anlagenbetreiber und Eigenversorger die VNB in die Lage alle wesentlichen Informationen im Rahmen eines Erhebungsprozesses bei der Anmeldung einer PVA mittels Steckbrief zu erfragen. Hierbei zeigen die Ergebnisse des Forschungsvorhabens ENERGIE [36] sowie Scheffler in [58], dass insbesondere die Ausrichtung und Neigung der Modulfläche einen signifikanten Einfluss auf den Energiewandlungsprozess der PVA haben.

Eine Übersicht des erarbeiteten Einspeisemodells einer PVA zeigt Abbildung 2.12. Der schematischen Darstellung ist eine temperaturabhängige Wirkungskette zu entnehmen, welche den Energiewandlungsprozess zwischen der globalen Bestrahlungsstärke auf geneigter Fläche $E_g(\vartheta_M, t)$ und der elektr. Einspeiseleistung $P(t)$ visualisiert. Die Wirkungskette wird durch

$$P(t) = E_g(\vartheta_M, t) \cdot A_{\text{PVA}} \cdot \eta_M \cdot \eta_{\text{WR}} \cdot f_{\text{BG}}(t) \quad (2.15)$$

beschrieben und approximiert die resultierende Einspeiseleistung $P(t)$ der PVA in das elektr. Netz, mit der Fläche der verbauten Module A_{PVA} sowie dem Wirkungsgrad des Wechselrichters und der PV-Module, η_{WR} und η_M . Die Scheinleistung wird mit Hilfe eines

korrespondierenden Leistungsfaktors $\cos(\varphi)$ bestimmt. Zudem ist der Gleichung ein zusätzlicher Faktor $f_{\text{BG}}(t)$ zu entnehmen. Dieser soll die Berücksichtigung des Bedeckungsgrades sowie Wolkenreflexionen über dem Netzgebiet ermöglichen und demnach einen direkten Einfluss auf die Einspeiseleistung der PVA modellieren.

Die Wirkungskette nach Gleichung 2.15 basiert im Wesentlichen auf [61]. Einzelne Komponenten der Wirkungskette sind nachfolgend im Detail beschrieben. Eine umfangreiche Beschreibung meteorologischer Modelle im Kontext der Einspeisemodellierung einer PVA sind bspw. [61] sowie [89] zu entnehmen.

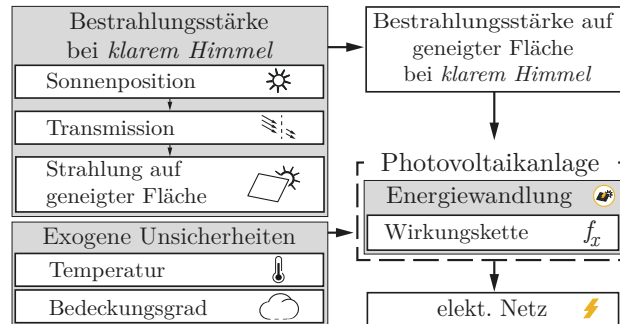


Abbildung 2.12.: Wirkungskette des Energiewandlungsprozesses einer PVA. Exogene Unsicherheiten sowie Installationsparameter werden in der Modellierung von Strahlungsleistung in elektr. Einspeiseleistung berücksichtigt.

2.2.2. Modellierung der Bestrahlungsstärke auf geneigter Fläche

Die Modellierung der globalen Bestrahlungsstärke auf geneigter Fläche erfolgt im Rahmen der vorliegenden Arbeit nach der Beschreibung in [61]. Hiernach wird die augenblickliche Bestrahlungsstärke auf einer geneigten Fläche über

$$E_g(t) = E_{\text{dir}}(t) + E_{\text{diff}}(t) + E_{\text{ref}}(t) \quad (2.16)$$

bestimmt. Nach Gleichung 2.16 erfolgt die Approximation der Bestrahlungsstärke auf geneigter Fläche $E_g(t)$ in drei Schritten:

1. Berechnung des direkten Strahlungsanteils $E_{\text{dir}}(t)$ über dem Netzgebiet mit Hilfe der Sonnenposition.
2. Bestimmung der diffusen Bestrahlungsstärke $E_{\text{diff}}(t)$ bedingt durch Streueffekte bei der Transmission durch die Atmosphäre bei *klarem Himmel* (engl. clear sky [61]).
3. Berücksichtigung der vom Boden reflektierten Bestrahlungsstärke $E_{\text{ref}}(t)$ mit Hilfe eines definierten Albedo-Wertes.

Die Bestimmung der Sonnenposition im Kontext der Berechnung der direkten Strahlung basiert im Wesentlichen auf Ansätzen der Norm *Tageslicht in Innenräumen (DIN 5034 Teil 2)*. Abweichend zu [61] erfolgt die Bestimmung der diffusen Bestrahlungsstärke $E_{\text{diff}}(t)$ nach [91]. Für die Bestimmung der reflektierten Bestrahlungsstärke können unterschiedliche Albedo-Werte [61] entnommen werden.

Der Abbildung 2.13 ist die Grundüberlegung zur Bestimmung des Sonneneinfallswinkels auf einer geneigten PVA sowie der hierfür notwendigen Winkelbezeichnungen zu entnehmen. Mit Hilfe der DIN 5034 Teil 2 können Sonnenhöhe γ_S sowie Sonnenazimut α_S am Aufstellungsort der PVA zum Zeitpunkt t bestimmt werden. Zur Berücksichtigung der geneigten PVA muss die Drehung der Modulfläche um den Winkel α_F mit $\Delta\alpha = |\alpha_S - \alpha_F|$ sowie die Neigung γ_F bekannt sein. Anschließend kann durch eine geometrische Verrechnung nach [61] die resultierende globale Bestrahlungsstärke auf geneigter Fläche bestimmt werden.

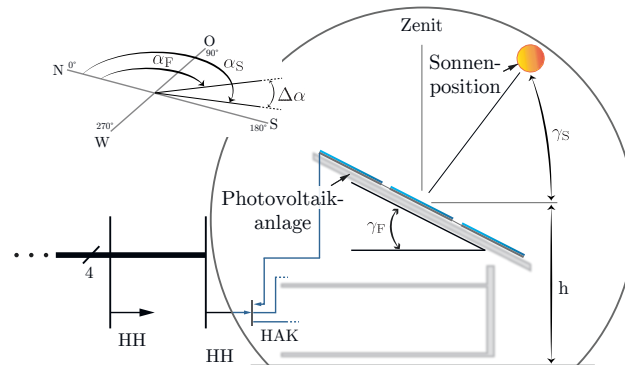


Abbildung 2.13.: Winkelbezeichnungen zur Bestimmung des Sonneneinfallswinkels einer geneigten PVA.

Im Rahmen der Netzsimulation wird während der Simulationslaufzeit die deterministische Sonnenposition über dem Netzgebiet bestimmt. Anschließend kann die globale Bestrahlungsstärke, welche die auf der Erde ankommende Bestrahlungsstärke auf einer horizontalen Fläche beschreibt [61], unter Annahme eines *klaren Himmels* berechnet werden. Auf Grundlage der Bestrahlungsstärke auf einer horizontalen Fläche wird anschließend während der Simulationslaufzeit die Bestrahlungsstärke auf der geneigten Fläche $E_g(t)$ bestimmt. Dabei wird jeweils die projizierte Ausrichtung der PVA berücksichtigt. Die einzelnen algorithmischen Berechnungsschritte und analytischen Zusammenhänge des deterministischen Modells sowie Konstanten und Kennzahlen sind im Detail [61] sowie [91] zu entnehmen.

In Abbildung 2.14 ist im linken Diagramm in gelb die berechnete globale Bestrahlungsstärke unter Annahme eines *klaren Himmels* exemplarisch dargestellt. Zudem ist dem Diagramm jeweils eine Messung bei *bedecktem Himmel* in grün und *überwiegend klarem Himmel* in blau zu entnehmen. Im direkten Vergleich der simulierten sowie der gemessenen Zeitreihe bei überwiegend klarem Himmel zeigt die erste Tageshälfte eine sehr gute Übereinstimmung. Die zweite Tageshälfte weist eine zunehmende Bedeckung des Himmels auf, was in einer Abweichung der Zeitreihen zueinander resultiert. Der Messung bei stark bedecktem Himmel in grün sind insbesondere zur Mitte des Tages überhöhte Werte der Bestrahlungsstärke im Vergleich zu der simulierten Kennlinie zu entnehmen. Diese sind durch Wolkenreflexionen begründet (vgl. [61]).

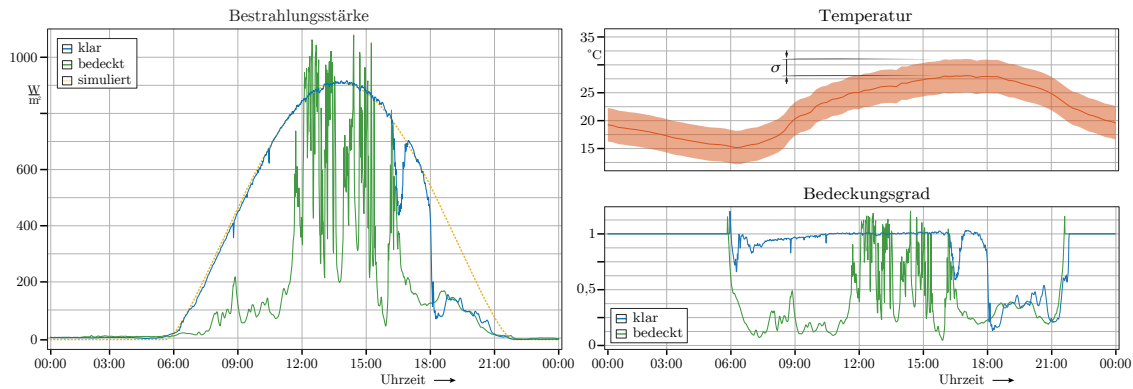


Abbildung 2.14.: Links: Vergleich der simulierten (gelb) und gemessenen (grün und blau) Bestrahlungsstärke. Die blaue Kennlinie zeigt einen gemessenen Tag im Sommer bei überwiegend klarem Himmel und die grüne Kennlinie bei bedecktem Himmel. Insbesondere dem Vergleich der ersten Tageshälfte der simulierten und gemessenen Zeitreihe bei überwiegend klarem Himmel in blau ist eine deutliche Übereinstimmung zu entnehmen. Rechts: Temperaturmessung mit konstanter Standardabweichung σ (oben) sowie ein exemplarischer Verlauf des Bedeckungsgrades $f_{BG}(t)$ (unten) für einen überwiegend klaren Himmel in blau und einem stark bedeckten Himmel in grün.

2.2.3. Exogene Unsicherheiten

Umgebungs- und Modultemperatur

Neben der Umgebungstemperatur $\vartheta_U(t)$ zeigt auch die Modultemperatur $\vartheta_M(t)$ einen signifikanten Einfluss auf die resultierende Einspeiseleistung $P(t)$ der PVA. Begründet durch den absorbierten Strahlungsanteil, welcher nicht in elektr. Energie umgewandelt wird, weicht die Modultemperatur der PVA von der Umgebungstemperatur ab [61]. Mit

$$\vartheta_M(t) = \vartheta_U(t) + c \cdot \frac{E_g(t)}{1000 \frac{W}{m^2}} \quad (2.17)$$

kann nach [61] die Modultemperatur $\vartheta_M(t)$ überschlagsweise aus der Umgebungstemperatur $\vartheta_U(t)$ und der globalen Bestrahlungsstärke auf geneigter Fläche $E_g(t)$ approximiert werden. Der Faktor c nach Gleichung 2.17 beschreibt eine Proportionalitätskonstante, welche von der Montageart der Module abhängt und mit der Wärmeabführung korreliert. Dem Leitfaden nach [92] sind Proportionalitätskonstanten für verschiedene Einbauvarianten, wie bspw. Dachintegration oder Fassadenintegration mit guter, schlechter oder ohne Hinterlüftung, zu entnehmen.

Mit Hilfe der resultierenden Modultemperatur $\vartheta_M(t)$ nach Gleichung 2.17 wird die nutzbare globale Bestrahlungsstärke auf geneigter Fläche $E_g(t)$, ausgehend vom Arbeitspunkt bei STC mit einer Modultemperatur $\vartheta_{M, STC} = 25^{\circ}C$ sowie einem Temperaturkoeffizienten α_P , über

$$E_g(\vartheta_M, t) = E_g(t)(1 + \alpha_P(\vartheta_M - 25^{\circ}C)) \quad (2.18)$$

korrigiert. Typische Temperaturkoeffizienten α_P unterschiedlicher Modultypen handelsüblicher PVA sind [61] zu entnehmen.

Der signifikante Einfluss der Umgebungs- und Modultemperatur auf die Einspeiseleistung der PVA findet im Zuge der Netzsimulation Berücksichtigung. Die Vorgabe der Umgebungstemperatur erfolgt mit Hilfe einer (gemessenen) Zeitreihe, exemplarisch im oberen Diagramm der rechten Seite der Abbildung 2.14 visualisiert. Dem Diagramm ist eine Standardabweichung σ zu entnehmen, welche eine Unsicherheit der Messung der Umgebungstemperatur sowie Modellierung der Modultemperatur während der Simulationslaufzeit adressieren soll. Insbesondere die unterschiedlichen Modultypen und deren korrespondierende Temperaturkoeffizienten α_P können durch diese Herangehensweise einbezogen werden.

Bedeckungsgrad

In mitteleuropäischen Breitengraden sind PVA von einer volatilen Einspeisecharakteristik geprägt. Diese auf einer stark wechselnden Bewölkung beruhende Eigenschaft muss in der Netzsimulation Berücksichtigung finden. Darüber hinaus kann die max. Bestrahlungsstärke an einem stark bedeckten Tag im Winter auf 30 W/m^2 sinken (vgl. Abbildung 2.14 mit ca. 1000 W/m^2 im Sommer). Im Gegensatz hierzu kann im Sommer, bedingt durch Wolkenreflexionen bei wechselnder Bewölkung, eine kurzzeitige Erhöhung der globalen Bestrahlungsstärke von ca. 20 % auftreten [61].

Im Rahmen der vorliegenden Arbeit wird auf eine detaillierte Modellierung meteorologischer Prozesse zur Beschreibung des Bedeckungsgrades und der damit einhergehenden Wolkenreflexion, wie bspw. in [93], verzichtet. Im Fokus steht eine realitätsnahe Modellierung des Einspeiseprozesses der PVA, welche die beschriebene volatile Charakteristik im Tagesverlauf innerhalb der Netzsimulation abbildet. Dies erfolgt mit dem in Gleichung 2.15 definierten Faktor $f_{BG}(t) \in \mathbb{R}$, mit $f_{BG}(t) \in (0, f_{BG, \max.}]$. Dieser beschreibt eine Skalierung der globalen Bestrahlungsstärke auf geneigter Fläche $E_g(\vartheta_M, t)$ und ermöglicht mit $f_{BG, \max.} > 1$ Wolkenreflexionen explizit zu berücksichtigen. Demnach ist zwischen den folgenden drei Fällen zu unterscheiden:

1. $f_{BG}(t) < 1$: Reduzierung der globalen Bestrahlungsstärke hervorgerufen durch einen *bedeckten Himmel* oder einer systematischen Verschattung.
2. $f_{BG}(t) = 1$: *Klarer Himmel*.
3. $f_{BG}(t) > 1$: Simulierte Erhöhung der globalen Bestrahlungsstärke bedingt durch Wolkenreflexionen.

Analog zur zuvor beschriebenen Umgebungs- und Modultemperatur erfolgt die Berücksichtigung des Bedeckungsgrades innerhalb der Netzsimulation während der Simulationslaufzeit durch Vorgabe einer Zeitreihe unter Verwendung einer statistischen Unsicherheit. Der rechten Seite des unteren Diagramms der Abbildung 2.14 ist exemplarisch eine Zeitreihe des Bedeckungsgrades $f_{BG}(t)$ zu entnehmen.

Das Erzeugen der Zeitreihe kann bspw. mit Hilfe eines Pyranometers erfolgen [61]. Alternativ kann ein gemessenes Wirkleistungsprofil einer PVA als Referenzanlage herangezogen

werden. Der Faktor $f_{BG}(t)$ wird anschließend unter Verwendung der Wirkungskette nach Gleichung 2.15 zum Zeitpunkt t bestimmt. Die resultierende Zeitreihe wird anschließend als Bedeckungsgrad über dem Netzgebiet angenommen und beeinflusst alle im Netzgebiet projektierten PVA in gleichem Maße.

2.2.4. Generierung synthetischer Einspeiseprofile

Zur realitätsnahen Projektierung einer PVA innerhalb der Netzsimulation mit dem zuvor beschriebenen Einspeisemodell sind neben der resultierenden Fläche der verbauten Module A_{PVA} , dem Wirkungsgrad des Wechselrichters η_{WR} und der PV-Module η_M (nach Gleichung 2.15) weitere Parameter zu definieren. Der im Rahmen der vorliegenden Arbeit verwendete Ansatz sieht zudem spezifische Informationen über den Aufstellungsort (als geographische Koordinaten sowie der Höhe über dem Meeresspiegel), Ausrichtung α_F und Neigung γ_F der PVA, den Modultypen (zur Ableitung des Temperaturkoeffizienten α_P nach [61]) sowie die Montageart c vor. Diese erweiterte Datenbasis ist zwingend notwendig, um mit Hilfe des modellierten Einspeiseprozesses der PVA wesentliche Netzkapazitäten im Tagesverlauf zu identifizieren und im Rahmen von Betriebsführungsstrategien im Sinne der integrierten Energiewende zu nutzen.

Verifizierung und Validierung eines synthetisch erzeugten Einspeiseprofils

Zur Verifizierung des Ansatzes soll der Fokus zunächst auf den Einfluss der Ausrichtung und Neigung der PVA gelegt werden. Abbildung 2.15 sind synthetisch generierte Einspeiseprofile mit dem hier vorgestellten Einspeisemodell einer PVA zu entnehmen. Der Abbildung ist die Beeinflussung der Einspeiseleistung an einem Tag d im Sommer

- bei unterschiedlicher Ausrichtung der Modulflächen $\alpha_F := \{90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ\}$ mit $0^\circ \leq \alpha_F < 360^\circ$ sowie einer festen Neigung $\gamma_F = 35^\circ$ im linken Diagramm der Abbildung 2.15 bzw.
- bei unterschiedlicher Neigung $\gamma_F := \{0^\circ, 10^\circ, 20^\circ, 30^\circ\}$ mit $0^\circ \leq \gamma_F \leq 90^\circ$ sowie einer festen Ausrichtung $\alpha_F = 180^\circ$ (rechtes Diagramm)

bei *klarem Himmel* zu entnehmen. Den Diagrammen ist ein signifikanter Einfluss auf das Einspeiseprofil zu entnehmen. Während die Neigung im Wesentlichen die max. Einspeiseleistung in der Höhe beeinflusst, sorgt die Ausrichtung der Modulflächen für eine Verschiebung des Zeitpunktes der max. Einspeiseleistung im Tagesverlauf. Demnach können PVA mit unterschiedlicher Ausrichtung ihre max. Einspeiseleistung nicht zeitgleich an das Netz abgeben.

Dieser Einfluss der detaillierten Datenbasis resultiert bei der Betrachtung eines Niederspannungsnetzes in deutlichen Kapazitätsgewinnen an übertragbarer Wirkleistung, visualisiert in Abbildung 2.16. Dargestellt ist jeweils der Außenleiter L1 des Abgangs 1, welcher mit Hilfe der Netzsimulation des Niederspannungsnetzes nach Anhang A.6 berechnet wurde. Die Kennlinie in Rot zeigt die resultierende Wirkleistung der Endkunden unter Berücksichtigung einer detaillierten Modellierung. Haushalts-Endkunden sind mit Hilfe des TDV nach

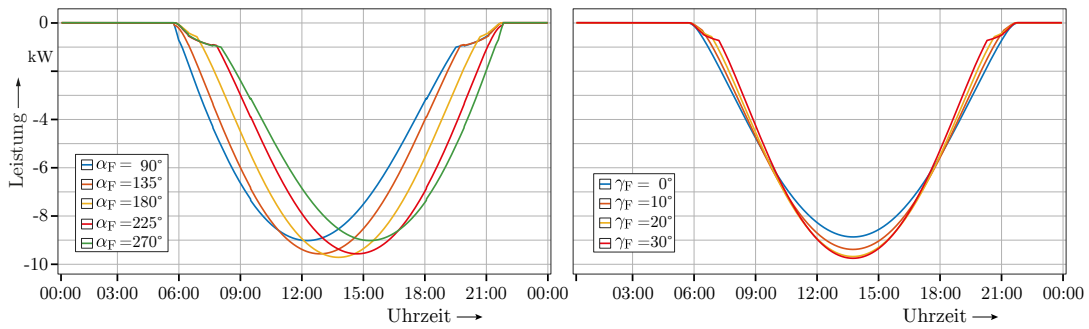


Abbildung 2.15.: Links: Einfluss der Ausrichtung der Modulfläche auf das Einspeiseprofil im Tagesverlauf für $\alpha_F := \{90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ\}$ und einer Neigung von $\gamma_F = 35^\circ$. Rechts: Einfluss der Neigung der Modulfläche auf die max. Einspeiseleistung für $\gamma_F := \{0^\circ, 10^\circ, 20^\circ, 30^\circ\}$ und einer Ausrichtung von $\alpha_F = 180^\circ$.

Abschnitt 2.1.2.1 und PVA mit dem beschriebenen Einspeisemodell modelliert. Als Referenz dient die in blau dargestellte Kennlinie. Zur Erzeugung des Referenzprofils wurden Haushalts-Endkunden mit Hilfe des SLP skaliert sowie Einspeiseprofile für die Modellierung der PVA angenommen, dessen Ausrichtung und Neigung identisch über alle Knotenpunkte mit PVA ist. Simuliert ist ein vollständiger Werktag d im Sommer bei *klarem Himmel* mit $f_{BG} = 1.0$ sowie einer vorgegebenen Temperatur nach Abbildung 2.14 mit $\sigma = 3\text{K}$. Das Simulationsergebnis weist eine Reduzierung der max. Einspeiseleistung von $15,48\text{ kW}$ auf $11,96\text{ kW}$ (ca. 23%) auf. Zudem tritt die max. Rückspeisung 115 Minuten früher ein. Des Weiteren ist der Abbildung 2.16 zu entnehmen, dass das Wirkleistungsprofil in Rot ein ausgeprägtes Plateau ausbildet, welches durch die unterschiedlichen Ausrichtungen der PVA begründet ist.

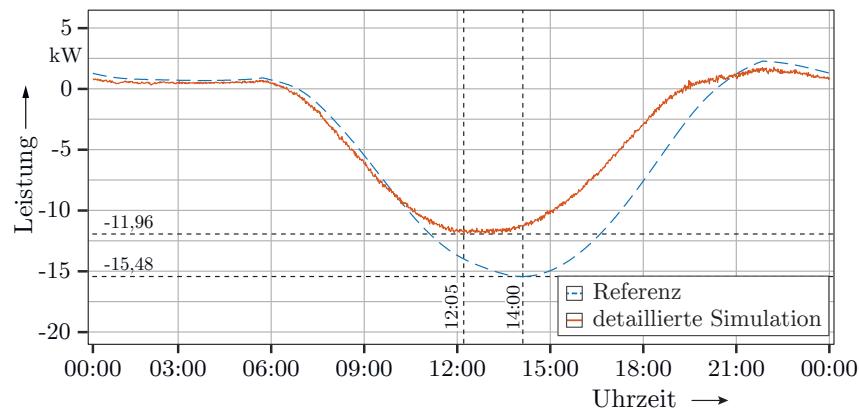


Abbildung 2.16.: Wirkleistung in Abgang 1 eines realen Ortsnetzes nach Abbildung A.5 des Anhangs A.6 auf Außenleiter L1. In Rot dargestellt ist das Simulationsergebnis unter Berücksichtigung des vorgestellten Einspeisemodells. Die Referenzkennlinie in blau weist eine identische Ausrichtung und Neigung aller PVA auf. Dem Diagramm ist ein Kapazitätsgewinn an übertragbarer Wirkleistung im Vergleich zur Referenz sowie eine zeitliche Verschiebung der max. Einspeiseleistung von ca. 23% um 115 Minuten zu entnehmen.

Im nachfolgenden Schritt soll das synthetisch erzeugte Einspeiseprofil mit Hilfe einer im Forschungsvorhaben aufgenommenen Messung verglichen und validiert werden. Die notwendigen Parameter wurden im Rahmen des Forschungsvorhabens bei der Montage der Referenzmessung mit Hilfe eines Steckbriefes durch den Monteur erfasst. Da eine exakte Erfassung aller Parameter durch den Monteur nicht möglich war, wie bspw. die Neigung des

Daches, wurden einige Parameter approximiert. Die definierten Parameter des Steckbriefs sind [36] zu entnehmen. Ein Auszug der aufgenommenen Parameter einer Referenzanlage ist Tabelle 2.5 zu entnehmen. Die gemessene PVA ist am Niederrhein verortet und weist eine südöstliche Ausrichtung auf. Die Module bestehen aus polykristallinem Silizium (Poly-Si) und sind dachintegriert mit guter Hinterlüftung montiert.

Tabelle 2.5.: Parameter einer PVA in südöstlicher Ausrichtung (Modultyp Poly-Si). Die Anlage ist am Niederrhein verortet und dachintegriert mit guter Hinterlüftung montiert.

P_{el}	α_F	γ_F	A_{PVA}	α_P	c	η_M	η_{WR}
kW _p	°	°	m ²	%/K	K	%	%
12,22	222	35	84,94	-0,45	29	14,4	97,9

Die Messung der Referenzanlage (grün) sowie das synthetisch generierte Einspeiseprofil der Anlage ohne Berücksichtigung eines Bedeckungsgrades (gelb) sind der Abbildung 2.17 zu entnehmen. Im direkten Vergleich weisen beide Zeitreihen optisch einen ähnlichen Verlauf auf. Sowohl die Referenzmessung als auch das synthetisch generierte Einspeiseprofil weisen eine max. Einspeiseleistung um ca. 15 Uhr auf. Jedoch ist der Abbildung bei ① eine wesentlich höhere max. Einspeiseleistung der Referenzanlage zu entnehmen. Dies könnte zum einen an der zuvor beschriebenen Wolkenreflexion liegen, zum anderen an den durch den Monteur teilweise approximierten Parametern. So spiegelt ein optimierter Parametersatz, dargestellt in orange, mit einer um 5 m² erhöhten Fläche der verbauten Module die max. Einspeiseleistung sowie die Form der Messung optisch wesentlich besser wider.

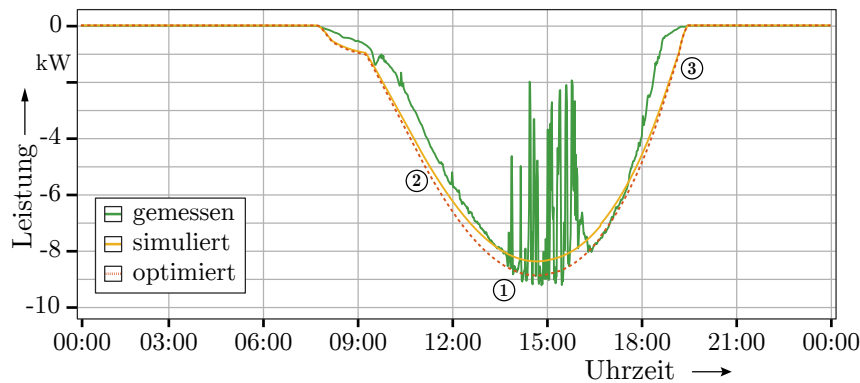


Abbildung 2.17.: Vergleich des Einspeisemodells einer PVA mit der Messung einer PV-Referenzanlage (grün). Die gelbe Kennlinie stellt die simulierte PVA mit den aufgenommenen Parametern dar. In orange ist die optimierte Modellrechnung mit einer um 5 m² vergrößerten Fläche der verbauten Module. Abweichungen zwischen der simulierten und gemessenen Zeitreihe sind jeweils markiert.

Zudem ist der Abbildung ein schmalerer Verlauf der Referenzmessung zu entnehmen, was durch einen bedeckten Himmel begründet sein kann. Dies scheint für die linke Flanke, markiert mit ②, der Fall zu sein. So ist dem Diagramm zwischen dem optimierten Einspeiseprofil in orange sowie der Referenzmessung in grün ein äquidistanter Abstand zu entnehmen. Während die rechte Flanke, in Abbildung 2.17 mit ③ markiert, mit fortschreitender Zeit eine zunehmende Abweichung zum synthetisch generierten Einspeiseprofil aufweist. Eine mögliche Erklärung ist der schematischen Darstellung der Referenzanlage nach Abbildung

2.18 (links) zu entnehmen. Zu erkennen ist ein Schattenwurf des Nachbargebäudes im Südwesten, welches zum Sonnenuntergang in einer systematischen Verschattung einzelner PV-Module der Referenzanlage resultiert.

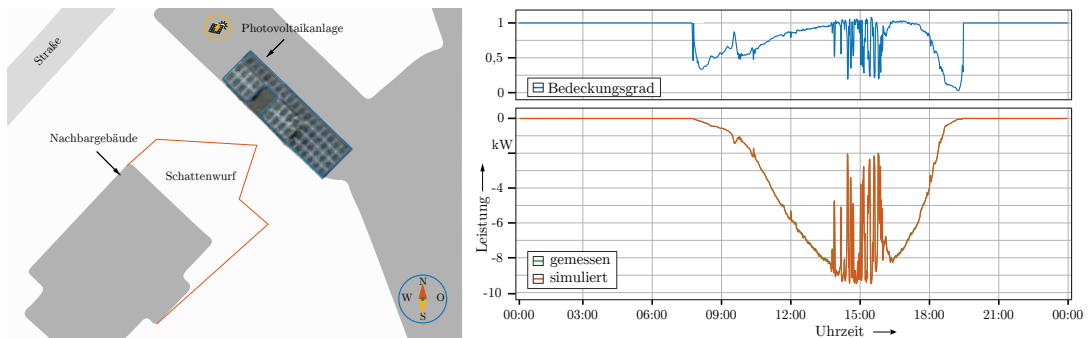


Abbildung 2.18.: Links: Schematische Darstellung der Referenzanlage. Der Schattenwurf des Nachbargebäudes ist gekennzeichnet (aus Datenschutzgründen nachgezeichnet). Rechts: Diagramm des resultierenden Bedeckungsgrades zur Korrektur der Simulationsrechnung des Einspeisemodells (oben) sowie die simulierte und gemessene Referenzanlage unter Berücksichtigung des Bedeckungsgrades (unten).

Die Berücksichtigung der beschriebenen Abweichungen erfolgt im Rahmen der Netzsimulation mit Hilfe des Faktors f_{BG} . Unter Anwendung der Wirkungskette nach Gleichung 2.15 kann ein resultierender Faktor $f_{BG}(t)$ bestimmt und das Einspeiseprofil entsprechend korrigiert werden. Dies ist im Diagramm der Abbildung 2.18 dargestellt. Aufgrund der beschriebenen systematischen Verschattung der Referenzmessung ist der Einsatz des abgeleiteten Faktors $f_{BG}(t)$ als allgemeiner Bedeckungsgrad über dem Netzgebiet jedoch zu hinterfragen.

Resümee

Das vorgestellte Einspeisemodell einer PVA basiert auf einer Wirkungskette und berücksichtigt detaillierte Anlageninformationen. Gegenüber klassischen Ansätzen mit globalen Referenzprofilen ermöglicht das vorgestellte Modell wesentliche Netzkapazitäten im Tagesverlauf zu identifizieren. Hierbei zeigt insbesondere das Einbeziehen der Neigung und Ausrichtung der PVA die Möglichkeit einer Optimierung der zeitlichen Zusammenhänge im Kontext von Netzführungsstrategien und Entscheidungsprozessen auf Niederspannungsebene. Dabei ist die Berücksichtigung des Bedeckungsgrades obligatorisch. Erfolgt die Ableitung des Bedeckungsgrades mit Hilfe einer PVA als Referenzanlage, ist neben der Erfassung von exakten Parametern auf mögliche Verschattung der PV-Module zu achten.

Die Erfassung der notwendigen Parameter für eine realitätsnahe Abbildung der PVA im Rahmen der Netzsimulation, insbesondere für bereits installierte Anlagen, stellt einen erheblichen Aufwand dar. Zudem sind einige Parameter schwer zu erfassen und können nur approximiert werden. Im Rahmen des Forschungsvorhabens ENERGIE erfolgt die Erfassung der Parameter für 20 Referenzanlagen durch den Menschen. Für weitere ca. 120 PVA wurden Luftbildaufnahmen und Informationen öffentlicher Katasterdaten zur Ableitung notwendiger Parameter herangezogen [36]. Beide Ansätze resultieren in einer größeren Unsicherheit, welche im Rahmen der Modellierung einbezogen werden sollte.

2.3. Modellierung der Spannungsabhängigkeit

Aufbauend auf Untersuchungen zur spannungsabhängigen Verbraucherleistung im elektr. Energieversorgungssystem in den 1980er bis Mitte der 1990er Jahre konnten unterschiedliche statische und dynamische Modelle erarbeitet und etabliert werden [17]. Im Wesentlichen sind der einschlägigen Literatur im Zusammenhang der spannungsabhängigen Verbraucherlastmodellierung und den damit einhergehenden Netzführungsstrategien zwei Modelle für den stationären Fall zu entnehmen [94]:

- Das ZIP-Modell sowie
- das Exponential-Modell.

Während letzteres der Bezeichnung entsprechend die Modellierung mit Hilfe eines Exponenten vornimmt, erfolgt die Modellierung spannungsabhängiger Prozesse beim ZIP-Modell mit Hilfe eines Polynoms 2.-Ordnung [87]. Die Namensgebung korrespondiert mit den jeweiligen Koeffizienten Z , I und P womit das jeweilige Verhalten der elektr. Anlagen und Gerätschaften bei einer Spannungsänderung beschrieben wird (im folgenden Abschnitt 2.3.1 konkretisiert).

Im Rahmen der vorliegenden Arbeit wird das ZIP-Modell zur Modellierung der Spannungsabhängigkeit im Zusammenhang mit dem Einspeiseprozess durch PVA sowie der Verbraucherlastmodellierung einzelner Haushalts-Endkunden mit dem erarbeiteten BUV genutzt. Dies ist in der besseren Verfügbarkeit aktueller Modell-Koeffizienten für einzelne elektr. Haushaltsgeräte begründet. Die Modellierung der Spannungsabhängigkeit im Zuge des TDV berücksichtigt beide Modelle, jedoch wird im Rahmen der vorliegenden Arbeit das Augenmerk auf die Modellierung mittels ZIP-Modell gelegt.

Vor diesem Hintergrund wird nachfolgend auf die analytische Beschreibung des Verhaltens der spannungsabhängigen Verbraucherlast mit Hilfe des ZIP-Modells eingegangen sowie auf die notwendigen Anpassungen der Modelle. Hierfür werden die einzelnen Ergänzungen der zuvor beschriebenen Last- und Einspeiseprozesse vorgestellt. Zum einen wird in Abschnitt 2.3.1 die Nutzung des anerkannten ZIP-Modells [95], und zum anderen eine optionalen Belegung von energiekonstanten Lasten beschrieben (in Abschnitt 2.3.2). Wie einleitend in Abschnitt 1.2 thematisiert, weisen die Autoren in [21] sowie in [48] auf die Notwendigkeit der Berücksichtigung des energiekonstanten Verhaltens ($E = \text{konstant}$) einzelner Geräte und Gerätegruppen im Kontext einer aktiven Anpassung der Netzspeisung hin. Insbesondere der damit einhergehende DCR, welcher mit einem zeitlichen Verzug zu beobachten ist, wird gesondert in Abschnitt 2.3.2 behandelt. Dieser wird in wenigen Untersuchungen benannt und i.d.R. vernachlässigt, sodass eine Bewertung einer expliziten Berücksichtigung des DCR im Rahmen der vorliegenden Arbeit thematisiert wird.

2.3.1. ZIP-Modell

Zur Berücksichtigung spannungsabhängiger Verbraucherleistungen im Rahmen der Netzsimulation werden die beschriebenen Knotenmodelle um das anerkannte ZIP-Modell [95]

$$P(U) = P_r \cdot \left[Z_P \cdot \left(\frac{U}{U_r} \right)^2 + I_P \cdot \frac{U}{U_r} + P_P \right] \quad (2.19)$$

$$Q(U) = Q_r \cdot \left[Z_Q \cdot \left(\frac{U}{U_r} \right)^2 + I_Q \cdot \frac{U}{U_r} + P_Q \right] \quad (2.20)$$

erweitert, um eine Änderung der Wirk- und Blindleistung, $P(U)$ respektive $Q(U)$, um den Arbeitspunkt bei Bemessungsspannung U_r mit den Bemessungsleistungen P_r und Q_r zu modellieren. Den Gleichungen nach 2.19 und 2.20 ist das Modell der Spannungsabhängigkeit direkt zu entnehmen. Demnach erfolgt die Modellierung mit Hilfe eines Polynoms 2.-Ordnung. Dabei beschreibt der Koeffizient $Z_{P,Q}$ ein *impedanzkonstantes*, $I_{P,Q}$ ein *stromkonstantes* und $P_{P,Q}$ ein *leistungskonstantes* Verhalten und wird jeweils für die Wirkleistung P sowie Blindleistung Q angegeben.

Bei mehreren zum Zeitpunkt t gleichzeitig aktiven spannungsabhängigen Prozessen, wie bspw. bei der Verrechnung mehrerer aktiver Gerätegruppen gg des BUV, erfolgt eine gewichtete Verrechnung der ZIP-Koeffizienten über

$$\begin{pmatrix} Z_{P,\text{res}}(t) \\ I_{P,\text{res}}(t) \\ P_{P,\text{res}}(t) \end{pmatrix} = \sum_{gg=1}^N \frac{P_{gg}(t)}{P_{\text{res}}(t)} \begin{pmatrix} Z_{P,gg}(t) \\ I_{P,gg}(t) \\ P_{P,gg}(t) \end{pmatrix} \quad (2.21)$$

$$\begin{pmatrix} Z_{Q,\text{res}}(t) \\ I_{Q,\text{res}}(t) \\ P_{Q,\text{res}}(t) \end{pmatrix} = \sum_{gg=1}^N \frac{Q_{gg}(t)}{Q_{\text{res}}(t)} \begin{pmatrix} Z_{Q,gg}(t) \\ I_{Q,gg}(t) \\ P_{Q,gg}(t) \end{pmatrix}, \quad (2.22)$$

mit der Wirk- und Blindleistung $P_{gg}(t)$ bzw. $Q_{gg}(t)$ der Gerätegruppe gg , der Anzahl N gleichzeitig aktiver Gerätegruppen gg und dem zum Zeitpunkt t resultierenden Wirk- und Blindleistungsbedarf $P_{\text{res}}(t)$ bzw. $Q_{\text{res}}(t)$. Dabei gelten stets die Normierungsbedingungen

$$Z_{P,\text{res}}(t) + I_{P,\text{res}}(t) + P_{P,\text{res}}(t) = 1 \quad (2.23)$$

und

$$Z_{Q,\text{res}}(t) + I_{Q,\text{res}}(t) + P_{Q,\text{res}}(t) = 1. \quad (2.24)$$

Demnach wird jedem Modell zur synthetischen Generierung einer Knotenleistung ein Satz an ZIP-Koeffizienten zur Beschreibung der Spannungsabhängigkeit hinzugefügt. Im Allgemeinen kann die Bestimmung der Koeffizienten über eine Messung einzelner Geräte erfolgen, wie in [97] und [98] veranschaulicht. Die resultierenden ZIP-Koeffizienten der einzelnen Geräte können anschließend, unter Berücksichtigung der Durchdringung und des Leistungsbezugs, gewichtet verrechnet und den einzelnen Gerätegruppen zugeordnet werden.

Für die Bewertung eines gesamten Ortsnetzes, wie es mit Hilfe des TDV notwendig wäre, kann wie in [99] vorgegangen werden. Dabei wird bei dem sog. *Notch Test* (vgl. [48]) für einen kurzen Augenblick eine Spannungsänderung vorgenommen. Parallel hierzu wird die stationäre Systemantwort gemessen und abschließend die Koeffizienten abgeleitet.

Zur simulationstechnischen Generierung zeitlich variierender und spannungsabhängiger Verbraucherlastprofile werden den Gerätegruppen gg eines synthetisch generierten Haushalts-Lastprofils nach dem BUV entsprechend Kompositionen von ZIP-Koeffizienten zugeordnet (vgl. Tabelle 2.6). Somit stellt jeder synthetisch generierte Haushalts-Endkunde zu einem beliebigen Simulationszeitpunkt t einen Satz resultierender ZIP-Koeffizienten nach Gleichung 2.21 und 2.22 zur Verfügung. Die Definition der Koeffizienten in Tabelle 2.6 erfolgt in Anlehnung an die Veröffentlichungen [87], [97], [98] sowie [100]. Fehlende elektr. Geräte zur Bildung der Gerätegruppen wurden mit Hilfe einer messtechnischen Untersuchung zur Ableitung von Modellkoeffizienten basierend auf der Regressionsanalyse in [101] ergänzt. Die Koeffizienten der Kühleinheiten sowie des Stand-by wurden der Untersuchung nach [102] entnommen und stellen den mittleren Verlauf der Nachtstunden dar.

Im Allgemeinen erfolgt die Berücksichtigung des ZIP-Modells im Rahmen einer Leistungsflussberechnung wahlweise über einen der folgenden zwei Wege:

1. Innerhalb der Modellierung der Last- und Erzeugerprozesse: Anpassung des Leistungsbezugs der Knotenmodelle nach Gleichung 2.19 und 2.20 durch Übermittlung der Knotenspannung.
2. Innerhalb der Leistungsflussberechnung: Direkte iterative Bestimmung der Knotenleistungen unter Berücksichtigung der resultierenden, gewichteten ZIP-Koeffizienten der Netzknoten nach [95].

Beide Ansätze führen zu einer resultierenden Anpassung des Leistungsbedarfs in Abhängigkeit der am Knotenpunkt anliegenden elektr. Spannung. Abbildung 2.19 zeigt das beschriebene ZIP-Modell im Rahmen eines Untersuchungsszenarios. Hierbei wurden drei identische Szenarien der Netztopologie nach Anhang A.7 simuliert, welche lediglich eine unterschiedliche Spannung am Bezugsknoten aufweisen. Simuliert wurde jeweils mit einer Spannungsanpassung von $\pm 8\%$ bei einer Nennspannung von $U_0 = 400\text{ V}$. Das obere Diagramm zeigt die kumulierte Wirkleistung P_{ONS} an der ONS im Tagesverlauf. Die untere Darstellung zeigt den Einfluss der modellierten Spannungsabhängigkeit $P(U)$ eines einzelnen Netzknotens der simulierten Netztopologie. Beiden Diagrammen ist eine im Tagesverlauf variierende Spannungsabhängigkeit zu entnehmen. Dies fällt insbesondere beim Vergleich der Mittags- und Abendstunden ins Gewicht. Darüber hinaus ist hervorzuheben, dass auch die Spannungsabhängigkeit einer Mittelung unterliegt. Während die Spannungsabhängigkeit des Netzknoten 93 in den Abendstunden (z. B. zwischen 22:00 – 23:00 Uhr) eine Tendenz hin zum *leistungskonstanten* Verhalten aufweist, kann der kumulierten Wirkleistung P_{ONS} des oberen Diagramms nach Abbildung 2.19 ein vermeintlich *stromkonstantes* Verhalten am *Morgen*, *Mittag* und *Abend* entnommen werden. Im Gegensatz hierzu zeigt der Verlauf der Wirkleistung des unteren Diagramms nach Abbildung 2.19 innerhalb des *Zeitbereichs*

Mittag unterschiedliche Ausprägungen der Spannungsabhängigkeit. So weisen die Wirkleistungsprofile ein *leistungskonstantes* Verhalten zwischen 12:00 – 13:00 Uhr auf, während zwischen 11:00 – 11:45 Uhr eine deutliche Ausprägung der Spannungsabhängigkeit zu erkennen ist.

Einspeiseprozesse, vornehmlich durch PVA, sind in Anlehnung an [10] und [72] als *leistungskonstant* modelliert.

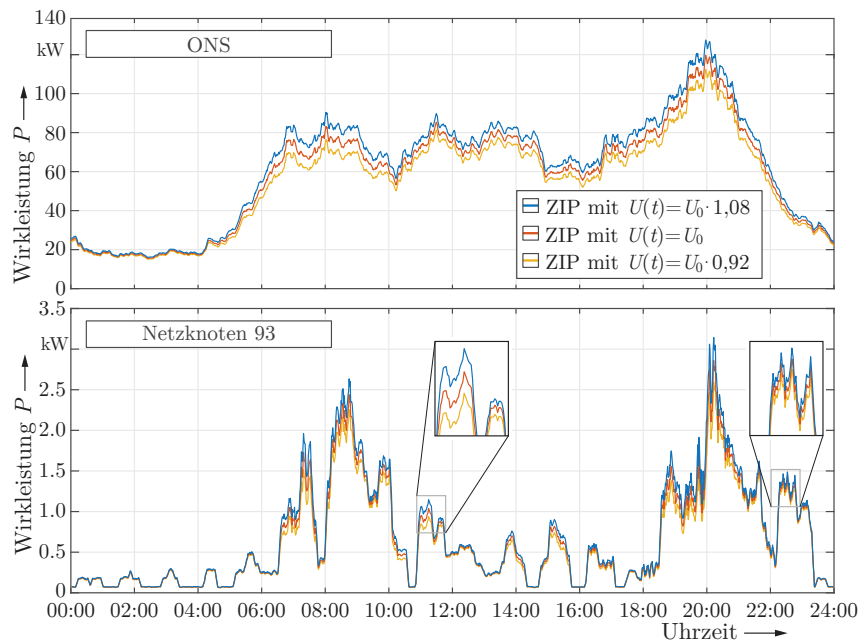


Abbildung 2.19.: Darstellung der Spannungsabhängigkeit $P(U)$ eines Niederspannungsnetzes im Tagesverlauf an zwei Netzknoten. Visualisiert ist der Wirkleistungsbezug an der ONS sowie am Netzknoten 93 im oberen respektive unteren Diagramm. Simuliert wurde jeweils mit einer Spannungsanpassung von $\pm 8\%$ sowie bei einer Nennspannung von $U_0 = 400\text{ V}$. Die unterschiedliche Ausprägung der Spannungsabhängigkeit $P(U)$ im Tagesverlauf sowie an unterschiedlichen Netzknoten ist der Abbildung zu entnehmen.

2.3.2. ZIP(E)-Modell zur Berücksichtigung der Energiekonstanz

Das zuvor beschriebene ZIP-Modell findet häufig in der Beschreibung und Bewertung von spannungsoptimierenden Maßnahmen Anwendung. Hierbei werden die augenblicklichen Folgen einer Spannungsanpassung mit einer hohen Güte beschrieben, jedoch werden langfristige Folgen vom Modell nicht korrekt wiedergegeben [21]. Diese Ungenauigkeit basiert vornehmlich auf Anlagen und Gerätschaften, welche unter die Kategorie der *Prozesswärme* oder *Wärmezwecke im Allgemeinen* fallen und mit Hilfe eines Thermostates sowie einer Steuerung oder Regelung betrieben werden, wie z. B. ein Untertischgerät, elektr. Heizung, elektr. Herd oder ein Wasserkocher. Auch wenn Maßnahmen zur Steigerung der Energieeffizienz in einer allgemeinen Reduzierung des Energiebedarfs der Haushalts-Endkunden resultieren (9% in den vergangenen zehn Jahren), so werden nach wie vor 30% des JEB auf Anwendungen der *Prozesswärme* und ca. 50% auf *Wärmezwecke im Allgemeinen* angewendet [103]. Dies gilt es insbesondere in der Modellierung korrekt zu berücksichtigen.

Demnach ist die durch Pinney in [21] beschriebene Abweichung im Wesentlichen auf eine konstante Energiebedingung

$$E = \text{konstant} \quad (2.25)$$

zurückzuführen (vgl. Kapitel 1.2). Abbildung 2.20 soll rein qualitativ die Auswirkung auf den Leistungsbezug, hervorgerufen durch die konstante Energiebedingung nach Gleichung 2.25 des ZIP(E)-Modells, im direkten Vergleich zum ZIP-Modell veranschaulichen. Schemenhaft dargestellt ist dreimal derselbe Prozess bei unterschiedlicher Spannung, modelliert mit dem ZIP- (links) sowie dem ZIP(E)-Modell (rechts). Während beim ZIP-Modell eine Anpassung der Wirkleistungsaufnahme für die Spannungen $U > U_r$, $U = U_r$ und $U < U_r$ zu beobachten ist, so ist beim ZIP(E)-Modell unter Einhaltung der konstanten Energiebedingung nach Gleichung 2.25 eine zusätzliche Anpassung der DOU explizit modelliert. Demnach kann zwischen den kurz- und langfristigen Folgen einer Spannungsanpassung auf die Verbraucherleistungen unterschieden werden. Der kurzfristige Effekt wird durch das ZIP-Modell adäquat beschrieben, wie bereits in Abschnitt 2.3.1 erläutert und in Abbildung 2.19 und 2.20 (links) dargestellt. Zur Verdeutlichung der langfristigen Folgen einer Spannungsanpassung sollen nachfolgend zwei mögliche Szenarien beschrieben werden, welche unter Einhaltung der konstanten Energiebedingung nach Gleichung 2.25 eine Spannungsabsenkung und Spannungserhöhung gegenüber der Referenzspannung U_r darstellen.

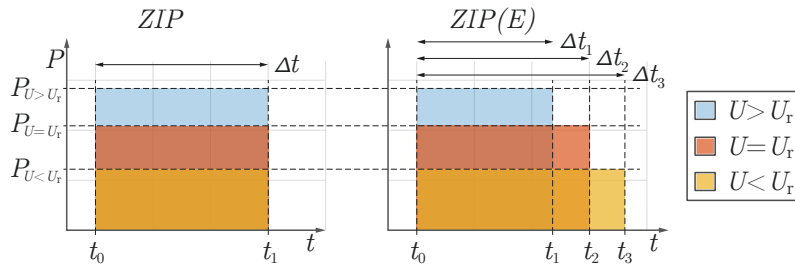


Abbildung 2.20.: Vergleich der Modelle ZIP (links) und ZIP(E) (rechts). Die konstante Energiebedingung $E = \text{konstant}$ und der damit einhergehenden längeren Nutzungsdauer (mit Δt_3 bei $U < U_r$) respektive kürzeren (mit Δt_1 bei $U > U_r$) ist dem Vergleich zu entnehmen.

Wie oben beschrieben, verlängert eine langfristig anhaltende Spannungsabsenkung die Nutzungsdauer einzelner elektr. Geräte. So führt bspw. eine Reduzierung der Spannung $U < U_r$ zu einer direkten, den ZIP-Koeffizienten folgenden Anpassung des Leistungsbezugs zum Zeitpunkt t . Als langfristige Folge führt diese Anpassung bei jenen Geräten zu einer Verlängerung der Nutzungsdauer $\Delta t_{\text{ZIP(E)}} > \Delta t$, welche eine konstante Energiebedingung $E = \text{konstant}$ aufweisen. Umgekehrt kann eine Spannungserhöhung $U > U_r$ dazu führen, dass bedingt durch diese Spannungsanpassung eine sofortige Zunahme des Leistungsbezugs erfolgt sowie eine kürzere Nutzungsdauer einzelner elektr. Anlagen und Gerätschaften eintritt (vgl. Abbildung 2.20 rechtes Diagramm). In Konsequenz führt dieses beschriebene energiekonstante Verhalten zum DCR [48].

Die Ursache für den DCR könnte in Folge einer Spannungsabsenkung dadurch begründet sein, dass elektr. Geräte parallel betrieben werden, welche ohne Anpassung der Spannung

Tabelle 2.6.: Ergänzung der Datengrundlage des Verbraucherlastmodells nach dem BUV nach Tabelle 2.4. Ein Satz resultierender ZIP-Koeffizienten je Gerätegruppe sowie die Belegung der energiekonstanten Gerätegruppen sind der Tabelle zu entnehmen.

KLASSIFIKATION	JAHRESARBEIT ANTEIL (%)	ZEITBEREICH			KOEFFIZIENTEN						
		MORGEN	MITTAG	ABEND	Z_P	I_P	P_P	Z_Q	I_Q	P_Q	$E = \text{konstant}$
Kühleinheiten	16,7		Automatik		0,17	0,01	0,82	0,93	0,00	0,07	✓
Unterhaltung, Büro	25,7	-	✓	✓	0,08	0,07	0,85	0,04	0,00	0,96	-
Warmwasser	14,8	✓	✓	✓	1,00	0,00	0,00	0,80	0,16	0,04	✓
WTS ¹	12,4	✓	✓	✓	0,10	0,10	0,80	1,54	-1,43	0,89	✓
Kochen	9,8	✓	✓	✓	1,00	0,00	0,00	0,69	0,25	0,06	✓
Beleuchtung	8,1	✓	-	✓	0,14	0,77	0,09	-0,06	-0,34	1,40	-
Sonstige	12,5	✓	✓	✓	0,41	0,30	0,29	3,10	-3,90	1,80	✓
Stand-By	11,0 ²		Dauerbetrieb		0,17	0,01	0,82	0,93	0,00	0,07	-

¹Waschmaschine, Trockner und Spülmaschine

²Der Anteil bezieht sich auf die gesamte Jahresarbeit

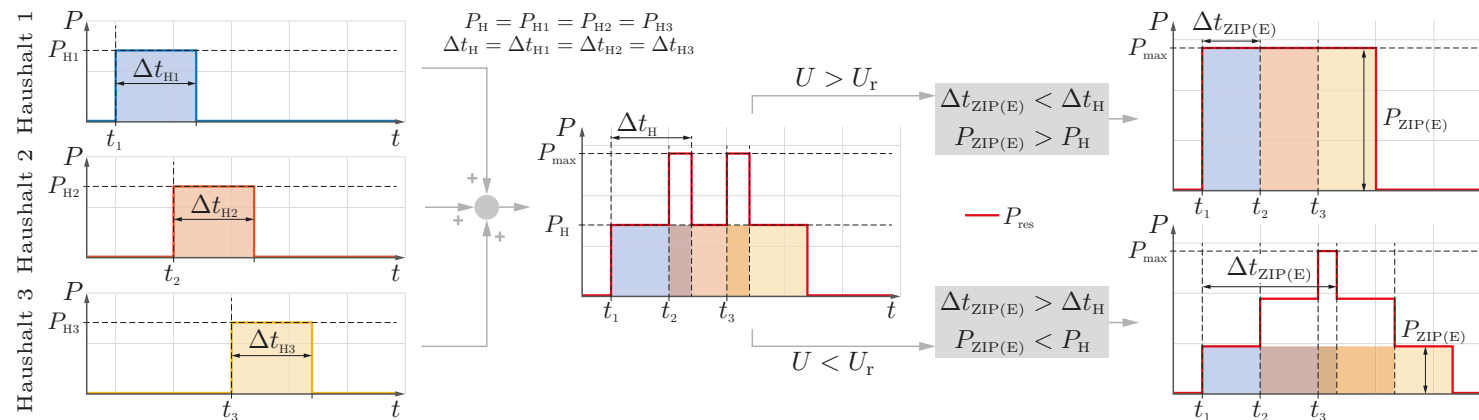


Abbildung 2.21.: Schematische Darstellung des DCR. Der resultierende Leistungsbezug P_{max} bei einer Spannungsabsenkung $U < U_r$ wird teilweise kompensiert, obwohl $P_{ZIP(E)} < P_H$ gilt. Dies ist durch das energiekonstante Verhalten einzelner Geräte bedingt, welches eine Verlängerung der Nutzungsdauer $\Delta t_{ZIP(E)}$ zur Folge hat. Entsprechend umgekehrtes Verhalten tritt bei einer Spannungsanhebung $U > U_r$ ein.

nicht gleichzeitig eingeschaltet wären. Dies hat einen direkten Einfluss auf die bezogene Leistung. Die Reduktion der Spitzenlast bspw. kann durch diesen Effekt teilweise kompensiert werden oder im schlechtesten Fall eine neue Leistungsspitze P_{\max} zur Folge haben [21]. Umgekehrt kann eine Spannungserhöhung zur Folge haben, dass Prozesse, die teils parallel ablaufen würden, nun bedingt durch eine kürzere Nutzungsdauer $\Delta t_{\text{ZIP(E)}} < \Delta t$ nacheinander betrieben werden. Dieses führt letztlich zu einer Kompensation des resultierenden Leistungsbezugs. Die Abbildung 2.21 visualisiert den beschriebenen Zusammenhang. Der Abbildung ist ein Summenknoten zu entnehmen, an welchem drei Haushalte mit einem Leistungsbezug $P_{\text{H}} = P_{\text{H1}} = P_{\text{H2}} = P_{\text{H3}}$ zu unterschiedlichen Zeitpunkten $t_1 \neq t_2 \neq t_3$ verbunden sind. Die Berücksichtigung der Nutzungsdauer $\Delta t_{\text{H}} = \Delta t_{\text{H1}} = \Delta t_{\text{H2}} = \Delta t_{\text{H3}}$ einzelner Haushalts-Endkunden führt im zeitlichen Verlauf zu einem resultierenden Leistungsbezug P_{res} , dargestellt in rot. Mittig in der Abbildung visualisiert ist der resultierende Leistungsbezug bei Bemessungsspannung U_{r} . Auf der rechten Seite der Abbildung sind die beschriebenen Szenarien mit einer Spannungserhöhung $U > U_{\text{r}}$ und der daraus resultierenden Nutzungsdauer $\Delta t_{\text{ZIP(E)}} < t_{\text{H}}$ (oben) sowie einer Spannungsabsenkung $U < U_{\text{r}}$ und der daraus resultierenden längeren Nutzungsdauer $\Delta t_{\text{ZIP(E)}} > t_{\text{H}}$ (unten) dargestellt.

Einige Untersuchungen im Kontext der Modellierung spannungsabhängiger Prozesse vernachlässigen oder thematisieren den oben beschriebenen Einfluss nicht (vgl. [10], [31], [32], [87] sowie [96]), obwohl die Notwendigkeit zur Berücksichtigung dieser Charakteristik und dessen Relevanz in Publikationen wie [21], [48] und [104] sowie nach der Richtlinie zur *Berechnung der Steigerung der elektrischen Energieeffizienz durch den Einsatz von elektrischen Energiereglern nach dem Prinzip der Spannungsabsenkung (VDE-AR-E 2055-1:2009-10)* entnommen werden kann. Insbesondere in [48] weist der Autor explizit auf den beschriebenen *Energie-Kompensations-Effekt* hin. Willis deutet an, dass der reduzierte Leistungsbezug im Zuge einer Spannungsabsenkung einen „Kampf“ entgegen dieser Anpassung zur Folge hat. Dieses *Wettmachen* von Energieeinsparung in Folge des DCR wirkt dabei negativ im Sinne von Netzführungsstrategien, wie z. B. dem CVR und begründen letztlich die von Pinney in [21] dokumentierten Abweichungen. Demnach kann festgehalten werden, dass das Vernachlässigen der energiekonstanten Lasten und der damit einhergehende DCR zu einer Abweichung der Güte der Netzführungsstrategie führt. So ist bspw. beim CVR mit einem schlechteren CVR-Faktor nach Gleichung 3.1 zu rechnen als dies durch eine Beschreibung mit Hilfe gängiger analytischer Vorschriften suggeriert wird. Zudem fehlen umfangreiche Untersuchungen im Kontext der intelligenten Netzführung unter Berücksichtigung des DCR.

Diese Lücke soll im Rahmen der vorliegenden Arbeit adressiert und im nachfolgenden Kapitel im Detail beleuchtet werden. Zudem sollen Prozesse mit einer konstanten Energiebedingung explizit modelliert und im Kontext spannungsoptimierender Maßnahmen innerhalb von Netzführungsstrategien berücksichtigt werden. Hierfür wird das in Abschnitt 2.1.2 vorgestellte BUV neben der in Abschnitt 2.3.1 beschriebenen Berücksichtigung des ZIP-Modells mit einer Funktion ausgestattet werden, die sicherstellt, dass die umgesetzte Ener-

gie auch bei angepasster Spannung konstant bleibt. Dies führt zum hier vorgestellten ZIP(E)-Modell. Die Gerätegruppen, welche im Rahmen der Modellierung einen energiekonstanten Anteil besitzen, sind in Tabelle 2.6 gekennzeichnet. Die Belegung der energiekonstanten Gerätegruppen nach Tabelle 2.6 basieren im Allgemeinen auf [21], [48] und [103]. WTS bspw., haben eine Heizphase, dessen Nutzungsdauer im Zuge einer Spannungsanpassung alterniert. Auf der anderen Seite ist die Beleuchtung von den jeweiligen Individuen abhängig und wird aktiv zu- und abgeschaltet. Demnach ist die Nutzungsdauer unabhängig von einer Spannungsanpassung und entsprechend nicht als energiekonstant modelliert.

Spannungsanpassung als Flexibilitätsoption im elektr. Verteilnetz

Die Entwicklung neuer Betriebsführungsstrategien ist durch die integrierte Energiewende sowie die Sektorenkopplung getrieben und resultiert in einer zunehmenden Flexibilisierung des Energieversorgungssystems. Dies erfordert insbesondere auf den unteren Spannungsebenen neue Betrachtungsweisen und Konzepte in der Netzführung und -planung. Hierbei wird der Fokus vornehmlich auf die Einhaltung der Spannungsbandgrenzen und Vermeidung von Betriebsmittelüberlastungen gelegt, was u.a. durch ein gezieltes Ansteuern von PVA oder den Einsatz rONT erreicht wird. Oftmals ermöglichen diese Maßnahmen eine Verzögerung oder sogar eine Vermeidung von Netzverstärkungsmaßnahmen [9] sowie eine Erhöhung der Netzanschlusskapazität [20].

Die zuvor in Abschnitt 2.3 beschriebenen Eigenschaften der spannungsabhängigen Verbraucherleistung kann hierbei einen zusätzlichen, passiven Beitrag leisten. Dabei gilt es die inhärenten Eigenschaften der elektr. Lasten dahingehend zu bewerten, wie diese kostensenkend in die bestehenden Prozesse der VNB integriert werden können. Je nach Beschaffenheit und technologischem Aufbau einer elektr. Last kann der resultierende Leistungsbezug aufgrund von Schwankungen der Versorgungsspannung unterschiedlich sein [96]. Diese Eigenschaft soll durch eine bewusst herbeigeführte Spannungsanpassung gezielt als Flexibilitätsoption adressiert werden.

Der einschlägigen Literatur sind unterschiedliche Betriebsführungsstrategien, wie z. B. das Lastmanagement, Blindleistungsmanagement oder CVR, zu entnehmen. Insbesondere Letzteres beruht auf den inhärenten Eigenschaften elektr. Lasten. Im Fokus steht hierbei eine bewusst herbeigeführte Spannungsanpassung mit folgenden Zielen:

- Eine Verringerung des Energiebedarfs sowie
- die Verschiebung von Lastspitzen bzw. Lastzeiten.

Demnach kann das CVR im Allgemeinen dem Bereich des Lastmanagements bzw. der Laststeuerung zugeordnet werden [48]. Eine umfassende Aufarbeitung der Ziele und die ursprüngliche Intention des CVR werden in [10] und [104] thematisiert.

Die Bewertung einer gezielt herbeigeführten Spannungsanpassung im Kontext des CVR erfolgt unter Anwendung des CVR-Faktors [104]

$$CVR_f = \frac{\% \text{ Veränderung der Optimierungsgröße}}{\% \text{ Anpassung der Spannung}} \quad (3.1)$$

Dieser beschreibt die prozentuale Veränderung der Optimierungsgröße, z. B. des Energiebedarfs, in Abhängigkeit der prozentualen Spannungsanpassung. Der resultierende CVR_f wird hierbei zum Vergleich und der Bewertung der Maßnahme im jeweiligen Netzabschnitt im Zuge des CVR herangezogen [104].

Das Kapitel 3 der vorliegenden Arbeit untersucht die kurz- und langfristigen Folgen einer Spannungsanpassung, welche als Flexibilitätsoption im elektr. Niederspannungsnetz Berücksichtigung finden sollen. Eine umfangreiche Bewertung der Prozesse, insbesondere im Kontext des DCR, ist der einschlägigen Literatur nicht zu entnehmen. Vielmehr wird dieser häufig vernachlässigt (vgl. bspw. [31] oder [104]). Die Untersuchung innerhalb des Kapitels 3 der vorliegenden Arbeit soll zum einen diese Lücke erstmals im Detail adressieren, zum anderen die Notwendigkeit der expliziten Berücksichtigung des DCR im Kontext von Netzführungsstrategien untermauern. Hierfür wird zunächst die Implementierung der spannungsabhängigen Modelle nach Abschnitt 2.3 verifiziert. Anschließend wird eine aktiv herbeigeführte Spannungsanpassung als Flexibilitätsoption im elektr. Verteilnetz im Zuge von zwei Untersuchungsszenarien bewertet. Dabei wird das Augenmerk auf den direkten, kurzfristigen Effekt sowie auf die langfristigen Auswirkungen der bewusst herbeigeführten Spannungsanpassung gelegt. Für die beiden Untersuchungsszenarien wird ein reales Niederspannungsnetz herangezogen und zum einen als *passives* Netz (Netz *ohne* DEA), zum anderen als *aktives* Netz betrachtet und bewertet.

In Anbetracht der Bandbreite möglicher Untersuchungs- und Optimierungsszenarien stehen im Kapitel 3 folgende drei Anwendungen im Vordergrund der Betrachtung, um die kurz- und langfristigen Folgen einer Spannungsanpassung im Kontext der Flexibilisierung des elektr. Verteilnetzes zu bewerten und die Modellannahmen nach Abschnitt 2.3 zu verifizieren:

1. Spitzenlastabsenkung,
2. Reduktion der Energieaufnahme des Netzes sowie
3. Konsequenzen der Schalthandlungen auf Netzverluste

am *Morgen*, *Mittag* und *Abend*. Zur Minderung des numerischen Aufwands werden die Nachtstunden per Definition ausgeschlossen. Zudem ist mit Bezug auf die Leistungsaufnahme des Netzes mit einem vergleichsweise geringen Einfluss zu rechnen (vgl. Abbildung A.2 des Anhangs).

Das wesentliche Ziel der nachfolgenden Untersuchung ist das Untermauern einer expliziten Berücksichtigung des DCR mit Hilfe der folgenden drei Thesen:

- Die spannungsabhängige Leistungsaufnahme elektr. Verbraucher kann vorteilhaft im Rahmen der Betriebsführung der VNB im Tagesverlauf genutzt werden.
- Die Zeitpunkte einzelner Schalthandlungen sind entscheidend für die jeweilige Optimierungsgröße.
- Eine gleichzeitige Reduktion der Spitzenlast, Energieaufnahme und Netzverluste ist möglich.

Vorweggenommen sei an dieser Stelle, dass die empirische Verifizierung der Ergebnisse im realen System eine herausfordernde Aufgabe darstellt, welche nur mit einem erheblichen Aufwand umzusetzen ist. Jedes Netz muss individuell untersucht werden, da die beschriebenen Effekte stark von der Verbraucherstruktur sowie DEA des jeweiligen Netzgebietes abhängen. Dies beinhaltet eine aufwändige messtechnische Identifikation der ZIP-Koeffizienten mit Hilfe des *Notch Test* [99], welche eine Vielzahl von Spannungsänderungen im Tagesverlauf vorsieht. Dessen ungeachtet sollen die folgenden simulativen Untersuchungen das Potenzial aufzeigen.

3.1. Netzsimulation spannungsabhängiger Prozesse

Nachfolgend sollen die in Abschnitt 2.3 vorgestellten Erweiterungen der Netzsimulation nach [51] verifiziert und die beschriebenen Effekte veranschaulicht werden. Hierfür wird zunächst der Rahmen der Netzsimulation vorgestellt. Anschließend wird der Terminus *Schaltzeit-tupel* sowie die Zielfunktionen definiert, welche zur numerischen Lösung der Untersuchungs- und Optimierungsszenarien des nachfolgenden Abschnitts 3.2 notwendig sind. Abschließend werden die Auswirkungen des ZIP(E)-Modells auf die drei oben definierten, zu optimierenden elektr. Größen (Spitzenlast, Energieaufnahme sowie Netzverluste) im Einzelnen beschrieben und der Unterschied zum ZIP-Modell hervorgehoben. Die für die nachfolgenden Untersuchungsszenarien notwendigen Leistungssenken werden mit Hilfe des vorgestellten BUV erzeugt. Dies ist in einer besseren Verfügbarkeit aktueller Modell-Koeffizienten einzelner elektr. Gerätschaften begründet.

3.1.1. Definition der Netzsimulation

Zur Verifizierung der Modellierung der Spannungsabhängigkeit innerhalb der Netzsimulation soll nachfolgend das in Anhang A.7 visualisierte Niederspannungsnetz herangezogen werden. Im Rahmen der Netzsimulation wird der Tagesverlauf in minütlicher Auflösung $t \in \mathcal{T} := \{1, 2, \dots, 1440\}$ min nachgebildet. Die Bewertung wird dabei auf die Bereiche reduziert, in welchen eine ausgeprägte Spannungsabhängigkeit zu erwarten ist (Vgl. Abschnitt 2.3.1). Demnach gilt per Definition: *Zeitbereich* $\in \{\text{Morgen}, \text{Mittag}, \text{Abend}\}$ mit *Morgen* $:= \{360 \dots 615\}$ min $\subset \mathcal{T}$, *Mittag* $:= \{616 \dots 940\}$ min $\subset \mathcal{T}$ und *Abend* $:= \{941 \dots 1335\}$ min $\subset \mathcal{T}$. Dabei umfasst das Simulationsszenario je *Zeitbereich* die Schalthandlungen SH₁ und SH₂ sowie die jeweiligen Zeitpunkte der Schalthandlungen $t_{\text{SH}_1}, t_{\text{SH}_2} \in \mathcal{T}$. Die Spannungsanpassung $f_{\Delta U_{\text{SH}}}$ erfolgt je Schalthandlung durch eine Anpassung der

Netzspeisespannung am Bezugsknoten (Slack Bus) für jeden *Zeitbereich* nach

$$U(t) = \begin{cases} U_0 & t < t_{\text{SH}_1} \\ U_0 \cdot f_{\Delta U_{\text{SH}}} & t_{\text{SH}_1} \leq t < t_{\text{SH}_2} \\ U_0 & t_{\text{SH}_2} \leq t \end{cases} \quad (3.2)$$

mit der Nennspannung U_0 und der relativen Spannungsanpassung $f_{\Delta U_{\text{SH}}}$ im Moment der Schalthandlung. Gleichung 3.2 ist zu entnehmen, dass eine Schalthandlung mit $U(t) = U_0 \cdot f_{\Delta U_{\text{SH}}}$ zum Zeitpunkt t_{SH_1} eine zweite Schalthandlung zurück zum Ausgangszustand mit $U(t) = U_0$ zum Zeitpunkt t_{SH_2} zur Folge hat. Demnach erfolgen per Definition je zwei Schalthandlungen je *Zeitbereich*.

3.1.2. Schaltzeitupel und Zielfunktionen

Für einen numerischen Lösungsansatz, welcher die Problemstellung der Ableitung optimaler Schaltzeitpunkte für die oben definierten Optimierungsgrößen je *Zeitbereich* adressiert, bedarf es einer formalen Beschreibung. Da per Definition einer ersten Schalthandlung SH_1 zum Zeitpunkt t_{SH_1} eine zweite Schalthandlung SH_2 zum Zeitpunkt t_{SH_2} folgt, wird zur Beschreibung der Schaltzeitpunkte ein Schaltzeitupel je *Zeitbereich* wie folgt definiert:

$$(t_{\text{SH}_1}, t_{\text{SH}_2})_{\text{Zeitbereich}} \cdot \quad (3.3)$$

Aufbauend auf der Definition der Schaltzeitupel in Gleichung 3.3 liegt die Identifizierung optimaler Schaltzeitupel $(t_{\text{SH}_1}, t_{\text{SH}_2})_{\text{opt, Zeitbereich}}$ im jeweiligen *Zeitbereich* im Fokus. Hierfür erfolgt für die Untersuchungs- und Optimierungsszenarien des nachfolgenden Abschnitts eine Bewertung der Schalthandlungen je *Zeitbereich* unter Berücksichtigung der folgenden drei Zielfunktionen:

- Die optimale Spitzenlastreduktion an der ONS

$$\min_{(t_{\text{SH}_1}, t_{\text{SH}_2})_{\text{Zeitbereich}}} P_{\max}((t_{\text{SH}_1}, t_{\text{SH}_2})) \quad (3.4)$$

- Die optimale Reduktion der Energieaufnahme des Gesamtnetzes

$$\min_{(t_{\text{SH}_1}, t_{\text{SH}_2})_{\text{Zeitbereich}}} E_{\max}((t_{\text{SH}_1}, t_{\text{SH}_2})) \quad (3.5)$$

- Die optimalen Schaltzeitpunkte zur Reduktion der Gesamtverlustenergie

$$\min_{(t_{\text{SH}_1}, t_{\text{SH}_2})_{\text{Zeitbereich}}} E_{\text{Verluste}}((t_{\text{SH}_1}, t_{\text{SH}_2})) \quad (3.6)$$

Die Auffindung eines optimalen Schaltzeitupels erfolgt auf Basis der im vorangegangenen Abschnitt 3.1.1 beschriebenen Netzsimulation. Grundlage der Analyse ist ein numerisches Experiment über 1000 Realisationen je Schaltzeitupel. Der jeweils hieraus resultierende gemittelte Verlauf wird anschließend hinsichtlich der Zielfunktionen nach Gleichung 3.4,

3.5 und 3.6 einzeln ausgewertet. Zur abschließenden Bewertung des Einflusses der Spannungsanpassung werden die einzelnen Zielfunktionen je *Zeitbereich* unabhängig voneinander betrachtet.

3.1.3. Verifizierung der Modellannahmen

Zur Verifizierung der Modellannahmen nach Abschnitt 2.3 soll Abbildung 3.1 herangezogen werden. Dargestellt ist jeweils der *Zeitbereich = Morgen* mit $Morgen := \{360 \dots 615\} \text{ min} \subset \mathcal{T}$ eines vollständig simulierten Tages d mit $t \in \mathcal{T} := \{1, 2, \dots, 1440\} \text{ min}$. Je Diagramm sind vier unterschiedliche Kennlinien visualisiert, welche wie folgt zu interpretieren sind:

- Die gestrichelten Verläufe in gelb, blau und schwarz resultieren auf der Simulation mit Hilfe des ZIP-Modells, ohne Berücksichtigung eines energiekonstanten Verhaltens einzelner elektr. Geräte.
- Die rote Kennlinie berücksichtigt einzelne Betriebsmittel mit $E = \text{konstant}$ und folgt demnach dem ZIP(E)-Modell.
- Für den gelben Verlauf gilt eine dauerhafte Spannungsanpassung $f_{\Delta U_{SH}} = 1,08$ respektive $f_{\Delta U_{SH}} = 0,92$ für den blauen Verlauf.
- Der schwarze Verlauf zeigt den Referenzwert bei U_0 , ohne Spannungsanpassung mit $f_{\Delta U_{SH}} = 1,0$.
- Die rote Kennlinie weist jeweils eine temporäre Spannungsanpassung $U(t) = U_0 \cdot 0,92$ für das Zeitintervall $t_{SH1} \leq t < t_{SH2}$ auf.

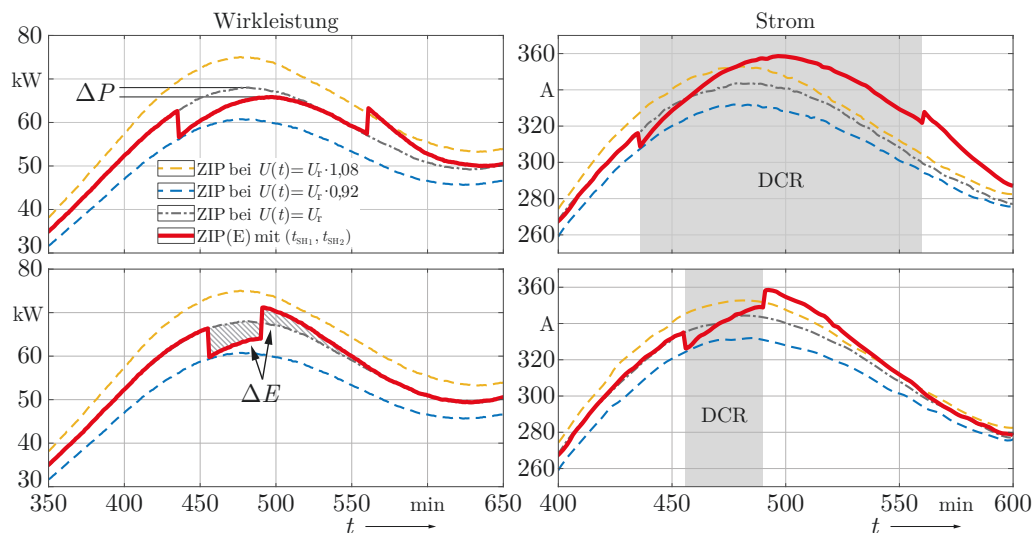


Abbildung 3.1.: Gegenüberstellung resultierender Verläufe der Wirkleistung und des Stromes an der ONS infolge einer Schalthandlung mit Schaltzeitupel (435 min, 560 min)_{Morgen} (oben) und (455 min, 490 min)_{Morgen} (unten). Zusätzlich sind die Referenzverläufe ohne Berücksichtigung von $E = \text{konstant}$ dargestellt. Dem Vergleich sind die Unterschiede bei einer langfristig anhaltenden Spannungsänderung zwischen der Beschreibung mittels ZIP(E)- und ZIP-Modell zu entnehmen.

Festzuhalten ist, dass die Rahmenbedingungen zur Generierung der vier Kennlinien bis auf die genannten Punkte identisch sind. Darüber hinaus stellen die Verläufe ein gemittelt

Verhalten dar, resultierend aus einer Vielzahl von Realisationen. Dies ist auf die hohe Diversität einzelner Realisationen nach dem BUV zurückzuführen, was einer systematischen Analyse entgegen spricht (vgl. Abbildung 2.8). Zudem bleibt die Bewertung des DCR bei der Betrachtung einer einzigen Realisation ohne wesentliche Aussage, da der DCR eine zeitliche Abhängigkeit im Tagesverlauf liefert.

Die vier Diagramme der Abbildung 3.1 verdeutlichen anhand von exemplarischen Schalthandlungen wie unterschiedlich diese im statistischen Mittel auf das Netz wirken. Die beiden oberen Diagramme zeigen die mittleren Auswirkungen auf den Wirkleistungsbezug sowie den resultierenden Strom an der ONS infolge von Schalthandlungen zu den Zeitpunkten $t_{SH_1} = 435$ min und $t_{SH_2} = 560$ min. Für die beiden unteren Diagramme wurden Schalthandlungen zu den Zeitpunkten $t_{SH_1} = 455$ min und $t_{SH_2} = 490$ min durchgeführt. Es ist zu erkennen, dass jeweils zum Zeitpunkt t_{SH_1} eine im Verhältnis zum Referenzwert deutliche Reduzierung der Wirkleistung eintritt. Ein abweichendes Verhalten ist der Stromstärke zu entnehmen. Die Reduzierung des Stroms erfolgt nicht im gleichen Maße. Dieses Verhalten deutet darauf hin, dass zum Zeitpunkt der Schalthandlung ein tendenziell *stromkonstantes* Verhalten vorliegt. Im Allgemeinen ist festzuhalten, dass zum Zeitpunkt t_{SH_1} der Schalthandlung SH_1 das ZIP(E)-Modell (rote Kennlinie) dem Verhalten des ZIP-Modells folgt (blaue Kennlinie). Demnach ist die Beschreibung einer kurzzeitigen Spannungsänderung durch das ZIP-Modell hinreichend genau beschrieben [21]. Erst mit fortschreitender Zeit treten bei einer anhaltenden Spannungsänderung signifikante Abweichungen zwischen der Beschreibung durch das ZIP(E)- und ZIP-Modell auf.

Zudem ist der Abbildung die Wirkung des zeitlich versetzten DCR zu entnehmen. Durch die Berücksichtigung des ZIP(E)-Modells erfolgt bei langfristiger Betrachtung ein sukzessiver Anstieg der Wirkleistung dem Referenzwert entgegen. Ein entsprechend umgekehrtes Verhalten ist bei der Schalthandlung zum Zeitpunkt t_{SH_2} zu beobachten. Hierbei ist hervorzuheben, dass im Zuge des DCR der Verlauf des Stromes den des Referenzwerts überschreitet. Dies lässt den Schluss zu, dass begründet durch den langfristigen Effekt der Schalthandlung ein Übergang in ein tendenziell *leistungskonstantes* Verhalten an der ONS zu beobachten ist. Diese Eigenschaft und das beschriebene Verhalten entsprechen den Beobachtungen der Untersuchung nach [21].

Mit Bezug auf die definierten, zu optimierenden elektr. Größen (s. o.) kann bei der Gegenüberstellung beider Schaltkombinationen in Abbildung 3.1 festgehalten werden, dass sowohl die Spitzenlast (um den Wert ΔP), die Energieaufnahme des Netzes (um den Wert ΔE), als auch die Stromstärke I und demnach die Netzverluste beeinflusst werden. Darüber hinaus ist festzuhalten, dass die Zeitpunkte der Schalthandlung SH_1 sowie SH_2 einen wesentlichen Einfluss auf die Größenordnung von ΔP und ΔE sowie die Stromstärke I haben. Dies soll im Zuge der Untersuchungs- und Optimierungsszenarien des nachfolgenden Abschnitts 3.2 näher analysiert und bewertet werden.

Abschließend bleibt festzuhalten, dass die Verwendung des ZIP(E)-Modells signifikante Unterschiede zum ZIP-Modell aufzeigt. Den Lastverläufen in Abbildung 3.1 ist die kompensierende Wirkung des DCR auf eine aktive Spannungsänderung mit fortschreitender Zeit zu entnehmen. Hiervon sind die oben definierten Optimierungsgrößen direkt beeinflusst. Die Auswirkung auf die Spitzenlastreduktion ΔP und Reduktion der Energieaufnahme ΔE ist hervorgehoben. Zudem resultiert die Erhöhung der Stromstärke in einer Erhöhung der Gesamtverlustenergie.

3.2. Untersuchungsszenarien

Zur Untermauerung der oben genannten Thesen soll nachfolgend die Berücksichtigung der Spannungsanpassung als Handlungsoption im Rahmen der Flexibilisierung des Energieversorgungssystems bewertet werden. Die Untersuchungen basieren auf der Simulation eines ländlichen Netzgebietes an einem *Werktag* im *Sommer*. Die reale Netztopologie ist am Niederrhein verortet und weist eine hohe Durchdringung an PVA auf [36]. Diese besteht aus ca. 2700 m Kabelstrecke sowie 104 Netzknoten. Eine Verletzung des Spannungsbandes wird zu jedem Simulationszeitpunkt t an den Enden der einzelnen Abgänge sowie an jedem Netzknoten mit generatorischer Leistung überwacht. Der Abbildung A.6 des Anhangs A.7 ist die Ausdehnung des Niederspannungsnetzes sowie die Positionen von PVA und des rONT zu entnehmen.

Mit Hilfe der Untersuchungsszenarien sollen bei einer bewusst herbeigeführten Spannungsanpassung an der ONS die Auswirkungen auf den Leistungsbezug des projektierten Niederspannungsnetzes untersucht werden. Zwei Szenarien stehen dabei im Fokus der Untersuchung. Zum einen wird das Niederspannungsnetz *ohne* DEA (*passives* Netz, Szenario 1) betrachtet und bewertet, zum anderen als *aktives* Netz (Szenario 2). Beim aktiven Netz sorgt die erhöhte Anzahl an DEA im Netzgebiet für eine Rückspeisung in das darüber gelagerte Mittelspannungsnetz. Begründet in einer besseren Verfügbarkeit aktueller Modellkoeffizienten einzelner elektr. Gerätschaften erfolgt die Erzeugung von Leistungssenken an spezifischen Netzknoten der Netzsimulation mit Hilfe des probabilistischen Verbraucherlastmodells nach dem BUV. Für die in Szenario 2 notwendigen Leistungsquellen findet das Modell der PVA des Abschnitts 2.2 Anwendung. Die Einspeisung der PVA wird im Rahmen des Szenarios 2 als *leistungskonstant*, mit Koeffizient $P_P = 1$, modelliert (vgl. [10] sowie [72]). Zusammen mit einer nichtlinearen Leistungsflussberechnung wird je Szenario der Einfluss einer Spannungsanpassung entlang der Zeitachse dahingehend untersucht, welchen Beitrag diese im Zuge der Flexibilisierung des Energieversorgungssystems leisten kann. Dabei liegt der Fokus insbesondere auf einer statistischen Bewertung der Vorgänge um den Einfluss des DCR zu berücksichtigen.

3.2.1. Szenario 1: passives Netz

Im Zuge der Untersuchung des passiven Netzes sollen mit Hilfe der vollständigen Enumeration (der sog. Brute-Force-Methode) alle möglichen Schaltzeitpunkte bewertet werden. Zur

besseren numerischen Handhabung und der Auffindung einer Lösung in einem angemessenen Zeitraum soll der Suchraum mit Hilfe der folgenden vier Punkte begrenzt werden:

- Je *Zeitbereich* $\in \{\text{Morgen}, \text{Mittag}, \text{Abend}\}$ wird ein Raster einer Auflösung $\Delta t_A = 5$ min definiert.
- Die einzelnen *Zeitbereiche* werden jeweils in eine steigende und fallende Flanke, für SH₁ respektive SH₂, aufgeteilt.
- Die Spannungsanpassung wird auf $f_{\Delta U_{SH}} = 0.92$ festgelegt.
- Auf SH₁ mit $U(t) = U_0 \cdot f_{\Delta U_{SH}}$ folgt zeitlich verortet SH₂ zurück auf $U(t) = U_0$.

Mit Bezug auf die ersten beiden Restriktionen gilt z. B. für SH₁ im *Zeitbereich Morgen* := $\{360 \dots 615\}$ für $t_{SH_1} \in \{360, 365 \dots, 475\}$ und für SH₂ $t_{SH_2} \in \{480, 485 \dots, 615\}$. Per Definition wird mit Hilfe der unteren drei Punkte erzwungen, dass eine Spannungsabsenkung vor und die Korrektur dieser nach dem Auftreten einer Leistungsspitze im jeweiligen *Zeitbereich* erfolgt. Zur Verdeutlichung soll Abbildung 3.2 herangezogen werden. Dargestellt sind drei Realisationen der definierten Netzsimulation am Übergabepunkt (ONS). Analog zur zweiten Restriktion ist jeweils *Zeitbereich* $\in \{\text{Morgen}, \text{Mittag}, \text{Abend}\}$ und dessen Unterteilung in eine steigende und fallende Flanke, für SH₁ respektive SH₂, hervorgehoben. Zusammen mit Restriktion drei und vier wird eine Absenkung der Spannung $U(t)$ vor dem Auftreten der jeweiligen Leistungsspitze im *Zeitbereich* gewährleistet.

Demzufolge wird im Rahmen des Szenarios die Auswirkung einer Spannungsabsenkung auf das passive Netz hinsichtlich der Zielfunktionen nach Abschnitt 3.1.2 untersucht. Zur Identifizierung des optimalen Schaltzeitpunkts $(t_{SH_1}, t_{SH_2})_{\text{opt, Zeitbereich}}$ im jeweiligen *Zeitbereich* wird unter Berücksichtigung der beschriebenen Restriktionen jede mögliche Schaltkombination simuliert und hinsichtlich der maximal aufgetretenen Leistung P_{max} , Energieaufnahme E_{max} sowie der Netzverluste E_{Verluste} bewertet. Anschließend wird jeweils das optimale Schaltzeitpunkt je *Zeitbereich* und Zielfunktion nach Abschnitt 3.1.2 bestimmt.

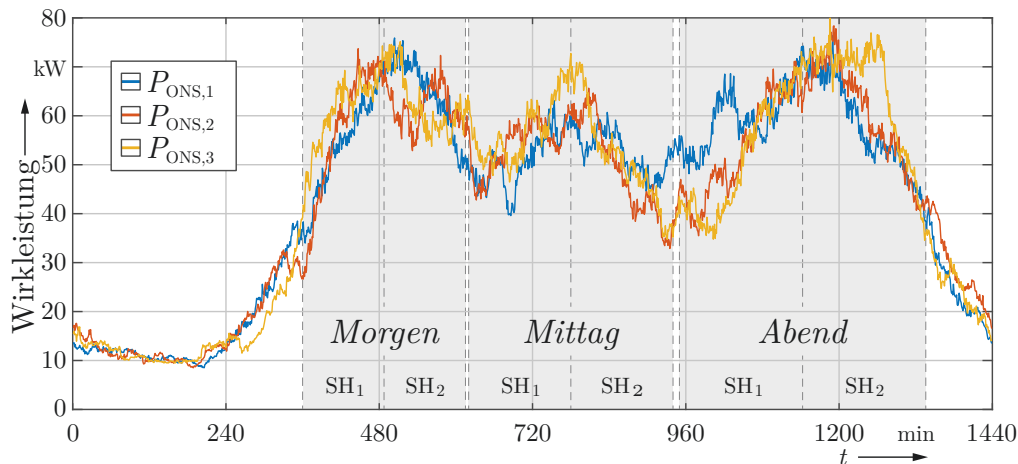


Abbildung 3.2.: Unterteilung je *Zeitbereich* $\in \{\text{Morgen}, \text{Mittag}, \text{Abend}\}$ in eine steigende und fallende Flanke für SH₁ respektive SH₂.

Darstellung der Ergebnisse

Abbildung 3.3 zeigt exemplarisch die Auswertung zur Auffindung eines optimalen Schaltzeitpels mit dem Ziel, der Spitzenlastreduktion für *Zeitbereich = Morgen*. In der dreidimensionalen Abbildung sind die Zeitpunkte $t_{SH_1}, t_{SH_2} \in \mathcal{T}$ der Schalthandlungen sowie der maximal aufgetretene Wirkleistungswert P_{\max} des jeweiligen Schaltzeitpels $(t_{SH_1}, t_{SH_2})_{Morgen}$ im *Zeitbereich Morgen* dargestellt. Darüber hinaus ist der Abbildung eine graue Ebene zu entnehmen. Diese ist als Referenz zu interpretieren und markiert den maximalen Wirkleistungswert, welcher ohne aktive Spannungsanpassung aufgetreten wäre. Es ist deutlich zu erkennen, dass t_{SH_1} im Bereich von Minute 420 bis 460 und t_{SH_2} im Bereich von Minute 545 bis 615 zu einer Verbesserung gegenüber dem Referenzwert (Ref) führen und eine Spitzenlastreduzierung entsprechend der Zielfunktion nach Gleichung 3.4 erzielt wird.

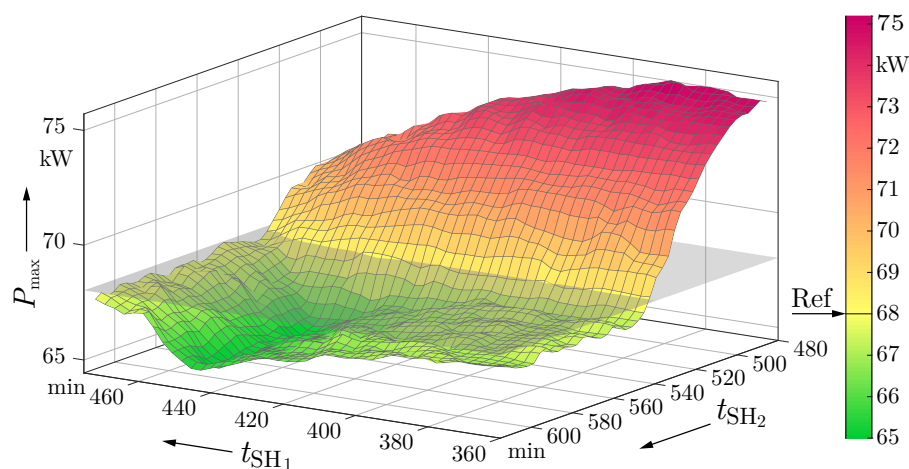


Abbildung 3.3.: Darstellung des Einflusses der Schalthandlung im *Zeitbereich Morgen* auf die Lastspitze im Vergleich zur Referenz (Ref, graue Ebene) ohne Schalthandlung [19].

Zur besseren Darstellung der Ergebnisse und Bewertung von Schalthandlungen mit Bezug auf alle drei Zielfunktionen je *Zeitbereich* werden die Ergebnisse mit Hilfe eines Konturdiagramms in Abbildung 3.4 visualisiert. Hierbei sind die jeweils kleinsten im Untersuchungsszenario aufgetretenen Werte mit grün markiert respektive die größten mit rot. Wiederum ist der jeweilige Referenzwert je Konturdiagramm angegeben. Die Markierung des Referenzwerts im Zusammenhang mit dem gewählten Farbgradienten ermöglichen eine direkte Bewertung der Ergebnisse.

Im direkten Vergleich der Ergebnisse können einige wesentliche qualitative Erkenntnisse für das in diesem Abschnitt beschriebene Szenario hervorgehoben werden:

- Für jeden *Zeitbereich* sind ähnliche Verläufe hinsichtlich der Zielfunktion erkennbar.
- Für die Spitzenlastreduktion P_{\max} können optimale Schaltkombinationen bestimmt werden (rechte obere Ecke).
- Im gewählten Szenario 1 wirken Schalthandlungen positiv auf die Energieaufnahme E_{\max} (der Referenzwert liegt weit im roten Bereich) und negativ auf die Netzverluste E_{Verluste} (der Referenzwert liegt weit im grünen Bereich).

Tabelle 3.1.: Optimale Schaltzeitupel $(t_{SH_1}, t_{SH_2})_{opt, Zeitbereich}$ je $Zeitbereich \in \{Morgen, Mittag, Abend\}$ der einzelnen Zielfunktionen des Szenarios 1.

Zielfunktion	Morgen					Mittag					Abend				
	$(t_{SH_1}, t_{SH_2})_{opt}$ min	ΔP_{max} %	ΔE_{max} %	$\Delta E_{Verluste}$ %	CVR_f	$(t_{SH_1}, t_{SH_2})_{opt}$ min	ΔP_{max} %	ΔE_{max} %	$\Delta E_{Verluste}$ %	CVR_f	$(t_{SH_1}, t_{SH_2})_{opt}$ min	ΔP_{max} %	ΔE_{max} %	$\Delta E_{Verluste}$ %	CVR_f
P_{max}	(445, 560)	-4,53	-0,42	3,93	0,57	(735, 840)	-2,42	-0,25	3,46	0,30	(1130, 1275)	-3,71	-0,46	4,66	0,46
E_{max}	(385, 610)	-1,25	-0,67	7,67	0,08	(620, 925)	-0,93	-0,80	8,37	0,10	(950, 1335)	-1,77	-1,23	9,78	0,15
$E_{Verluste}$	(475, 485)	1,28	-0,06	0,27	-0,03	(755, 765)	2,33	0,04	0,42	-0,05	(1165, 1170)	0,54	0,00	0,16	-0,02

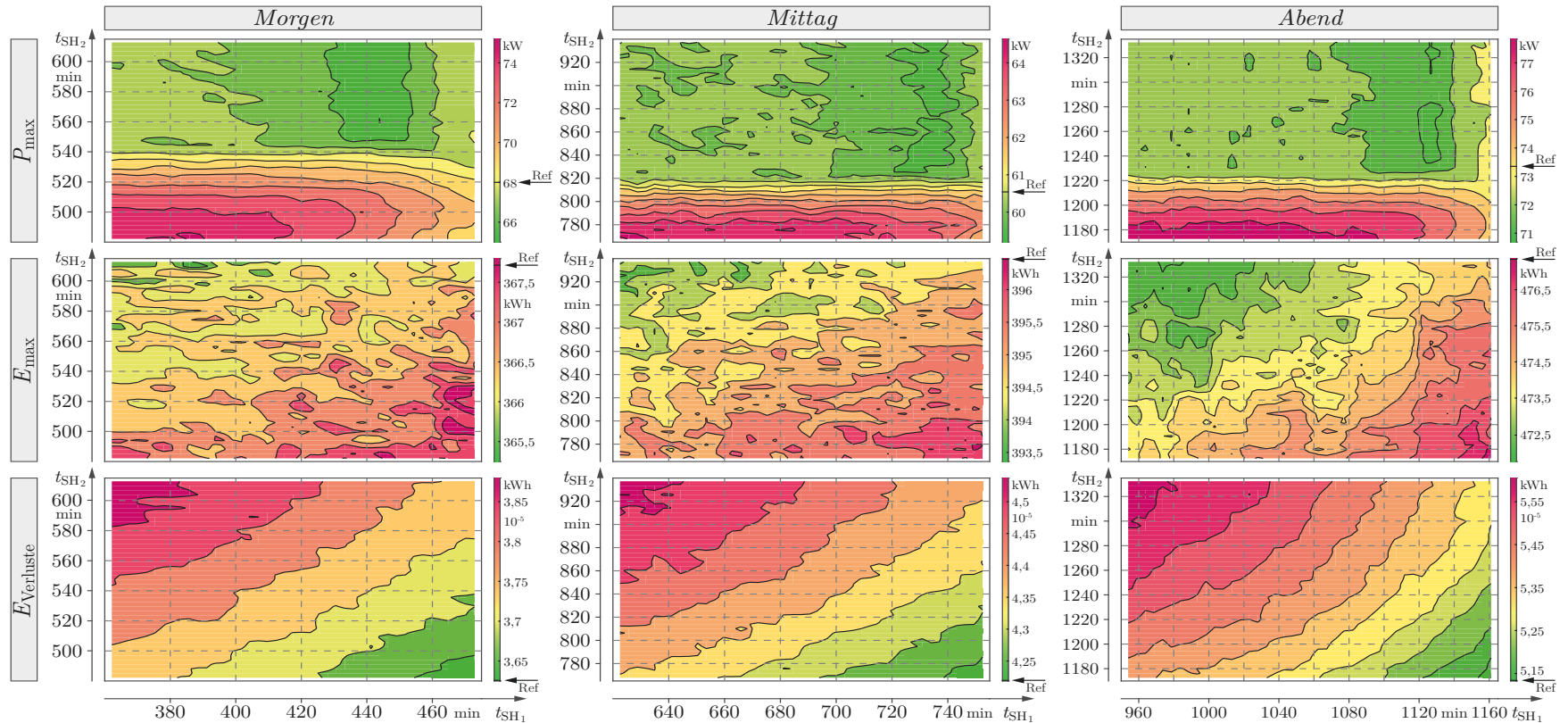


Abbildung 3.4.: Konturdiagramm zur Visualisierung des Einflusses der Schaltzeitupel $(t_{SH_1}, t_{SH_2})_{Zeitbereich}$ je $Zeitbereich$ des Szenarios 1 auf die einzelnen Zielfunktionen. Der jeweilige Referenzwert (Ref) markiert den Wert, welcher ohne aktiver Spannungsanpassung aufgetreten wäre.

- Die reduzierte Energieaufnahme E_{\max} weist im *Zeitbereich Abend* ein ausgeprägtes Minimum für weit auseinander liegende Schaltzeitpunkte auf (linke obere Ecke).
- Eine *gleichzeitige* Reduktion der Spitzenlastreduktion P_{\max} , der Energieaufnahme E_{\max} und der Netzverluste E_{Verluste} ist im Szenario 1 durch eine Schaltkombination nicht zu erreichen.

Zudem weisen die Diagramme der Auswertung hinsichtlich einer reduzierten Energieaufnahme E_{\max} eine Vielzahl lokaler Minima auf. Eine Begründung hierfür kann durch den kleinen Wertebereich sowie der Interpolation zur Erzeugung des Konturdiagramms gegeben sein.

Die optimalen Schaltzeitupel über alle *Zeitbereiche* und Zielfunktionen können Tabelle 3.1 entnommen werden. Zur besseren Bewertung und für einen allgemeinen Vergleich sind ebenfalls die prozentualen Veränderungen und der jeweilige CVR_f nach Gleichung 3.1 dargestellt. Hinsichtlich der Spitzenlastreduktion P_{\max} erreicht der $CVR_{f,P_{\max}}$ im Vergleich zu [104] eine typische Größenordnung (0,30 – 0,57). Der $CVR_{f,E_{\max}}$ der max. Energieaufnahme ist im gewählten Simulationsszenario weniger stark ausgeprägt (0,08 – 0,15). Dies liegt zum einen am DCR und des kompensierenden Einflusses, zum anderen ist dies durch das gewählte Simulationsszenario und die damit verbundenen ZIP(E)-Koeffizienten des BUV begründet. Letztlich bieten nur spannungsabhängige Verbraucher ein Potenzial zur Energieeinsparung. Hierfür ist eine höhere Durchdringung an *impedanzkonstanten* Verbrauchern notwendig, die gleichzeitig kein energiekonstantes Verhalten aufweisen (vgl. *VDE-AR-E 2055-1:2009-10*).

Darüber hinaus ist der Auswertung in Tabelle 3.1 zu entnehmen, dass im Szenario 1 kein Schaltzeitupel eine Reduktion hinsichtlich der Gesamtverlustenergie E_{Verluste} herbeiführt. Dieser Sachverhalt ist mit dem gewählten Simulationsszenario und dem DCR zu begründen. Eine Spannungsabsenkung führt mit fortschreitender Zeit zu einem Stromanstieg in Abhängigkeit der resultierenden Spannungsabhängigkeit, wie in Abbildung 3.1 dargestellt. Dies wiederum führt zu einer Erhöhung der Netzverluste gegenüber dem Referenzszenario.

3.2.2. Szenario 2: aktives Netz

In diesem Abschnitt soll das eingangs beschriebene aktive Netz im Rahmen der vorgestellten Modellierung untersucht werden. Das Szenario 2 unterstellt einen klaren Himmel um den maximalen Einfluss der PVA zu untersuchen. Im Vergleich zum zuvor beschriebenen Szenario 1 sollen neben der Berücksichtigung von generatorischer Leistung in Form von PVA die genannten Restriktionen abgeändert werden. Das Raster der möglichen zeitlichen Positionierung $t_{\text{SH}_1}, t_{\text{SH}_2} \in \mathcal{T}$ der jeweiligen SH_1 und SH_2 wird hierbei von der Netzsimulation mit einer minütlichen Auflösung $\Delta t_A = 1 \text{ min}$ für $t \in \mathcal{T} := \{1, 2, \dots, 1440\}$ min vorgegeben. Des Weiteren wurde im Szenario 1 der Konvention gefolgt, dass der SH_1 zum Zeitpunkt t_{SH_1} mit einer Spannungsanpassung von $f_{\Delta U_{\text{SH}}} = 0.92$ eine zweite SH_2 zum Zeitpunkt t_{SH_2} mit einer korrigierenden Spannungsanpassung zurück auf die Nennspannung U_0 folgt. Diese Konvention wird dahingehend angepasst, dass per Definition SH_1 eine

Spannungsabsenkung zur Folge hat und SH_2 eine Spannungserhöhung. Zudem entfällt die Restriktion vier des Szenario 1 in Gänze, sodass keine zeitliche Reihenfolge von SH_1 und SH_2 besteht. Diese Anpassung führt zu folgenden drei möglichen Ausgangssituationen:

1. Bei $t_{SH_1} < t_{SH_2}$ wird die Spannung zunächst abgesenkt und im weiteren Verlauf zum Zeitpunkt t_{SH_2} zurück auf die Nennspannung U_0 angehoben.
2. Mit $t_{SH_1} > t_{SH_2}$ erfolgt zunächst eine Spannungserhöhung zum Zeitpunkt t_{SH_2} , welcher eine Spannungsabsenkung zum Zeitpunkt t_{SH_1} zurück auf Nennspannung U_0 folgt.
3. Für den Fall $t_{SH_1} = t_{SH_2}$ erfolgt keine Spannungsanpassung.

Hierdurch ist gewährleistet, dass sowohl eine Spannungsabsenkung als auch eine Spannungserhöhung gegenüber der Nennspannung U_0 mit Hilfe der Schaltzeitupel beschrieben werden kann und eine Analyse hinsichtlich der Zielfunktionen nach Gleichung 3.4, 3.5 und 3.6 ermöglicht wird. Zudem wird für das Szenario 2 die Spannungsanpassung durch den rONT auf $\pm 5\%$ festgelegt. Der Abbildung 3.5 sind exemplarisch drei Realisationen nach dem beschriebenen Szenario 2 zu entnehmen. Optisch hervorgehoben sind die jeweiligen *Zeitzbereiche*.

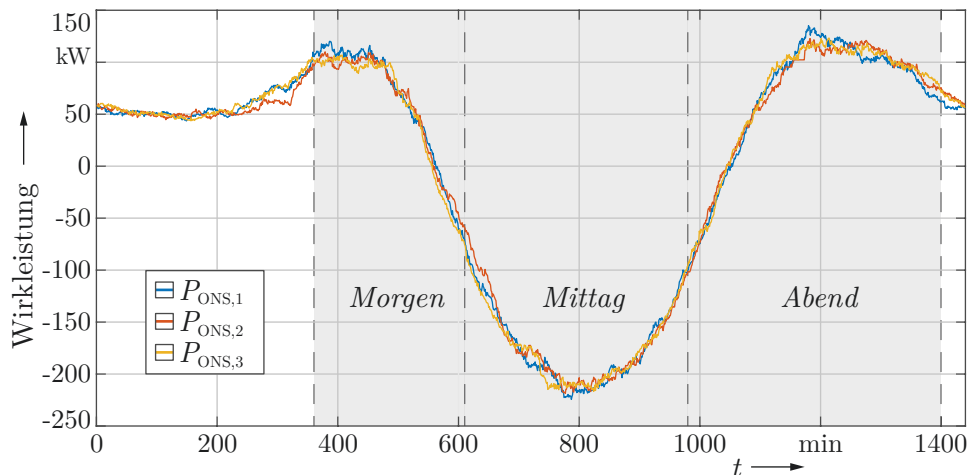


Abbildung 3.5.: Exemplarische Darstellung von drei Realisationen des Szenario 2. Hervorgehoben ist jeweils *Zeitzbereich* $\in \{Morgen, Mittag, Abend\}$.

Die genannten Anpassungen haben eine Vergrößerung des Suchraums zur Folge. Das Auswerten der stochastischen Zeitreihen einer jeden möglichen Kombination von Schaltzeitupel im Rahmen der vollständigen Enumeration erscheint im Zuge der Vergrößerung des Suchraums impraktikabel. Zudem kann den Ergebnissen des Szenario 1 in Abbildung 3.4 eine Wahrscheinlichkeit für lokale Minima im Suchraum entnommen werden, sodass im Zuge des Optimierungsproblems des Szenarios 2 die Partikelschwarmoptimierung (PSO) herangezogen wird [105]. Der Einsatz der Metaheuristik soll das Risiko in einem lokalen Minimum zu verbleiben reduzieren und die Wahrscheinlichkeit der Auffindung des optimalen Schaltzeitupels $(t_{S1}, t_{S2})_{opt, Zeitzbereich}$ unter Berücksichtigung eines akzeptablen numerischen Aufwands erhöhen.

Die beschriebene Optimierungsaufgabe des Szenarios 2 spannt einen zweidimensionalen Suchraum auf, welcher bspw. für den *Zeitbereich = Morgen* in $\mathcal{T}_{\text{Morgen}} = [360, 615) \times [360, 615)$ resultiert. Mit Hilfe der PSO wird der kontinuierliche Suchraum in Iteration n mit m Partikeln $\mathbf{x}_i(n) = (x_{i1}(n), x_{i2}(n))^T$ für $i = 1, 2, \dots, m$ durchsucht. Nach einer zufälligen initialen Positionierung der Partikel in Iteration $n = 0$ werden die Zielfunktionen evaluiert und minimiert. Basierend auf den Ergebnissen erfolgt eine Aktualisierung der Position \mathbf{x}_i in Iterationsschritt $n + 1$ mit:

$$x_{ij}(n+1) = x_{ij}(n) + v_{ij}(n+1) \quad (3.7)$$

wobei für $v_{ij}(n+1)$ gilt

$$\begin{aligned} v_{ij}(n+1) = & wv_{ij}(n) + \\ & + r_1c_1(p_{ij}(n) - x_{ij}(n)) \\ & + r_2c_2(g_{ij}(n) - x_{ij}(n)) . \end{aligned} \quad (3.8)$$

Die Summation, beschrieben durch Gleichung 3.8, wird im Kontext der PSO als *Geschwindigkeit* interpretiert, mit dem Trägheitskoeffizienten w sowie der Geschwindigkeit $v_{ij}(n)$ im Iterationsschritt n . Der Bestwert eines jeden einzelnen Partikels wird durch $p_{ij}(n)$ in Richtung der Komponente j repräsentiert. Der Bestwert aller Partikel ist durch $g_{ij}(n)$ beschrieben. Sowohl die Differenz $p_{ij}(n)$ zur aktuellen Position des Partikels x_{ij} als auch $g_{ij}(n)$ werden mit Hilfe eines konstanten Beschleunigungskoeffizienten c_1 respektive c_2 skaliert. Zusammen mit den Zufallszahlen $r_1, r_2 \sim \mathcal{U}(0, 1)$ einer stetigen Gleichverteilung \mathcal{U} folgend wird sichergestellt, dass die einzelnen Partikel den Suchraum auf einem stochastischen Pfad durchlaufen [105].

Bezogen auf das Untersuchungsszenario korrespondiert die Position des Partikels, beschrieben durch die Komponente \mathbf{x}_i , mit der zeitlichen Positionierung der Schalthandlungen $t_{\text{SH}_1} = x_{i1}(n)$ sowie $t_{\text{SH}_2} = x_{i2}(n)$ und kann mit Hilfe des Schaltzeitupels beschrieben werden. Für die Bestimmung der optimalen Schaltzeitupel $(t_{\text{SH}_1}, t_{\text{SH}_2})_{\text{opt, Zeitbereich}}$ eines einzelnen *Zeitbereichs* wird die Netzsimulation hinsichtlich der einzelnen Zielfunktionen separat evaluiert.

In der schematischen Darstellung der Abbildung 3.6 sind die notwendigen Simulationsschritte illustriert. Im Zuge der Initialisierung werden das Netzabbild aus der Datenbank geladen und die Simulationsparameter gesetzt. Wie oben beschrieben werden im initialen Iterationsschritt $n = 0$ in Summe m Partikel zufällig im Suchraum positioniert. Jeder einzelne Partikel $i = 1, 2, \dots, m$ repräsentiert dabei einen Schaltzeitupel $(t_{\text{SH}_1}, t_{\text{SH}_2})_{\text{Zeitbereich}}$. Zur Bestimmung der Geschwindigkeit nach Gleichung 3.8 eines einzelnen Partikels wird für alle m Partikel die in Abschnitt 3.1.1 beschriebene Netzsimulation durchgeführt und hinsichtlich der jeweiligen Zielfunktion evaluiert. Anschließend werden die Ergebnisse verglichen und der globale Bestwert $g_{ij}(n)$ identifiziert. Hierauf aufbauend wird für alle m Partikel eine neue Position mit Hilfe des Schaltzeitupels der Gleichung 3.7 folgend bestimmt und

der Iterationsschritt mit $n+1$ erhöht. Die Umpositionierung der Partikel und die damit einhergehenden Iterationsschritte werden fortgeführt, solange neue globale Bestwerte $g_{ij}(n)$ gefunden werden oder bis die max. Anzahl an Iterationsschritten bzw. ein Konvergenzkriterium erreicht wurde.

Zur weiteren Effizienzsteigerung, mit dem primären Ziel der Reduzierung der Rechenzeit, erfolgt die Netzsimulation und die Evaluierung der Zielfunktion eines jeden Partikels $i = \{1, 2, \dots, m\}$ verteilt im Netzwerk. Die Verwaltung des Partikelschwarms und Zuweisung der einzelnen Berechnungsaufgaben im Rechnernetz erfolgt zentral von einem Rechner im Netzwerk, dem sog. *Kopfknoten* (aus dem engl. head node), unter Zuhilfenahme von [106]. Der *Kopfknoten* wertet die statistisch evaluierten Ergebnisse des Schwarms aus und berechnet die neue Position $x_{ij}(n+1)$ eines jeden Partikels.

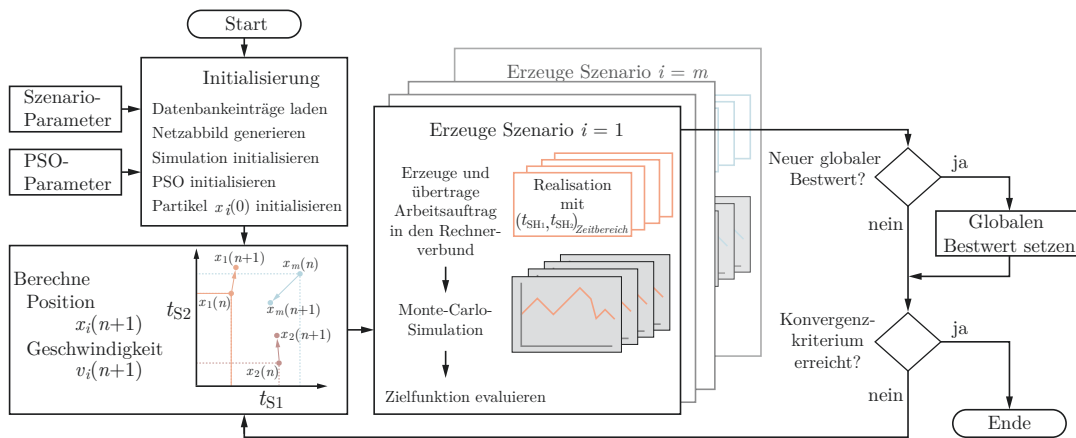


Abbildung 3.6.: Schematische Darstellung zur Lösung des Optimierungsproblems des Szenario 2 mit Hilfe der PSO.

Darstellung der Ergebnisse

Da das Szenario 2 über eine Vielzahl an PVA verfügt, wurde vorab eine Analyse durchgeführt um sicherzustellen, dass keine der Schalthandlungen zu einem invaliden Netzstatus führt. Hierbei stehen insbesondere Schalthandlungen und die daraus resultierende Spannungsanpassung durch den rONT von $f_{\Delta U_{SH}} = 1.05$ für $t_{SH_1} > t_{SH_2}$ (Spannungserhöhung) im Vordergrund. Diese können im beschriebenen Szenario zu einer Verletzung des Spannungsbandes an einzelnen Netzknoten führen. Daher wurde vorab eine Monte-Carlo-Simulation mit einer konstanten Spannungserhöhung durchgeführt. In Abbildung 3.6 (a) ist die statistische Auswertung der Voruntersuchung in Form eines Histogramms dargestellt. Der Abbildung ist die Häufigkeit der im Tagesverlauf aufgetretenen Spannungen dargestellt. Hinzuzufügen ist, dass bei der Auswertung nur Netzknoten berücksichtigt wurden, welche über eine PVA verfügen. Dem Histogramm ist zu entnehmen, dass es im Rahmen des festgelegten Szenarios zu keinem Zeitpunkt zu einer Verletzung der Spannungsqualitätsmerkmale nach DIN EN 50160 kommt. Die max. beobachtete Spannung liegt innerhalb des Grenzwerts von +10%. Dieses spannungsstabile Verhalten ist durch Modernisierungsarbeiten im Netz begründet, welche in der jüngsten Vergangenheit im Netzgebiet durchgeführt

wurden. Dabei wurde das gesamte Netz bei der Übernahme des Konzessionsgebietes verstärkt. Eine prozentuale Verteilung der im Netz zum Einsatz kommenden Kabeltypen ist in der Abbildung 3.6 (b) veranschaulicht. Der Kabelverteilung ist die Dominanz des Kabeltypen NAYY-J $4 \times 150 \text{ mm}^2$ zu entnehmen, welche teils im Netzgebiet als Doppelleitung (parallel) ausgeführt wurde. Dies erscheint jedoch bei einer installierten generatorischen Gesamtleistung $P_{\text{PVA}} = 398 \text{ kW}_{\text{peak}}$ im Netzgebiet als zwingend erforderlich. Insbesondere unter Berücksichtigung der Häufigkeitsverteilung und deren maximalen Werte nach Abbildung 3.6 (a).

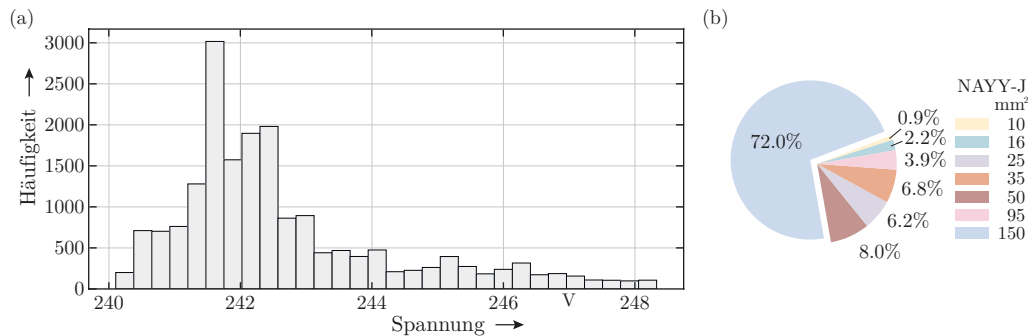


Abbildung 3.6.: (a) Darstellung der Spannungsverteilung im Tagesverlauf an Netzknotten mit PVA nach einer Spannungsanpassung von +5% des rONT. (b) Prozentuale Verteilung der Kabeltypen des Niederspannungsnetzes (eigene Darstellung nach [107]).

Auf Basis der evaluierten Zielfunktionen aller Partikel des Untersuchungsszenarios 2 erfolgt die Visualisierung der Ergebnisse wiederum mittels Konturdiagramm. Der Abbildung 3.7 sowie der Tabelle 3.2 ist die Auswertung zu entnehmen. Im Vergleich zur Darstellung der Ergebnisse des Szenarios 1 in Abbildung 3.4 zeigt die Auswertung des Szenarios 2 den CVR_f über alle *Zeitbereiche* und Zielfunktionen. Darüber hinaus sind die möglichen zeitlichen Positionierungen der Schalthandlungen t_{SH_1} und t_{SH_2} identisch, was durch die oben definierte Konvention begründet ist. Demnach sind die Werte der einzelnen Diagramme der Abbildung 3.7 durch eine imaginäre Winkeldiagonale getrennt und wie folgt zu interpretieren: Werte unter der Winkeldiagonale ($t_{\text{SH}_1} > t_{\text{SH}_2}$) entsprechen einer Spannungserhöhung mit $f_{\Delta U_{\text{SH}}} = 1.05$ zum Zeitpunkt t_{SH_2} und einer zeitlich folgenden Spannungsabsenkung zum Zeitpunkt t_{SH_1} auf Nennspannung U_0 und umgekehrt für Werte oberhalb der Winkeldiagonale. Für Werte entlang der Winkeldiagonale ($t_{\text{SH}_1} = t_{\text{SH}_2}$) gilt ein $\text{CVR}_f = 0$. Diese können als Referenzwerte ohne jegliche Spannungsanpassung interpretiert werden. Leichte Abweichungen hiervon sind durch die Interpolation zwischen den Werten zur Erzeugung des Konturdiagramms begründet. Zur Verdeutlichung ist der beschriebene Sachverhalt in der Abbildung 3.7 für den *Zeitbereich = Morgen* und der Zielfunktion P_{max} visuell gekennzeichnet.

Anhand der Darstellung der Ergebnisse können im Rahmen des Szenarios 2 folgende Kernaussagen getroffen werden:

Tabelle 3.2.: Optimale Schaltzeitupel $(t_{SH_1}, t_{SH_2})_{opt, Zeitbereich}$ je *Zeitbereich* $\in \{Morgen, Mittag, Abend\}$ der einzelnen Zielfunktionen des Szenarios 2.

Zielfunktion	<i>Morgen</i>					<i>Mittag</i>					<i>Abend</i>				
	$(t_{SH_1}, t_{SH_2})_{opt}$ min	ΔP_{max} %	ΔE_{max} %	$\Delta E_{Verluste}$ %	CVR_f	$(t_{SH_1}, t_{SH_2})_{opt}$ min	ΔP_{max} %	ΔE_{max} %	$\Delta E_{Verluste}$ %	CVR_f	$(t_{SH_1}, t_{SH_2})_{opt}$ min	ΔP_{max} %	ΔE_{max} %	$\Delta E_{Verluste}$ %	CVR_f
P_{max}	(389, 615)	-4.45	-0.28	0.55	0.89	(964, 788)	-0.9	-0.07	-1.23	0.18	(1178, 1400)	-2.25	-0.13	0.91	0.45
E_{max}	(360, 559)	-3.55	-0.45	0.48	0.09	(980, 615)	-0.68	-0.25	-2.35	0.05	(1056, 1400)	-1.32	-0.4	1.1	0.08
$E_{Verluste}$	(615, 360)	4.30	0.75	-0.05	0.01	(980, 615)	-0.68	-0.25	-2.35	0.47	(1400, 980)	3.55	0.65	-0.6	0.13

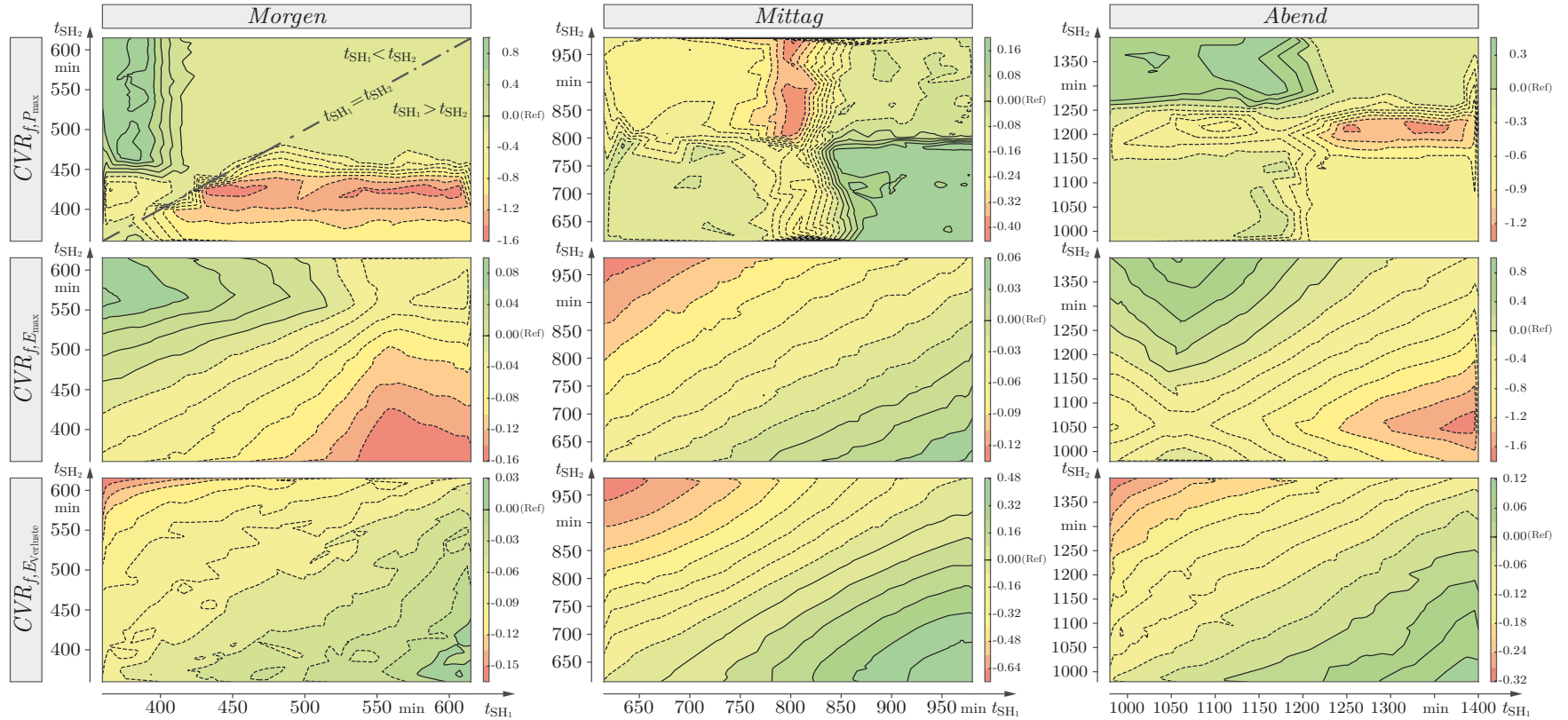


Abbildung 3.7.: Konturdiagramm zur Visualisierung des Einflusses der Schaltzeitupel $(t_{SH_1}, t_{SH_2})_{Zeitbereich}$ je *Zeitbereich* des Szenarios 2 auf die einzelnen Zielfunktionen. Der jeweilige Referenzwert (Ref) markiert den Wert, welcher ohne aktiver Spannungsanpassung aufgetreten wäre.

- Je Zielfunktion kann im jeweiligen *Zeitbereich* ein optimales Schaltzeittupel identifiziert werden.
- In den Morgen- und Abendstunden kann ein Schaltzeittupel zeitgleich netzdienlich auf P_{\max} und E_{\max} wirken.
- Eine Reduktion der Netzverluste in den Morgen- und Abendstunden hat eine Erhöhung der Spitzenlast P_{\max} sowie der Energieaufnahme E_{\max} des Netzes zur Folge.
- Während der Mittagszeit können Schaltzeittupel zeitgleich positiv auf alle drei Zielfunktionen wirken.

Die zweite und dritte Aussage sind auf den DCR zurückzuführen. Wie in Abbildung 3.1 dargestellt ist, kann der DCR zu einer Erhöhung der resultierenden Stromstärke führen, welches letztlich eine Erhöhung der Verluste im Netz zur Folge hat. Dieses Verhalten gleicht den Ergebnissen des Szenarios 1. Dies ist dahingehend zu begründen, dass die PVA erst über den *Zeitbereich Mittag* einen nennenswerten Einfluss auf den Leistungsfluss im Netz nimmt. Demzufolge ist die letzte Kernaussage das Resultat der hohen Einspeiseleistung der im Netz installierten PVA. Abbildung 3.7 können zudem Indikatoren entnommen werden, welche auf ein aktives Netz schließen lassen. Im *Zeitbereich Morgen* und *Abend* ist der Zielfunktion zur Reduktion der Energieaufnahme E_{\max} nach Gleichung 3.5 der Nulldurchgang der Wirkleistung bei Schaltzeittupel $(560, 560)_{\text{Morgen}}$ respektive $(1050, 1050)_{\text{Abend}}$ zu entnehmen, welcher den Wechsel von passiven zum aktiven Netz am *Morgen* kennzeichnet und umgekehrt am *Abend*. Der zweite Indikator kann den Zielfunktionen zur Spitzenlastreduktion P_{\max} nach Gleichung 3.4 und zur Reduktion der Gesamtverlustenergie E_{\max} nach Gleichung 3.5 im *Zeitbereich Mittag* entnommen werden. Im Vergleich zu den weiteren Diagrammen der Abbildung 3.7 sind diese annähernd an der Winkeldiagonalen gespiegelt.

Die optimalen Schaltzeittupel je *Zeitbereich* und Zielfunktionen können der Tabelle 3.2 entnommen werden. Optisch hervorgehoben sind die prozentualen Änderungen gegenüber des Referenzwertes. Im Vergleich über alle *Zeitbereiche* konnte für die Spitzenlastreduktion ein $CVR_{f,P_{\max}} = 0.89$ im *Zeitbereich = Morgen* im Rahmen des Szenarios 2 erreicht werden. Für das Schaltzeittupel $(389, 615)_{\text{Morgen}}$ resultiert dies im Zuge der Spannungsabsenkung ($t_{\text{SH}_1} < t_{\text{SH}_2}$) in einer Reduzierung der Spitzenlast P_{\max} um $\Delta P_{\max} = 4.45\%$. Im Vergleich zu den Ergebnissen des Szenarios 1 nach Tabelle 3.1 kann mit Hilfe optimaler Schaltzeitpunkte eine Reduktion der Gesamtverlustenergie von $CVR_{f,E_{\text{Verluste}}} = 0.47$ am *Mittag* erzielt werden. Dies stellt eine Reduzierung der Verlustenergie im Niederspannungsnetz von $\Delta E_{\text{Verluste}} = 2.35\%$ dar. Diese wurde durch eine Spannungserhöhung ($t_{\text{SH}_1} > t_{\text{SH}_2}$) mit Hilfe des Schaltzeittupels $(980, 615)_{\text{Mittag}}$ erzielt. Ergänzend ist hervorzuheben, dass das Schaltzeittupel $(980, 615)_{\text{Mittag}}$ positiv auf die Gesamtenergieaufnahme und -verlustenergie wirkt. Die Reduktion der Gesamtenergieaufnahme entspricht weitestgehend den Ergebnissen des Szenarios 1. Der dominante Einfluss der PVA, welche mit einer *leistungskonstanten* Spannungsabhängigkeit modelliert sind, sowie die Berücksichtigung des *energiekonstanten* Verbraucherlastverhaltens begründen die geringeren $CVR_{f,E_{\max}}$ im jeweiligen *Zeitbereich*.

3.3. Resümee

Die Ergebnisse der Szenarien 1 und 2 verdeutlichen, dass Betriebsführungsstrategien, welche im Kontext eines intelligenten Netzbetriebs spannungsabhängige Verbraucherleistungen als Handlungsoption berücksichtigen, einen Beitrag zur Flexibilisierung des Energiesystems auf unterer Spannungsebene leisten können. Eine gezielt herbeigeführte Spannungsanpassung bietet das Potenzial zur Spitzenlastreduktion, Energieeinsparung sowie Senkung der Netzverluste. Aufgrund der steigenden Komplexität in der Netzführung und -planung, primär hervorgerufen durch die Volatilität DEA sind diese zusätzlichen Flexibilitäten als sinnvolle Ergänzung zu sehen.

Der Abschnitt 3.1.3 der vorliegenden Arbeit zeigt zudem, dass im direkten Vergleich die Berücksichtigung des DCR zu abweichenden Ergebnissen im Kontext von spannungsoptimierenden Maßnahmen auf der Niederspannungsebene führt. Die Abbildung 3.1 des Abschnitts 3.1.3 veranschaulicht diesen Sachverhalt. Der Abbildung ist zu entnehmen, dass die Verwendung des ZIP(E)-Modells signifikante Unterschiede zum ZIP-Modell in den Lastverläufen aufzeigt. Darüber hinaus kompensiert die Berücksichtigung des ZIP(E)-Modells und des zugehörigen DCR die Auswirkung einer aktiven Spannungsänderung und fügt eine zusätzliche Zeitabhängigkeit hinzu.

Unabhängig von dem kompensierenden Einfluss des DCR konnten im Rahmen der Szenarien im Kontext der definierten Zielfunktionen optimale Schaltzeitpunkte eines rONT identifiziert werden. Auf Grundlage der vorgestellten Ergebnisse der Untersuchungsszenarien können folgende Kernaussagen getroffen werden:

- Aufgrund des kompensierenden Effektes, hervorgerufen durch den DCR, ist die Spitzenlastreduktion an der ONS stark zeitabhängig.
- Schaltzeitpunkte mit dem Ziel der Spitzenlastreduktion sollten die Spitzenlast zeitlich umschließen und demnach nur eine temporäre Spannungsanpassung darstellen.
- Schalthandlungen mit weit auseinanderliegenden Schaltzeitpunkten sollten für Maßnahmen zur Reduzierung der Energieaufnahme und Netzverluste vorgesehen werden.
- Ein optimales Schaltzeitpunkte über alle Zielfunktionen ist für ein *passives* Niederspannungsnetz *nicht* zu erzielen.
- Eine gleichzeitige Reduktion der Spitzenlast, Energieaufnahme und Netzverluste ist in *aktiven* Niederspannungsnetzen *möglich*.
- Die Untersuchungsszenarien zeigen, dass optimale Schaltzeitpunkte passiver und aktiver Niederspannungsnetze unterschiedlich sind.
- Eine Spannungserhöhung in Zeiten hoher Einspeiseleistung eines Niederspannungsnetzes kann zu geringeren Netzverlusten führen.

Die Ergebnisse bestätigen einen positiven Einfluss der PVA im Kontext von spannungsoptimierenden Maßnahmen. Dies kann insbesondere dem Konturplot nach Abbildung 3.7 im *Zeitbereich = Mittag* entnommen werden. Der Vergleich beider Szenarien zeigt, dass eine

zeitgleiche Reduktion aller Zielfunktionen nur in aktiven Niederspannungsnetzen möglich ist. Vor diesem Hintergrund sollten Betriebsführungsstrategien und Entscheidungsprozesse an die neuen Gegebenheiten im Zuge des Umbaus des Energieversorgungsystems angepasst und spannungsabhängige Verbraucherleistungen explizit für den intelligenten Netzbetrieb im Zuge der integrierten Energiewende mit einbezogen werden. Die dynamische Optimierung des Schaltzustandes einzelner ONS oder Umspannanlage im Kontext neuer Betriebsführungsstrategien bietet hierbei das Potenzial einen Beitrag zur Flexibilisierung des Energiesystems zu leisten.

Abschließend ist festzuhalten, dass die Untersuchung des Kapitels 3 untermauern die Notwendigkeit einer expliziten Berücksichtigung des energiekonstanten Verbraucherverhaltens zur Ableitung einer Netzführungsstrategie. Demnach wird ein energiekonstantes Verhalten einzelner elektr. Gerätschaften innerhalb der Simulationsumgebung zur Ableitung einer Weitbereichsregelung im Niederspannungsnetz im Rahmen der vorliegenden Arbeit einbezogen.

Bestärkendes Lernen für die integrierte Energiewende

Methoden der KI, welche als Teilgebiet der Informatik, Statistik und Robotik verstanden werden [108], erleben derzeit eine Renaissance. So wird bspw. das Jahr 2017 als „Jahr der KI“ bezeichnet [109]. Laut [110] ist die weltweite Anzahl an jährlichen Veröffentlichungen im Kontext der KI im Zeitraum von zehn Jahren (von 2007 auf 2017) von 30.000 auf 60.000 angestiegen. Die Begrifflichkeiten KI, ML und tiefes Lernen (aus dem engl. deep learning) werden in der gegenwärtigen Zeit im Kontext intelligenter Maschinen genannt, teilweise sogar als Synonyme verwendet [109]. Dabei sind die Begrifflichkeiten als spezielle Ansätze sowie Sammelbegriffe zu verstehen. Die einzelnen Zusammenhänge sind in Abbildung 4.1 illustriert.

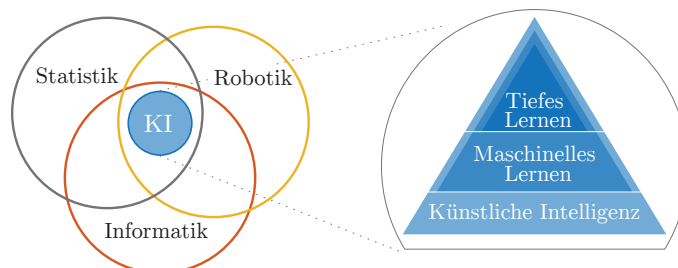


Abbildung 4.1.: Links: Verortung der künstlichen Intelligenz innerhalb verschiedener akademischer Disziplinen. Rechts: Zusammenhang der Begriffe künstliche Intelligenz, maschinelles und tiefes Lernen [108]. Der Pyramide ist zu entnehmen, dass das tiefe Lernen ein Teilbereich des maschinellen Lernens ist, was wiederum ein Teilgebiet der künstlichen Intelligenz ist.

Der Wunsch nach einer Maschine bzw. einem Algorithmus, der Entscheidungsstrukturen eines Menschen nachbildet, findet seine Ursprünge 1950 in den USA. In diesem Kontext wird häufig ein zwei monatiger Arbeitskreis am Dartmouth College in New Hampshire (USA) genannt, bei welchem der Begriff KI für einen solchen Ansatz erstmalig genannt wurde. Mc Carthy et al. benutzten KI in ihrem Projektantrag um Gelder für den Arbeitskreis zu akquirieren [111]. Dieser Arbeitskreis mit 10 Teilnehmern gilt bis heute als Ursprung dieser Forschungsdisziplin. Im Jahr 1968 definiert Marvin Minsky darüber hinaus den Begriff KI mit Hilfe des folgenden Satzes: „*the science of making machines do things that would require intelligence if done by men*“ [112] (engl.; frei übersetzt: die Wissenschaft der Herstellung von Maschinen, die Dinge bewerkstelligen, die Intelligenz erfordert, wenn diese durch den Menschen ausgeführt würden). Demnach beschreibt KI ein System, welches menschliche

Fähigkeiten nachahmt, speziell das strategische Denken.

Der Abbildung 4.1 ist zu entnehmen, dass das ML ein Teilgebiet der KI ist. Im Vergleich zu klassischen Ansätzen beschreibt das ML einen Ansatz, Muster aus Daten zu extrahieren, mit dem Ziel, verwandte Daten zu identifizieren und zukünftige Ergebnisse, Verhaltensweisen und Trends vorherzusagen. Hierbei sind dem Programmierparadigma wesentliche Unterschiede zu entnehmen, wie in Abbildung 4.2 visualisiert ist. Während in klassischen Ansätzen, dargestellt in Abbildung 4.2 (a), die Regeln durch den Menschen, basierend auf seiner Erfahrung, formuliert werden, sehen die datengetriebenen Ansätze des ML vor, dass die Maschine diese auf Grundlage von Daten und den zugehörigen Antworten ableitet. Zudem sehen die klassischen Ansätze i.d.R. eine Beschreibung des zugrunde liegenden (physikalischen) Modells vor, welches die Regeln der zu lösenden Problemstellung beschreibt. Dieses ist bei modellfreien Ansätzen des ML nicht erforderlich. Letztlich wird das zugrunde liegende Modell durch die Verarbeitung der Daten und Antworten innerhalb der Lernphase erarbeitet (vgl. Abbildung 4.2 (b)). Unabhängig des jeweiligen Ansatzes kann die erarbeitete Logik anschließend dahingehend genutzt werden, Antworten auf etwaige Systemzustände (Daten) mit Hilfe des ausführenden (Computer-) Programms zu bekommen. Ergänzend ist hinzuzufügen, dass auch die Erarbeitung der Logik mit Hilfe der klassischen Ansätze mehrere Iterationsschleifen beinhalten kann und in Abhängigkeit der Erfahrung des einzelnen Individuums zu sehen ist.

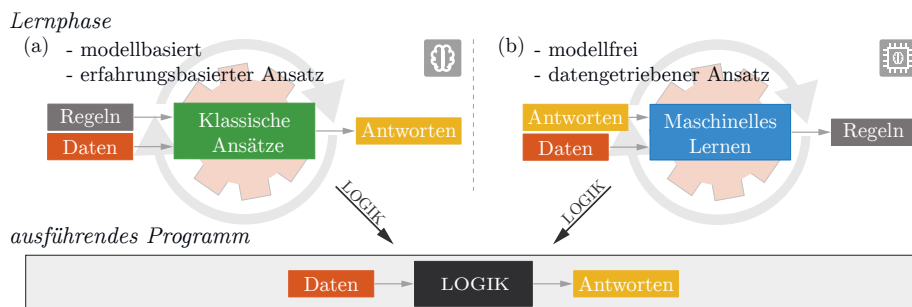


Abbildung 4.2.: Schematische Darstellung unterschiedlicher Programmierparadigmen. Der Abbildung ist ein Vergleich klassischer Ansätze (a) zu Methoden des ML (b) zu entnehmen. Während in klassischen Ansätzen die Regeln durch den Menschen formuliert werden, leitet die Maschine auf Grundlage von Daten und den zugehörigen Antworten diese bei Methoden des ML ab. Die innerhalb der Lernphase abgeleitete Logik, unabhängig ob nach Ansatz (a) oder (b), kann anschließend dahingehend genutzt werden, Antworten auf etwaige Systemzustände zu bekommen (vgl. ausführendes Programm).

Über die vergangenen Jahrzehnte wurden verschiedenste Algorithmen und Verfahren erarbeitet, welche mit Hilfe des Sammelbegriffs KI gruppiert werden. Dabei sorgen insbesondere die Fortschritte im Bereich des ML für die jüngsten Erfolge der KI. Speziell sind hiermit die Erfolge des 2010 gegründeten Unternehmens DeepMind gemeint, welches 2014 Teil von Google LLC wurde (damals noch Google Inc). DeepMind veröffentlichte u.a. im Jahr 2016 einen Algorithmus mit dem Namen *AlphaGo* [113]. Der Ansatz sorgte für ein großes mediales Echo, da es den Wissenschaftlern von DeepMind erstmalig gelungen ist, dass ein Algorithmus einen professionellen Spieler im Brettspiel Go besiegt. Auch wenn dieses Ereignis nur einen weiteren Meilenstein im Forschungsgebiet der KI und ML darstellt,

werden verschiedenste Ansätze in einem breiten Spektrum von Anwendungen im Bereich der Forschung und Entwicklung erprobt und teils im produktiven Einsatz getestet. Eine Übersicht der unterschiedlichen Bereiche im globalen Kontext kann den Veröffentlichungen nach [110] und [114] entnommen werden. Eine Bewertung des Potenzials mit Bezug auf die Energiewirtschaft ist den Studien nach [108] und [115] zu entnehmen. Zudem ist der Publikation nach [28] eine Gegenüberstellung von wiss. Untersuchungen im Kontext der Energieversorgung im Zeitraum von 2010 bis 2016 zu entnehmen. Diese werden unter Berücksichtigung der Reaktionszeiten, des Betriebszustands, der Problemstellung sowie anhand des verwendeten Lösungsansatzes klassifiziert und zusammengefasst. Ergänzend zu dieser Übersicht wurden in der jüngsten Vergangenheit Publikationen vorgestellt, welche zudem tiefe künstliche neuronale Netze (aus dem engl. deep neural networks; DNN) für Regelungsaufgaben und Entscheidungsprozesse im Energieversorgungssystem nutzen. So adressieren bspw. die Autoren der Publikation nach [45] die Problemstellung der Weitbereichsregelung im Niederspannungsnetz mit Hilfe des SARSAmix-Verfahrens. Als Grundlage der Untersuchung dienen reale Messdaten eines vollständig gemessenen Jahres mit einer Auflösung von 5 Minuten. Das in Brasilien bei São Paulo gelegene Niederspannungsnetz verfügt über einen rONT mit 17 Stufen (+5 % bis -15 % mit 1,25 % je Stufe des Stufenschalters) sowie einer hohen Durchdringung an PVA. Der in der Publikation vorgestellte Lösungsansatz reduziert die Spannungsschwankungen im Netz erfolgreich auf $< \pm 10\%$. Jedoch bemängeln die Autoren Restriktionen des genutzten tabellarischen Ansatzes und verweisen auf eine bessere Skalierbarkeit unter Anwendung des DRL.

Die Autoren der Publikation nach [42] adressieren die aktive Spannungshaltung im Mittelspannungsnetz mit Hilfe eines DDPG. Der Aktionsraum umfasst zum einen den Stufenschalter des Transformators, zum anderen die aktuelle Wirk- und Blindleistung einzelner Generatoren. Toubeau et al. präsentieren einen Ansatz für eine mit Unsicherheiten behaftete zentrale Netzführung im Mittelspannungsnetz. Hierfür wird mit Hilfe des Rahmenwerks des RL eine Kontrollstrategie erarbeitet, welche auf Basis zusätzlicher Szenarien fehlerhafte Einspeise- und Lastprognosen sowie Knotenmodell- und Netzinformationen berücksichtigt. Toubeau et al. verweisen darauf, dass der präsentierte Ansatz und das (D)RL im Allgemeinen die numerische Komplexität auf den Vorprozess, die *Lernphase* des Agenten, verschiebt. Dies resultiert in einer schnellen Reaktionszeit innerhalb des *ausführenden Modules* (siehe Abbildung 1.1 des Kapitels 1). Zudem wird darauf hingewiesen, dass das Rahmenwerk des RL die Notwendigkeit einer analytischen Repräsentation des elektr. Systems umgeht. Toubeau et al. beschreiben eine Entkopplung von der Modellgüte und der Güte der Kontrollstrategie. Hierdurch ist ein flexibles, adaptives System gegeben, welches sich leicht an Veränderung anpasst.

Eine weitere Publikation, welche Fragestellungen im Kontext des elektr. Energieversorgungssystems adressieren, stellt [46] dar. Wiederum mit Hilfe eines DDPG nutzen die Autoren den Ansatz des DRL zum Steuern von PV-Wechselrichtern im elektr. Niederspannungsnetz. Bei dem vorgestellten Ansatz werden eine Vielzahl von PV-Wechselrichtern an

unterschiedlichen Netzübergabepunkten orchestriert und jeweils eine Anpassung der Wirk- und Blindleistung zur Auffindung einer optimalen Lösung im ganzheitlichen Kontext bestimmt. Dieser zentrale Ansatz wird mit einem lokalen Ansatz verglichen, bei welchem den einzelnen Wechselrichter nur lokale Informationen vorliegen, womit der Status quo abgebildet ist. In Summe resultiert der vorgestellte Ansatz in einer reduzierten Abregelung der Einspeiseleistung sowie einer Reduzierung der Netzverluste. Zudem wird insbesondere der Nutzen des kontinuierlichen Aktionsraums des DDPG in den Fokus gestellt. Dieser kommt insbesondere bei der Koordinierung einer Vielzahl von Wechselrichtern zum Tragen, welche eine Diskretisierung des Aktionsraums nicht praktikabel macht.

In den folgenden Abschnitten werden zunächst die Grundlagen des RL diskutiert sowie Algorithmen vorgestellt, welche den Fokus auf die Lösung komplexer Aufgaben auf Grundlage eines unverarbeiteten, hochdimensionalen Datenstrom legen. Dies stellt eine der Kernaufgaben des Anwendungsfeldes der KI dar [49]. Anschließend wird im Hinblick auf den Umbau des elektr. Energieversorgungssystems ein Algorithmus des RL identifiziert, welcher in der Lage ist, die eingangs beschriebenen Herausforderungen im Zuge der integrierten Energiewende handzuhaben. Im Fokus der Bewertung stehen eine gute Skalierbarkeit des Ansatzes und Adaption auf unterschiedliche Problemstellungen. Abschließend wird ein Untersuchungsszenario vorgestellt, welches aufbauend auf der Verarbeitung einer Vielzahl von Messungen an unterschiedlichen Netzknoten eine Weitbereichsregelung im Niederspannungsnetz beschreibt. Vor dem Hintergrund eines praxisnahen Ansatzes steht eine möglichst geringe Schalthäufigkeit des rONT im Zentrum der Untersuchung. Diese Prämisse wird mit einem im Rahmen der vorliegenden Arbeit erstmals entwickelten Verfahren adressiert, welches auf Basis von Beobachtungen innerhalb eines zurückliegenden Intervalls erfolgt. Die simulative Untersuchung soll das Potenzial des Verfahrens im Kontext des intelligenten Netzbetriebs aufzeigen.

4.1. Grundlagen des bestärkenden Lernens

Die Rahmenstruktur des RL ist grundlegend auf einen Ansatz ausgelegt, der es ermöglicht, durch Interaktion mit einem zu kontrollierenden System ein spezifisches Ziel zu erreichen. Die theoretische Grundlage zur Beschreibung des zu kontrollierenden Systems bietet hierbei der Markow-Entscheidungsprozess. Mit Hilfe von Systemzuständen, Übergangswahrscheinlichkeiten, Aktionen und der Bewertung dieser kann im Rahmen des Markow-Entscheidungsprozesses ein System ganzheitlich beschrieben werden [117].

Die Algorithmen, welche für eine Zustandsänderung des Systems Aktionen zum jeweiligen Zeitpunkt t ausführen sind als Regler respektive Entscheidungsträger zu interpretieren. Dieser verfolgt im Kontext des RL eine *Kontrollstrategie* π (aus dem engl. policy). Die Ansätze zur Erarbeitung der Kontrollstrategie können im Allgemeinen in zwei übergeordnete Gruppen zusammengefasst werden, in *tabellarische Verfahren* und *Näherungsverfahren*.

Nachfolgend soll der Kerngedanke des zielorientierten Lernens innerhalb der Rahmenstruktur des RL sowie die Idee hinter beiden übergeordneten Gruppen der Verfahren des Reglers skizziert werden. Insbesondere die Schnittstelle zwischen dem zu kontrollierenden System und dem Regler sowie das Ableiten einer optimalen Kontrollstrategie π_{opt} des Reglers werden dabei näher beleuchtet. Vorweg genommen sei an dieser Stelle, dass mit Bezug auf die verwendeten Bezeichnungen und Nomenklatur einzelner Komponenten des RL der allgemeinen Formulierung nach [116] gefolgt wird. So wird bspw. der Regler als *Agent* bezeichnet und das zu kontrollierende System als *Umgebung* (aus dem engl. environment). Dies kann abweichen zu Formulierungen in den Ingenieurwissenschaften, welche u.a. den Regler als Controller bezeichnen [116]. Ziel dieser Annahme ist den Vergleich mit Beiträgen der einschlägigen Literatur zu vereinfachen.

4.1.1. Konzept und Rahmenstruktur

Wie in Kapitel 1 beschrieben, wird bei den Methoden des ML zwischen überwachtem, unüberwachtem und bestärkendem Lernen unterschieden. Anhand der Nomenklatur könnte davon ausgegangen werden, dass überwachtes und unüberwachtes Lernen den Bereich des ML ganzheitlich beschreiben. Insbesondere die Tatsache, dass dahingehend unterschieden wird, ob die Algorithmen ein beschreibendes Merkmal berücksichtigen oder nicht, verstärkt diese Annahme. Wie aufgezählt wird beim ML jedoch zwischen drei Ansätzen unterschieden. Auch wenn beim RL die verarbeiteten Informationen nicht über ein beschreibendes Merkmal verfügen (analog zum unüberwachten Lernen), ist das Ziel und die besondere Struktur der Beschreibung der Problemstellung ausschlaggebend für diese dritte Kategorie des ML. Zudem liegt der Fokus des RL bei Problemstellungen, welche Entscheidungsprozesse im Zuge von Regelungsaufgaben adressieren. Der Ansatz des RL kommt dabei der natürlichen Intuition eines spielerischen Lernens des Menschen am nächsten. Konkret ist hier die Interaktion mit der Umgebung gemeint, welche als fundamentale Idee hinter den Lernverfahren dieser Kategorie steckt [116]. Durch ein spielerisches Ausprobieren und Interagieren können Muster und Abläufe identifiziert werden. Dieser Kerngedanke wird beim RL aufgegriffen und für ein rechnergestütztes Verfahren formuliert.

Das RL stellt im Allgemeinen ein mathematisches Modell dar, dessen theoretische Rahmenstruktur aus einem Agenten und einer Umgebung \mathcal{E} besteht, wie in Abbildung 4.3 dargestellt. Auf Grundlage eines Systemzustandes S_t sowie der Belohnung R_t zum Zeitpunkt t führt der Agent eine Aktion A_t aus. Dem folgt die Systemantwort der Umgebung \mathcal{E} zum Zeitpunkt $t + 1$ mit einem Systemzustand S_{t+1} sowie einer Belohnung R_{t+1} . Die Belohnung R_{t+1} stellt hierbei einen Indikator zur Bewertung der Güte der Aktion A_t im Kontext der zu lösenden Problemstellung dar. Dieser in Abbildung 4.3 skizzierte Kreislauf wird zyklisch wiederholt bis ggf. ein Abbruchkriterium erreicht ist.

Zusammengefasst wird mit dem RL die Projektion einer Situation oder eines Zustandes S_t auf bestimmte Aktionen A_t beschrieben um eine Belohnung R_t zu maximieren [116]. Dabei lernt der Agent durch ausprobieren, welche der ausgeführten Aktionen A_t in be-

stimmten Zuständen S_t positiv im Sinne der Belohnung R_t sind. Hierbei wird zwischen *episodischen* und *kontinuierlichen Aufgaben* unterschieden. *Episodische Aufgaben* haben im Unterscheid zu *kontinuierlichen Aufgaben* einen Systemzustand S_T zum Zeitpunkt T , welcher einen terminalen Systemzustand darstellt und die jeweilige Episode beendet. Im Kontext des RL sind dies Aufgaben, wie bspw. das Erlernen eines Brettspiels. Eine Partie Schach wäre hier als eine Episode zu interpretieren. Im Kontrast hierzu sind *kontinuierliche Aufgaben* ohne definierten Endzustand zu verstehen. Ein Agent, welcher erlernt, Kauf- und Verkaufssignale auf Positionen am Finanzmarkt als Aktion durchzuführen, interagiert mit einer Umgebung \mathcal{E} , welche eine *kontinuierliche Aufgabe* darstellt.

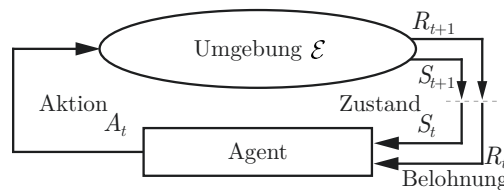


Abbildung 4.3.: Visualisierung des Konzeptes des RL. Die Interaktion zwischen Agent und Umgebung \mathcal{E} erfolgt mit Hilfe der Signale des Zustandes S_t , der Aktion A_t sowie einer skalaren Belohnung R_t zum jeweiligen Zeitpunkt t .

4.1.2. Signale und Zieldefinition

Unabhängig davon ob die Umgebung eine *episodische* oder *kontinuierliche Aufgabe* beschreibt, basiert das mathematische Modell zur Beschreibung der Problemstellung auf drei Signalen: Dem Systemzustand S , der Aktion A sowie der Belohnung R , welche jeweils eine ZV mit den Realisierungen s , a und r darstellt [116]. Die Signale sowie die Definition des übergeordneten Ziels des Rahmenwerks mit Hilfe des sog. Gewinns G , welcher eine gesonderte Stellung einnimmt, werden nachfolgend thematisiert.

4.1.2.1. Systemzustand

Der Systemzustand ist als aktueller Zustand S_t (aus dem engl. state) der Umgebung zum Zeitpunkt t zu interpretieren und dient als Entscheidungsgrundlage des Agenten. Formal gilt für die Zustände der Umgebung \mathcal{E} , dass die Gesamtheit aller Zustände $S \in \mathcal{S}$ den endlichen Zustandsraum \mathcal{S} der Umgebung definieren. Dabei sind die Zustände im Rahmen der jeweiligen Problemstellung als Kontext zu verstehen, welcher notwendig ist, damit der Agent eine Entscheidung hoher Güte mit Bezug auf die jeweilige Problemstellung treffen kann. So ist es üblich, dass der Zustandsraum nur eine Teilmenge der möglichen Informationen der Umgebung \mathcal{E} darstellt. Darüber hinaus ist festzuhalten, dass der Zustandsraum \mathcal{S} alle nicht terminalen Systemzustände beschreibt. Mit \mathcal{S}^+ werden bei *episodischen Aufgaben* alle Zustände angegeben, einschließlich der terminalen Zustände. Für *kontinuierliche Aufgaben* gilt, dass \mathcal{S} alle Zustände $S \in \mathcal{S}$ beinhaltet, da kein definiertes Ende gegeben ist [116].

4.1.2.2. Aktion und Aktionsraum

Die Aktion A_t (aus dem engl. action) beschreibt das Signal, welches der Agent zum interagieren mit der Umgebung \mathcal{E} getroffen hat, um den aktuellen Systemzustand zu beeinflussen. Die Wahl der Aktion $A_t = a$ basiert auf der im Moment gültigen Kontrollstrategie $\pi(A_t|S_t = s)$ des Agenten. Die Kontrollstrategie $\pi(A_t|S_t = s)$ beschreibt die bedingte Wahrscheinlichkeit, dass der Agent Aktion $A_t = a$ unter der Bedingung eines vorherrschenden Systemzustandes $S_t = s$ ausführt. Der Aktionsraum $\mathcal{A}(s)$ beschreibt alle möglichen Aktionen $A_t \in \mathcal{A}(s)$. Mit $\mathcal{A}(s)$ wird die Teilmenge der möglichen Aktionen im aktuellen Zustand $S_t = s$ angegeben. Sollte der Aktionsraum unabhängig des jeweiligen Systemzustandes S_t sein, so ist eine verkürzte Schreibweise mit $\mathcal{A}(s) = \mathcal{A}$ üblich [116].

4.1.2.3. Belohnungssignal

Das dritte Signal nach Abbildung 4.3 stellt den Belohnungsmechanismus des Lernprozesses dar. Die Belohnung $R_t = \mathcal{R}(S_t, A_t, S_{t+1})$ (aus dem engl. reward) stellt eine skalare Größe mit $R_t \in \mathbb{R}$ dar und wird mit Hilfe der Belohnungsfunktion $\mathcal{R} : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow \mathbb{R}$ beschrieben [118]. Das Signal R_t der Umgebung \mathcal{E} stellt die Grundlage der Bewertung der Aktion A_{t-1} im Zustand S_{t-1} des vergangenen Zeitpunktes $t - 1$ dar.

Im Allgemeinen erfolgt über die Formulierung der Belohnungsfunktion \mathcal{R} die Zieldefinition der Problemstellung. Demnach spielt die Belohnungsfunktion eine zentrale Rolle im Rahmenwerk des RL [46]. Zum einen muss das Belohnungssignal das definierte Ziel adressieren, zum anderen ist eine Konvergenz in einem adäquaten Zeitraum sicherzustellen [119].

4.1.2.4. Zieldefinition des Rahmenwerks

Basierend auf der Interaktion des Agenten mit der zugehörigen Umgebung \mathcal{E} ist das übergeordnete Ziel des RL durch die Ableitung einer optimalen Kontrollstrategie π_{opt} beschrieben, welche den Erwartungswert der kumulierten Belohnung

$$\mathbb{E}_{\pi_{\text{opt}}} \left[\sum_{t=1}^T R_t \right] \quad (4.1)$$

maximiert, mit Zeitpunkt T des terminalen Systemzustands. Der Gleichung 4.1 ist zu entnehmen, dass nicht die maximale Belohnung R_t in jedem Zeitpunkt t ausschlaggebend ist. Vielmehr ist die Maximierung der zu erwartenden kumulierten Belohnung, den sog. Gewinn G_t , das definierte Ziel. Dieser Zusammenhang wird im Kontext des Verfahrens als *Belohnungshypothese* (aus dem engl. reward hypothesis) bezeichnet und ist charakteristisch für das RL. Mit Hilfe des Gewinns G_t kann das Ziel des Agenten formal als Summe der Belohnung zum Zeitpunkt t durch

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \gamma^3 R_{t+4} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \quad (4.2)$$

beschrieben werden, mit einem Diskontierungsfaktor (aus dem engl. discount rate) $\gamma \in [0, 1]$. Der Diskontierungsfaktor γ der Gleichung 4.2 beschreibt eine Gewichtung der in der Zukunft zu erzielenden kumulierten Belohnung. Demnach ist eine Belohnung, welche in k Zeitschritten zu erwarten ist um den Faktor γ^{k-1} kleiner. Insbesondere bei Problemstellungen, welche eine *kontinuierliche Aufgabe* beschreiben, ist ein Diskontierungsfaktor γ obligatorisch. Ohne eine Gewichtung würde der Gewinn G_t nicht konvergieren ($G_t \rightarrow \infty$) [116]. Eine Bewertung einzelner Aktionen des Agenten im Sinne des RL wäre für eine *kontinuierliche Aufgabe* nicht möglich.

Zusammengefasst soll die Zieldefinition über den Gewinn G_t gewährleisten, dass der Agent in die Lage versetzt wird in einzelnen Zeitpunkten t Aktionen A_t von geringerer Güte im Sinne des Belohnungssignals R_t auszuführen, mit dem übergeordneten Ziel der Maximierung des Gewinns G_t .

4.1.3. Markow-Entscheidungsprozess

Dem Markow-Entscheidungsprozess zugrundeliegend ist der Markow-Prozess, auch bekannt unter der Markow-Kette [117]. Beim Markow-Prozess können im Allgemeinen einzelne Systemzustände eines bestehenden Systems beobachtet werden. Dabei kann das System über eine beliebige aber endliche Anzahl an Zuständen $S \in \mathcal{S}$ verfügen. Zudem gilt, dass ein Wechsel des Zustandes $S_t = s$ zum Zeitpunkt t in einen nachfolgenden Zustand $S_{t+1} = s'$ stattfindet, ohne dass Einfluss auf das System genommen wurde. Wird nun eine Sequenz an Zuständen

$$S_t, S_{t+1}, S_{t+2}, S_{t+3}, \dots \quad (4.3)$$

beobachtet und erfüllt diese Sequenz zudem die Markow-Eigenschaft, so wird von der sog. Markow-Kette gesprochen [117]. Die Markow-Eigenschaft besagt hierbei, dass die einzelnen Zustände in der Sequenz nach 4.3 jeweils nur von dem zuvor beobachteten Zustand und nicht von der gesamten Sequenz abhängig sind. Formal bedeutet dies, dass für die Übergangswahrscheinlichkeit $p_{s,s'}$, welche die Dynamik des Systems beschreibt,

$$\begin{aligned} p_{s,s'} &= \mathcal{P}(S_{t+1} = s' | S_t = s, S_{t-1} = s_{t-1}, \dots, S_1 = s_1, S_0 = s_0) \\ &= \mathcal{P}(S_{t+1} = s' | S_t = s) \end{aligned} \quad (4.4)$$

gelten muss [120]. Im Allgemeinen wird ein Markow-Prozess über die Untermenge an Zuständen $S \in \mathcal{S}$ des Zustandsraums \mathcal{S} sowie einer quadratischen $L \times L$ Übergangsmatrix \mathcal{P} beschrieben, wobei L die Anzahl der Zustände definiert [117].

Soll nun die Systembeschreibung von einem Markow-Prozess hin zu einem Markow-Entscheidungsprozess angepasst werden, müssen folgende zwei Ergänzungen Berücksichtigung finden. Zunächst wird der Wechsel eines Zustandes mit einem Wert belegt. Dieser Wert entspricht der Belohnung R_t nach Abschnitt 4.1.2. Formal würde eine solche Systembeschreibung als Markow-Belohnungsprozess (aus dem engl. Markov reward process) bezeichnet

werden [117]. Die zweite Ergänzung umfasst das Hinzufügen eines Aktionsraumes \mathcal{A} nach Abschnitt 4.1.2 und führt letztlich zum Markow-Entscheidungsprozess. Mit Hilfe einer Aktion $A_t = a$ kann abweichend zum Markow-Prozess Einfluss auf den Zustand $S_t = s$ des Markow-Entscheidungsprozesses genommen werden. Die Übergangswahrscheinlichkeit wird abschließend durch

$$p(s', r|s, a) = \mathcal{P}(S_{t+1} = s', R_{t+1} = r|S_t = s, A_t = a), \quad (4.5)$$

für alle $s', s \in \mathcal{S}, r \in \mathcal{R}$ sowie $a \in \mathcal{A}$, beschrieben [116].

Gleichung 4.5 beschreibt die Einzelschritt-Dynamik des Markow-Entscheidungsprozesses (aus dem engl. one-step dynamics). Der Gleichung folgend gilt für das Auftreten der Realisationen s' und r der ZV S_t respektive R_t eine Wahrscheinlichkeit p , unter der Bedingung, dass zum Zeitpunkt t der Systemzustand $S_t = s$ vorlag sowie der Agent die Aktion $A_t = a$ ausführte. Darüber hinaus gilt im Allgemeinen für die Wahrscheinlichkeitsdichte p

$$\sum_{s' \in \mathcal{S}} \sum_{r \in \mathcal{R}} p(s', r|s, a) = 1, \text{ für alle } s \in \mathcal{S}, a \in \mathcal{A}(s). \quad (4.6)$$

Der Markow-Entscheidungsprozess wird vollständig über das Tupel $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$ beschrieben [116] und resultiert, analog zur Sequenz nach 4.3, in

$$S_0, A_0, R_1, S_1, A_1, \dots, R_{t-1}, S_{t-1}, A_{t-1}, R_t, S_t, A_t, R_{t+1}, S_{t+1}, A_{t+1}, \dots \quad (4.7)$$

zur Beschreibung des Übergangs in einen Zustand zum jeweiligen Zeitpunkt $t \in \mathcal{T}$ mit $\mathcal{T} = \{1, 2, 3, \dots, T\}$.

4.1.4. Aufbau der Agenten

Mit Hilfe des im letzten Abschnitt beschriebenen Markow-Entscheidungsprozesses wird die Problemstellung im Kontext des RL über $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$ formell beschrieben. Der Zustandsraum \mathcal{S} , Aktionsraum \mathcal{A} sowie der *Diskontierungsfaktor* γ sind dem Agenten i.d.R. bekannt. Die Belohnung \mathcal{R} sowie die Übergangswahrscheinlichkeiten \mathcal{P} hingegen sind zunächst unbekannt und müssen durch Interaktion mit der Umgebung \mathcal{E} durch den Agenten erlernt werden. In dieser Interaktion des Agenten mit der Umgebung liegt die Kernüberlegung des RL. Hierfür muss der Agent zum einen in die Lage versetzt werden Zustände der Umgebung aufzunehmen und zum anderen Aktionen auszuführen, die eine Zustandsänderung in der Umgebung zur Folge haben.

Beim Interagieren mit der Umgebung \mathcal{E} wird zunächst ein „trial and error“-Ansatz (engl., frei übersetzt: Versuch und Irrtum) verfolgt, welcher das explorative *Erforschen* des Zustandsraumes ermöglicht. Dies ist darin begründet, dass zunächst keinerlei Informationen über die Problemstellung vorliegen. Mit Fortschreiten des Experimentes wird der Agent zunehmend in die Lage versetzt, erarbeitetes Wissen über die Umgebung \mathcal{E} bei der Wahl

der Aktionen zu nutzen. Durch Beobachten und Verarbeiten der Systemantwort in Folge einer Aktion wird auf Grundlage des Belohnungssignals R_t die Güte der jeweiligen Aktion mit Hilfe einer Wertfunktion bewertet und in einer resultierenden Kontrollstrategie π festgehalten. Die einzelne Kontrollstrategie π ist demnach als *ein* möglicher Weg durch die Problemstellung zu verstehen und stellt nicht zwingend die optimale Kontrollstrategie π_{opt} dar. Erst durch eine immer wiederkehrende Interaktion mit der Umgebung kann der Agent in die Lage versetzt werden, eine optimale Kontrollstrategie π_{opt} abzuleiten. Eine einzelne Episode des Experimentes ist hierbei als Stichprobe zu verstehen und spiegelt nicht das zugrunde liegende Modell wider. Erst mit zunehmender Anzahl an Stichproben kommt die Modellschätzung dem unterliegenden Modell näher [117].

In diesem Abschnitt liegt der Fokus auf dem Algorithmus, welcher den Agenten beschreibt und der damit einhergehenden Ableitung einer Kontrollstrategie π durch den Agenten. Im Rahmen der vorliegenden Arbeit wird das Augenmerk auf die Gruppe der modellfreien Ansätze gelegt. Die Ausgestaltung des Agenten stellt hierbei eine Struktur dar, welche es ermöglicht, dass der Agent durch kontinuierliche Interaktion mit der Umgebung \mathcal{E} eine optimale Kontrollstrategie π_{opt} ableitet, ohne dass weiterführende Informationen über die Problemstellung vorliegen [117]. In Abhängigkeit der Problemstellung und dessen Komplexität kann die Kontrollstrategie des Agenten im einfachsten Fall über ein *tabellarisches Verfahren* oder für komplexe Zusammenhänge mit Hilfe eines *Näherungsverfahrens* erfolgen. Hierbei ist der Aufbau des Agenten und die Wahl der Kontrollstrategie entscheidend für die effiziente Lösung der Problemstellung durch den Agenten und ausschlaggebend dafür, ob dieser eine optimale Kontrollstrategie π_{opt} findet. Hierzu wird zunächst auf die mathematischen Grundlagen zur Bewertung einer Kontrollstrategie π mit Hilfe von Wertfunktionen eingegangen. Im direkten Anschluss werden tabellarische sowie Näherungsverfahren vorgestellt und hinsichtlich eines gut skalierenden, generalisierten Ansatzes im Kontext einer integrierten Energiewende bewertet.

4.1.4.1. Optimale Kontrollstrategie

Die optimale Kontrollstrategie π_{opt} ist jeweils in Abhängigkeit der zu lösenden Problemstellung zu sehen. Eine definierte Kontrollstrategie π ist demzufolge nicht auf neue Problemstellungen zu adaptieren und muss je Problemstellung gesondert durch eine *Lernphase* erarbeitet werden. Während dieser Phase interagiert der Agent mit der Umgebung \mathcal{E} und erarbeitet eine Vielzahl an unterschiedlichen Kontrollstrategien π_i mit $i \in \mathbb{Z}^+$, welche in Abhängigkeit des jeweiligen Ansatzes am Ende einer Episode oder bereits nach jedem Zeitschritt evaluiert werden. Folglich ist eine Kontrollstrategie π des Agenten als Projektion eines jeden einzelnen Systemzustandes auf eine Aktion zu verstehen. Konkret ist hierdurch definiert, mit welcher Aktion A_t der Agent im jeweiligen Systemzustand S_t zu reagieren hat, um den Gewinn G_t zu maximieren. Formal werden die Kontrollstrategien in eine der folgenden zwei Gruppen an Algorithmen klassifiziert:

1. Deterministisch mit $\mu : \mathcal{S} \rightarrow \mathcal{A}$: Eine direkte Projektion des Zustandes $s \in \mathcal{S}$ auf eine Aktion $a \in \mathcal{A}(s)$.
2. Stochastisch mit $\pi : \mathcal{S} \times \mathcal{A} \rightarrow [0, 1]$: Die bedingte Wahrscheinlichkeit $\pi(a|s)$ eine Aktion a auszuführen während der Agent in Zustand s verweilt.

Hierbei gilt im Allgemeinen, dass eine deterministische Kontrollstrategie jederzeit durch eine stochastische Repräsentation beschrieben werden kann [116]. Demnach soll nachfolgend die allgemein gültige Schreibweise Anwendung finden.

Ein genereller Vergleich der Kontrollstrategien π erfolgt durch

$$\pi' \geq \pi \text{ für alle } \pi \quad (4.8)$$

und führt zur Definition einer höherwertigen Kontrollstrategie π' und letztlich zu einer optimalen Kontrollstrategie π_{opt} . Bei Ansätzen des modellfreien RL basiert die Kontrollstrategie π i.d.R. auf einer Bewertung der Systemzustände S_t mit Hilfe einer Wertfunktion, welcher der Gewinn G_t zugrunde liegt. Dieser Zusammenhang soll nachfolgend anhand der Wertfunktion des Zustandes $V_\pi(s)$ sowie des Zustand-Aktion-Tupels $Q_\pi(s, a)$ beleuchtet werden.

Wertfunktion des Zustandes

Mit Hilfe der Wertfunktion des Zustandes $V_\pi(s)$ wird jedem Zustand $S_t = s$ mit $S_t \in \mathcal{S}$ ein Wert zugeordnet, der dem erwarteten Gewinn $\mathbb{E}_\pi(G_t)$ entspricht. Es gilt für die Wertfunktion des Zustandes

$$V_\pi(s) = \mathbb{E}_\pi[G_t | S_t = s]. \quad (4.9)$$

Gleichung 4.9 beschreibt für jeden Zustand $S_t \in \mathcal{S}$ der Umgebung \mathcal{E} den Erwartungswert des nach Gleichung 4.2 diskontierten Gewinns G_t , unter der Bedingung, dass der Zustand der Umgebung $S_t = s$ vorliegt und der Agent der Kontrollstrategie π für alle nachfolgenden Schritte $t \in \mathcal{T}$ folgt. Zur vereinfachten numerischen Handhabung wird die Gleichung 4.9 unter Berücksichtigung des Optimalitätsprinzips nach Bellman [121] in eine rekursive Form umgeschrieben. Unter Berücksichtigung der *Belohnungshypothese* sowie der Gleichung 4.2 kann für generelle Markow-Entscheidungsprozesse die Gleichung 4.9 zu

$$V_\pi(s) = \mathbb{E}_\pi[R_{t+1} + \gamma V_\pi(S_{t+1}) | S_t = s] \quad (4.10)$$

umgeschrieben werden. Der Gleichung 4.10 ist zu entnehmen, dass die Wertfunktion des Zustandes $V_\pi(s)$ mit Hilfe der Wertfunktion des Zustandes $V_\pi(S_{t+1})$ des nachfolgenden Zeitpunktes $t+1$, unter Berücksichtigung des Diskontierungsfaktors $\gamma \in [0, 1]$, beschrieben wird. Für eine Aktion $A_t = a$ der Kontrollstrategie $\pi(a|s)$ im Zustand s folgend kann nach Bellman [121], unter Annahme der beschriebenen rekursiven Form der Wertfunktion des

Zustandes, der erwartete Gewinn $\mathbb{E}_\pi(G_t)$ durch die Summe

$$V_\pi(s) = \sum_{a \in \mathcal{A}(s)} \pi(a|s) \sum_{s' \in \mathcal{S}^+} \sum_{r \in \mathcal{R}} p(s', r|s, a) (r + \gamma V_\pi(s')) \quad (4.11)$$

beschrieben werden, mit $S_{t+1} = s'$ und $R_{t+1} = r$ [116].

Mit Bezug auf die Güte der Kontrollstrategie π kann mit Hilfe der Wertfunktion des Zustandes $V_\pi(s)$ eine im Kontext der Problemstellung *bessere* Kontrollstrategie π' durch

$$\pi' \geq \pi \text{ wenn } V_{\pi'}(s) \geq V_\pi(s) \text{ für alle } s \in \mathcal{S} \quad (4.12)$$

ermittelt werden [116]. Gleichung 4.12 folgend ist eine höherwertige Kontrollstrategie π' jene Strategie, welche die Wertfunktion $V_{\pi'}(s)$ eines jeden Zustandes s maximiert. Eine Bewertung über alle Kontrollstrategien π hinweg kann letztlich zu einer optimalen Kontrollstrategie π_{opt} respektive einer optimalen Wertfunktion des Zustandes V_{opt} führen.

Wertfunktion des Zustand-Aktion-Tupels

Die zweite Funktion im Kontext des modellfreien RL, welche neben der Wertfunktion des Zustandes $V_\pi(s)$ im Fokus der vorliegenden Arbeit liegt, ist die Wertfunktion des Zustand-Aktion-Tupels $Q_\pi(s, a)$. Die Wertfunktion

$$Q_\pi(s, a) = \mathbb{E}_\pi[G_t | S_t = s, A_t = a] \quad (4.13)$$

beschreibt für jedes Zustand-Aktion-Tupel (s, a) den zu erwarteten Gewinn $\mathbb{E}_\pi(G_t)$ unter der Bedingung, dass der Agent in Zustand $S_t = s$ die Aktion $A_t = a$ ausführt und anschließend der Kontrollstrategie π über alle nachfolgenden Zeitschritte $t \in \mathcal{T}$ folgt. Demnach ist die Wertfunktion des Zustand-Aktion-Tupels $Q_\pi(s, a)$ dahingehend von der Wertfunktion des Zustandes $V_\pi(s)$ zu unterscheiden, dass jeder möglichen Aktion A_t im jeweiligen Zustand S_t ein Wert zugeordnet wird. Für einen finiten Zustands- und Aktionsraum, \mathcal{S} respektive \mathcal{A} , stellt die Gleichung nach 4.13 eine Tabelle dar, allgemein als Q-Tabelle bekannt. Jeder einzelne Wert der Tabelle beschreibt die Wertfunktion $Q_\pi(s, a)$ eines Zustand-Aktion-Tupels (s, a) und entspricht dem erwarteten Gewinn $\mathbb{E}_\pi(G_t)$ nach Gleichung 4.13, unter der Voraussetzung, dass der Agent im Zustand $S_t = s$ die Aktion $A_t = a$ ausführt und anschließend der Kontrollstrategie π folgt. Exemplarisch soll in Tabelle 4.1 eine Q-Tabelle eines fiktiven Markow-Entscheidungsprozesses mit $\mathcal{S} = \{s_1, s_2, s_3, s_4\}$ und $\mathcal{A} = \{a_1, a_2, a_3, a_4\}$ dargestellt sein. Die angewandte Kontrollstrategie des Beispiels entspricht π .

Unter der Annahme, dass der Agent durch Interaktion mit der Umgebung \mathcal{E} eine optimale Wertfunktion des Zustand-Aktion-Tupels $Q_{\pi_i}(s, a) \rightarrow Q_{\text{opt}}$ für Episode $i \rightarrow \infty$ abgeleitet hat, kann die optimale Kontrollstrategie π_{opt} unmittelbar durch Anwenden der Zielfunktion

$$\pi_{\text{opt}}(s) = \arg \max_{a \in \mathcal{A}(s)} Q_{\text{opt}}(s, a) \text{ für alle } s \in \mathcal{S} \quad (4.14)$$

aus der Q-Tabelle bestimmt werden [116]. Die Gleichung 4.14 nimmt im Kontext des RL eine gesonderte Rolle ein und wird insbesondere bei der Beschreibung der einzelnen Verfahren und Algorithmen im Zusammenhang eines *gierigen* Agenten (aus dem engl. greedy) thematisiert. Ein Agent handelt demnach *gierig*, wenn er nach Gleichung 4.14 die Aktion $A_t = a$ wählt, welche die Wertfunktion des Zustand-Aktion-Tupels $Q_\pi(s, a)$ im Zustand $S_t = s$ maximiert.

Tabelle 4.1.: Exemplarische Darstellung einer Q-Tabelle eines Markow-Entscheidungsprozesses mit $\mathcal{S} = \{s_1, s_2, s_3, s_4\}$ und $\mathcal{A} = \{a_1, a_2, a_3, a_4\}$ sowie der Kontrollstrategie π .

	a_1	a_2	a_3	a_4
s_1	$Q_\pi(s_1, a_1)$	$Q_\pi(s_1, a_2)$	$Q_\pi(s_1, a_3)$	$Q_\pi(s_1, a_4)$
s_2	$Q_\pi(s_2, a_1)$	$Q_\pi(s_2, a_2)$	$Q_\pi(s_2, a_3)$	$Q_\pi(s_2, a_4)$
s_3	$Q_\pi(s_3, a_1)$	$Q_\pi(s_3, a_2)$	$Q_\pi(s_3, a_3)$	$Q_\pi(s_3, a_4)$
s_4	$Q_\pi(s_4, a_1)$	$Q_\pi(s_4, a_2)$	$Q_\pi(s_4, a_3)$	$Q_\pi(s_4, a_4)$

4.1.4.2. Verfahren

In der einschlägigen Literatur sind eine Vielzahl an unterschiedlichen Algorithmen zu finden, welche den Agenten in die Lage versetzen, durch Interaktion mit der Umgebung \mathcal{E} die optimale Kontrollstrategie π_{opt} zu erarbeiten. Ein Überblick kann [116] sowie [117] entnommen werden. In diesem Abschnitt wird anhand einzelner konkreter Algorithmen die Funktionsweise vorgestellt sowie einzelne Probleme des RL adressiert, welche nach wie vor Gegenstand der Forschung sind. Hierfür soll zunächst anhand der Monte-Carlo-Methoden die Kernüberlegung manifestiert und anschließend der Ansatz mit Hilfe von essentiellen Anpassungen hin zu einem adaptiven System modifiziert und optimiert werden. Hierbei werden die Methoden der *Temporalen Differenz* (aus dem engl. temporal difference; TD) vorgestellt. Abschließend wird eine Brücke zu modernen Algorithmen unter Anwendung von Näherungsverfahren geschlagen, welche mit Hilfe von künstlichen neuronalen Netzen (KNN) komplexe Aufgabenstellungen bewältigen. Im Rahmen der vorliegenden Arbeit wird zwischen den klassischen tabellarischen Verfahren des RL sowie den Näherungsverfahren des DRL unterschieden.

1. Tabellarische Verfahren

Algorithmen, welche in die Gruppe der tabellarischen Verfahren eingeordnet werden, stützen i.d.R. auf die beschriebenen Wertfunktionen des Abschnittes 4.1.4.1. Während der *Lernphase* erfolgt mit fortschreitender Zeit eine Konsolidierung der Ergebnisse der Interaktionen des Agenten mit der Umgebung \mathcal{E} . Hierdurch wird eine kontinuierliche Verbesserung der Kontrollstrategie π' , hin zur optimalen Kontrollstrategie π_{opt} erzielt. Die konkrete Verarbeitung und Bewertung der ausgeführten Aktionen in Abhängigkeit der jeweiligen Methode erfolgt dabei zu unterschiedlichen Zeitpunkten. Dieser Sachverhalt wird nachfolgend thematisiert und führt zu einer grundlegenden Unterscheidung zwischen Monte-Carlo-Methoden und Methoden der TD.

Monte-Carlo-Methoden

Mit Hilfe der Monte-Carlo-Methoden werden im Kontext des RL vorwiegend episodische Aufgaben adressiert. Hierbei durchläuft der Agent eine beliebige Anzahl an Episoden und leitet unter Anwendung der Kontrollstrategie π eine Q-Tabelle mit einzelnen Wertfunktionen der Zustand-Aktion-Tupel $Q_\pi(s, a)$ ab. Die Evaluierung einer Kontrollstrategie π erfolgt dabei erst nach Abschluss einer Episode. Dem *Gesetz der großen Zahlen* der Stochastik folgend konvergieren die Monte-Carlo-Methoden im Allgemeinen für Episode $i \rightarrow \infty$ gegen die optimale Kontrollstrategie π_{opt} , wie in [116] verdeutlicht. Das einzelne Auftreten eines Zustand-Aktion-Tupels (s, a) wird in diesem Kontext allgemein als *Besuch* (aus dem engl. visit) bezeichnet.

Grundsätzlich durchlaufen alle Monte-Carlo-Methoden im Rahmen der *Lernphase* einen *Evaluierung-Verbesserung-Kreislauf* (EVK), welcher folgende zwei Schritte alterniert:

1. Wertfunktion evaluieren: Der aktuellen Kontrollstrategie π folgend wird die Wertfunktion des Zustand-Aktion-Tupels $Q_\pi(s, a)$ je Zustand-Aktion-Tupel (s, a) abgeleitet respektive die Q-Tabelle konstruiert.
2. Kontrollstrategie verbessern: Mit Hilfe der auf Grundlage der aktuellen Kontrollstrategie π konstruierten Q-Tabelle wird durch Anwenden der Gleichung 4.14 eine neue Kontrollstrategie π' mit $\pi' \geq \pi$ abgeleitet.

Dieser Kreislauf, welcher durch Sequenz

$$\pi_0 \xrightarrow{1.} Q_{\pi_0} \xrightarrow{2.} \pi_1 \xrightarrow{1.} Q_{\pi_1} \xrightarrow{2.} \pi_2 \xrightarrow{1.} \dots \xrightarrow{2.} \pi_{\text{opt}} \xrightarrow{1.} Q_{\text{opt}}, \quad (4.15)$$

mit $\xrightarrow{1.}$ stellvertretend für den Evaluierungsschritt nach Schritt 1 respektive $\xrightarrow{2.}$ für den Verbesserungsschritt beschrieben wird, wird allgemein als *generalized policy iteration* (engl., frei übersetzt: generalisierte Iteration der Kontrollstrategie; GPI) bezeichnet (vgl. [116]) und führt zu einer kontinuierlichen Verbesserung der Kontrollstrategie auf Grundlage der Wertfunktion. Dieser Kreislauf wird bis zum Auftreten eines Abbruchkriteriums, wie bspw. der max. Anzahl an Episoden oder bei Konvergenz, durchlaufen. Im Kern dieser Methode liegt die Überlegung, dass dem *Gesetz der Großen Zahlen* der Stochastik folgend der Agent gegen die optimale Kontrollstrategie π_{opt} konvergiert, sofern der EVK lange genug wiederholt wird.

Die skizzierte Alternierung des EVK zum Auffinden der optimalen Kontrollstrategie π_{opt} wird in der einschlägigen Literatur als *Monte-Carlo-Kontrollverfahren* (aus dem engl. monte carlo control) bezeichnet. Die Bestimmung der Wertfunktion, welche zum Konstruieren der Q-Tabelle innerhalb des Evaluierungsschritts nach Schritt 1 unabdingbar ist, wird allgemein als *Monte-Carlo-Vorhersage* (aus dem engl. monte carlo prediction) bezeichnet [116]. Während der Interaktion des Agenten mit der Umgebung \mathcal{E} im Evaluierungsschritt, folgt der Agent zunächst einer Kontrollstrategie $\pi \sim \mathcal{U}$, welche durch eine gleichverteilte Zufallsgröße \mathcal{U} beschrieben wird, die sog. *Equiprobable Random Policy* (engl., frei übersetzt:

gleichverteilte zufällige Kontrollstrategie; EPRP) [116]. Bei dieser einfachen stochastischen Kontrollstrategie haben alle Aktionen A_t im jeweiligen Zustand S_t die gleiche Auftrittswahrscheinlichkeit. Basierend auf der EPRP werden im Zuge der Monte-Carlo-Vorhersage während der *Lernphase* zunächst eine Vielzahl an Zustand-Aktion-Tupel (s, a) besucht und eine Q-Tabelle mit Hilfe der Wertfunktion des Zustand-Aktion-Tupels $Q_\pi(s, a)$ nach Gleichung 4.13 abgeleitet. Tritt ein Zustand-Aktion-Tupel (s, a) innerhalb einer Episode mehrfach auf, sehen die Monte-Carlo-Methoden die Verarbeitung des zu erwarteten Gewinns $\mathbb{E}[G_t]$ nach einer der beiden folgenden Optionen vor:

1. Monte-Carlo-Vorhersage beim erstmaligen Besuch (aus dem engl. first-visit monte carlo prediction): In jeder Episode wird nur der erste Besuch eines Zustand-Aktion-Tupels (s, a) in der Q-Tabelle berücksichtigt.
2. Monte-Carlo-Vorhersage bei jedem Besuch (aus dem engl. every-visit monte carlo prediction): Über die gesamte Episode hinweg finden alle Besuche der Zustand-Aktion-Tupel (s, a) über eine Mittelwertbildung in der Q-Tabelle Berücksichtigung.

Mit Fortschreiten des Experimentes werden Aktionen, auf Grundlage vergangener Interaktion und der Bewertung einzelner Zustände, mit einer höheren Wertfunktion des Zustand-Aktion-Tupels $Q_\pi(s, a)$ favorisiert. Nach [116] konvergieren beide Optionen für Episode $i \rightarrow \infty$. Zur Verdeutlichung der Monte-Carlo-Vorhersage des Evaluierungsschrittes sind exemplarisch die einzelnen Schritte der Option 1 dem Algorithmus 3 des Anhangs B.1 zu entnehmen.

Im Schritt 2, dem Verbesserungsschritt des EVK, kann anschließend eine neue Kontrollstrategie mit

$$\pi' \geq \pi \text{ wenn } Q_{\pi'}(s, a) \geq Q_\pi(s, a), \quad (4.16)$$

über alle $s \in \mathcal{S}$ und $a \in \mathcal{A}(s)$ bestimmt werden. Mit Hilfe von Gleichung 4.14 erfolgt dies direkt aus der Q-Tabelle. Diese Wahl der Aktion $A_t \in \mathcal{A}(s)$, welche den zu erwartenden Gewinn $\mathbb{E}[G_t]$ im jeweiligen Zustand $S_t \in \mathcal{S}$ maximiert, wird in der Terminologie des RL als *gierig* bezeichnet und führt bei der vorliegenden Q-Tabelle zur bestmöglichen Kontrollstrategie π' aller durchlaufenen Episoden und letztlich für $i \rightarrow \infty$ mit $i \in \mathbb{Z}^+$ zur optimalen Kontrollstrategie π_{opt} .

Die beschriebene Monte-Carlo-Methode stellt einen simplen und robusten Ansatz dar und konvergiert unter den Annahmen, dass der Agent $i \rightarrow \infty$ Episoden mit $i \in \mathbb{Z}^+$ durchläuft und jede Episode mit Hilfe der EPRP explorativ beginnt [116]. Grundsätzlich führen diese Annahmen zu folgenden weiterführenden Überlegungen, welche im konzeptionellen Ansatz der Ausgestaltung des Agenten Berücksichtigung finden müssen und nach wie vor Gegenstand aktueller Forschungstätigkeiten sind:

- Erforschen und Gelerntes einsetzen: Das in der einschlägigen Literatur unter *Exploration-Exploitation*-Dilemma (engl., frei übersetzt: Explorations- und Ausbeutungs-dilemma) beschriebene Problem adressiert eine Kompromissfindung zwischen zufälligen

Aktionen, welche eine Exploration des Suchraums ermöglicht und dem Anwenden des erarbeiteten Wissens über die Problemstellung bei der Wahl einer Aktion.

- **Schnelle Konvergenz:** Methoden sowie Verbesserungen bestehender Ansätze zur Auffindung der optimalen Kontrollstrategie π_{opt} werden im Kontext der RL u.a. an der Zuverlässigkeit sowie der numerischen Effizienz hinsichtlich der Konvergenz und der verwendeten Rechenzeit verglichen.

Die beiden genannten Punkte wirken dabei gegeneinander. So steht das ausgiebige Erforschen des Suchraums sowie die theoretische Annahme von $i \rightarrow \infty$ durchgeführten Episoden während der *Lernphase* im Kontrast zu einer schnellen Konvergenz, hin zur optimalen Kontrollstrategie π_{opt} .

ϵ -greedy Kontrollstrategie

Ein in der Praxis weit verbreiteter Ansatz, der die beschriebenen Überlegungen adressiert, stellt die sog. ϵ -greedy Kontrollstrategie dar. Die Motivation hinter einer ϵ -greedy Kontrollstrategie ist, dass ein Agent, welcher mit Blick auf die Maximierung des zu erwarteten Gewinns $\mathbb{E}[G_t]$ in jedem Zustand S_t eine *gierige* Aktion wählt (vgl. Gleichung 4.14), eine suboptimale Kontrollstrategie π ableitet. So wäre der erste Besuch eines Zustand-Aktion-Tupels (s, a) entscheidend, da das *gierige* Verhalten des Agenten im Sinne des RL keine weitere Exploration des Suchraums und das damit einhergehende Auffinden eines *besseren* Zustand-Aktion-Tupels (s, a) im Kontext der Problemstellung zulässt. Folglich muss im Zuge der schnellen Konvergenz gewährleistet sein, dass der Agent erlerntes Wissen nutzt sowie genug Freiraum zum weiteren *Erforschen* des Suchraums im Sinne des *Exploration-Exploitation*-Dilemma hat. Dieser Spagat erfolgt mit Hilfe des dynamischen Parameters ϵ . Der Parameter beschreibt eine Wahrscheinlichkeit, dass der Agent nicht *gierig* (aus dem engl. greedy), sondern zufällig agiert. Hierdurch ist gewährleistet, dass der Agent den Suchraum trotz gesammelter Erfahrungen in Form von Besuchen einzelner Zustand-Aktion-Tupel (s, a) weiter durchsucht und ggf. eine bessere Kontrollstrategie π' im Sinne des Ziels des Agenten findet. Formal wird dieser Sachverhalt mit Hilfe von

$$\pi(a|S_t) = \begin{cases} \arg \max_{a \in \mathcal{A}(s)} Q_{\text{opt}}(S_t, a) & \text{für } 1 - \epsilon \text{ Aktionen} \\ A_t \sim \mathcal{U} & \epsilon \text{ sonst} \end{cases} \quad (4.17)$$

ausgedrückt. Der Parameter $\epsilon \in [0, 1]$ beschreibt, ob der Agent mit $1 - \epsilon$ gierig agieren oder mit $\epsilon = 1$ der EPRP folgend zufällig eine beliebige Aktion A_t wählen soll. Üblicherweise beginnt der Agent innerhalb der *Lernphase* mit $\epsilon = 1$ den Suchraum mit Hilfe der EPRP zu durchsuchen (Exploration). Mit Fortschreiten des Experimentes soll zunehmend das erarbeitete Wissen über die Problemstellung in die Entscheidungsfindung des Agenten einfließen (*Exploitation* über *Exploration* favorisieren) und mit Bezug auf die Wertfunktion des Zustand-Aktion-Tupels Q_π zunehmend *gierig* agieren. Dies erfolgt durch eine kontinuierliche Verkleinerung des Parameters $\epsilon \rightarrow 0$ bis hin zu einem *gierigen* Verhalten des Agenten bei $\epsilon = 0$.

Konvergenzbedingung

Um zu gewährleisten, dass der Agent im Rahmen des Monte-Carlo-Kontrollverfahrens in der optimalen Kontrollstrategie π_{opt} konvergiert, müssen nach [116] folgende Bedingungen gegeben sein:

- Jedes Zustand-Aktion-Tupel (s, a) über alle $s \in \mathcal{S}$ und $\mathcal{A}(s)$ wird unendliche Male besucht.
- Mit Bezug auf die Wertfunktion des Zustand-Aktion-Tupels Q_π konvergiert die angenommene Kontrollstrategie π in einem *gierigen* Verhalten des Agenten.

Diese Bedingungen werden in der einschlägigen Literatur unter *Greedy in the Limit with Infinite Exploration* (engl., frei übersetzt: Gierig im Grenzfall bei unendlicher Erkundung; GLIE) zusammengefasst [116]. Mit Bezug auf die ϵ -greedy Kontrollstrategie erfolgt dies durch ein kontinuierliches Anpassen des Parameters ϵ_i für Episode i während der *Lernphase*. Es gilt $\epsilon_i > 0$ über alle Episoden $i \in \mathbb{Z}^+$ sowie $\lim_{i \rightarrow \infty} \epsilon_i = 0$. Diese rein mathematische Forderung ist in der Praxis impraktikabel. Insbesondere das Ziel einer numerisch effizienten Lösung (mit Bezug auf die Rechenzeit) kann den GLIE-Bedingungen entgegenstehen. So nutzen die Autoren in [47] im Rahmen einer ϵ -greedy Kontrollstrategie während der *Lernphase* ein ϵ , welches zunächst für die erste Million Zustände in Form von Bildinformationen von 1.0 auf 0.1 linear angepasst wurde. Episoden darüber hinaus wurden mit einem festen $\epsilon = 0.1$ trainiert. In Summe wurde der Agent mit bis zu 50 Millionen Bilddaten trainiert.

Die vorgestellte Methode ist in der Lage, lediglich durch Interaktion mit der Umgebung \mathcal{E} , die optimale Kontrollstrategie π_{opt} zu erarbeiten. Die Monte-Carlo-Methode zeichnet dabei ein *Episode-für-Episode* EVK aus, welcher auf einer Mittelwertbildung des zu erwarteten Gewinns $\mathbb{E}[G_t]$ basiert. Hierfür sind keinerlei Kenntnisse über die Problemstellung von Nöten, wie bspw. ein Modell der Übergangswahrscheinlichkeiten \mathcal{P} . Der einschlägigen Literatur sind verschiedenste Optimierungen und Variationen der Monte-Carlo-Methoden zu entnehmen. Dem Anhang B.2 sind zwei Optimierungsansätze, der Inkrementelle Mittelwert sowie die Konstante-Alpha-Methode, näher beschrieben. Im Kontext des RL nimmt insbesondere der Hyperparameter α (vgl. Konstante-Alpha-Methode) eine gesonderte Stellung ein. Dieser beschreibt die Schrittweite und wird allgemein als *Lernrate* interpretiert [122]. Weiterführende Informationen zum Hyperparameter α sind dem Anhang B.2 zu entnehmen.

Temporale Differenz

Das zuvor beschriebene Verfahren samt Optimierungsschritten unterliegt einer wesentlichen Restriktion. So erfolgt das Verarbeiten von *Wissen* in Form von evaluieren und aktualisieren der Q-Tabelle erst nach Abschluss von mindestens einer vollständigen Episode. Dies steht im Kontrast zu einem adaptiven System, welches auf Grundlage des aktuellen Informationsflusses lernt. Einen wesentlichen Schritt in Richtung eines solchen Systems sowie einen wichtigen Meilenstein in der Entwicklung von Methoden des RL stellen die Methoden der TD dar [116]. Wie bei den Monte-Carlo-Methoden basiert das Lernen mit

Hilfe der TD auf einer Interaktion mit der Umgebung \mathcal{E} ohne weiterführende Informationen über das zugrunde liegende Modell. Jedoch erfolgt das Update der Wertfunktion im Evaluierungsschritt des EVK nicht auf einer *Episode-für-Episode*-Basis, sondern auf einer *Zeitschritt-für-Zeitschritt*-Basis.

Konkret bedeutet dies, dass für den Evaluierungsschritt der TD keine vollständige Sequenz einer vollständigen Episode nach 4.7 von Nöten ist. Für Ansätze der TD kann diese auf einen Ausschnitt

$$S_t, A_t, R_{t+1}, S_{t+1}, A_{t+1} \quad (4.18)$$

verkürzt werden. Lediglich das Tupel $(S_t, A_t, R_{t+1}, S_{t+1}, A_{t+1})$ ist notwendig um einen Eintrag $Q(S_t, A_t)$ der Q-Tabelle zum Zeitpunkt t zu aktualisieren. Die drei bekanntesten Verfahren des TD-Lernens, welche nach der zugrunde liegenden Sequenz nach 4.18 benannt wurden, sind SARSA(0), SARSAmax und Expected SARSA. Während der *Lernphase* erfolgt die Wahl der Aktion A_t zum Zeitpunkt t bei allen drei Verfahren i.d.R. nach der vorgestellten ϵ -greedy Kontrollstrategie. Zu unterscheiden sind die Ansätze in der jeweiligen Aktualisierung der Wertfunktion des Zustand-Aktion-Tupels $Q(S_t, A_t)$ der Q-Tabelle wie folgt (vgl. [116]):

1. SARSA(0): Mit Hilfe der zuvor beschriebenen ϵ -greedy Kontrollstrategie erfolgt die Wahl der Aktion A_{t+1} zum Zeitpunkt $t + 1$, welche anschließend mit

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha [R_{t+1} + \gamma Q(S_{t+1}, A_{t+1}) - Q(S_t, A_t)] \quad (4.19)$$

zur Aktualisierung des Eintrags der Q-Tabelle zum Zeitpunkt t Anwendung findet.

2. SARSAmax: Abweichend zum SARSA(0) erfolgt die Wahl der Aktion A_{t+1} zur Aktualisierung des Eintrags der Q-Tabelle dieses Verfahrens der TD nicht mit Hilfe einer ϵ -greedy Kontrollstrategie, sondern einer *gierigen* Wahl der Aktion A_{t+1} zum Zeitpunkt $t + 1$ nach Gleichung 4.14. Demnach erfolgt mit

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha [R_{t+1} + \gamma \max_{a \in \mathcal{A}} Q(S_{t+1}, a) - Q(S_t, A_t)] \quad (4.20)$$

die Aktualisierung des Eintrags $Q(S_t, A_t)$ der Q-Tabelle zum Zeitpunkt t und kann von der ausgeführten Aktion zur Interaktion mit der Umgebung \mathcal{E} abweichen. In der einschlägigen Literatur wird das SARSAmax u.a. als *Q-Learning* bezeichnet [123].

3. Expected SARSA: Abweichend zu SARSAmax wird hierbei nicht die Aktion A_{t+1} gewählt, welche die Wertfunktion des Zustand-Aktion-Tupels maximiert, sondern über alle Aktionen im Schritt S_{t+1} und deren Auftrittswahrscheinlichkeit $\pi(a|S_{t+1})$ summiert. Die Aktualisierung der Q-Tabelle erfolgt durch

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha [R_{t+1} + \gamma \sum_{a \in \mathcal{A}} \pi(a|S_{t+1}) Q(S_{t+1}, a) - Q(S_t, A_t)] \quad (4.21)$$

mit der bedingten Wahrscheinlichkeit $\pi(a|S_{t+1})$ die Aktion $a \in \mathcal{A}$ in Zustand S_{t+1}

auszuführen.

Der Hyperparameter α der Gleichung 4.19, 4.20 sowie 4.21 beschreibt die Lernrate (vgl. Anhang B.2). Zudem kann der TD-Fehler δ jeweils aus der Summe innerhalb der eckigen Klammern bestimmt werden, wie z. B. für SARSA(0) mit $\delta = R_{t+1} + \gamma Q(S_{t+1}, A_{t+1}) - Q(S_t, A_t)$ (vgl. Anhang B.2).

Wie die vorgestellte Monte-Carlo-Methode konvergieren alle drei Verfahren der TD unter Einhaltung der oben genannten Bedingungen in einer optimalen Kontrollstrategie π_{opt} (vgl. [123], [116] sowie [124]). Im Allgemeinen ist bei Ansätzen der TD ein optimistischer Ansatz bei der Initialisierung der Q-Tabelle empfohlen, womit das initiale Belegen der Q-Tabelle mit großen Werten zu verstehen ist. Die Autoren in [125] zeigen, dass dieser Schritt die Performanz erheblich verbessert.

Im Vergleich zu den Optimierungsansätzen der Monte-Carlo-Methoden nach Anhang B.2 ist der Aktualisierung der Q-Tabelle der TD eine wesentliche Änderung zu entnehmen. So wird abweichend zur Gleichung B.2 des Anhangs B.2 nicht der Gewinn G_t zum Aktualisieren der Q-Tabelle zum Zeitpunkt t herangezogen, sondern lediglich die Informationen des Zustandes S_{t+1} sowie der Aktion A_{t+1} des nachfolgenden Zeitpunktes $t + 1$. Diese fundamentale Anpassung der Methoden der TD ermöglicht das Intervall der Aktualisierung der Q-Tabelle von einer *Episode-für-Episode*-Basis hin zu einer *Zeitschritt-für-Zeitschritt*-Basis zu reduzieren. Diese Verbesserung sowie die simple Struktur der TD führt zu einem breiten Anwendungsfeld und einer kontinuierlichen Anpassung und Entwicklung weiterer Algorithmen auf Basis dieser Methoden, wie bspw. TD(λ). Dieses Verfahren beschreibt mit Hilfe des Parameters λ einen Übergang eines TD-basierten Ansatzes mit $\lambda = 0$ hin zu einem Ansatz basierend auf den Monte-Carlo-Methoden mit $\lambda = 1$ [116].

2. Näherungsverfahren

Im Allgemeinen ist das RL auf finite Markow-Entscheidungsprozesse begrenzt, bei welchen die Anzahl an Systemzuständen und Aktionen limitiert ist [116]. Bei Problemstellungen mit großem Zustands- sowie Aktionsraum, \mathcal{S} respektive \mathcal{A} , kann es zu Problemen hinsichtlich der numerischen Handhabbarkeit kommen. Dieser Sachverhalt tritt bereits bei Problemstellungen mit bspw. stetigen Zuständen hervor, da traditionelle Ansätze, wie die zuvor beschriebenen tabellarischen Verfahren, eine finite Struktur nutzen. Diese resultieren im Falle eines stetigen Zustandsraums \mathcal{S} in einer unendlich großen Tabelle. Zudem kann bei großen Zustandsräumen nicht garantiert werden, dass Zustände im Rahmen der *Lernphase* mehrfach oder überhaupt besucht werden, was gegen einer akkuraten Entscheidungsfindung spricht.

Der einschlägigen Literatur sind mehrere mögliche Herangehensweisen zur Berücksichtigung der beschriebenen Problematik zu entnehmen [116]. Nachfolgend sollen zwei Ansätze betrachtet werden:

1. Diskretisierung: Mit Hilfe der Diskretisierung wird der Zustand- oder Aktionsraum in eine endliche, diskrete Teilmenge unterteilt. Ein möglicher, generischer Ansatz zur Diskretisierung beschreibt die sog. *Kachel-Codierung* (aus dem engl. tile coding) [49]. Bei der *Kachel-Codierung* wird der kontinuierliche Zustands- oder Aktionsraum in mehrere, zueinander versetzte *Kacheln* zerlegt. Die Position eines Zustandes oder einer Aktion wird anschließend durch alle Kacheln, welche den Zustand bzw. die Aktion beinhalten, beschrieben. Zur Steigerung der numerischen Effizienz erfolgt dies i.d.R. mit Hilfe eines Bits, welches die jeweiligen Zustände oder Aktionen über die Werte 1 und 0 maskiert. Eine Erweiterung der *Kachel-Codierung* ist durch die *adaptive Kachel-Codierung* beschrieben. Hierbei erfolgt die Ausdehnung der einzelnen Kacheln dynamisch, in Abhängigkeit der jeweiligen Informationsdichte.

Neben der exemplarisch beschriebenen *Kachel-Codierung* und deren Erweiterung sind der einschlägigen Literatur weitere generische Ansätze zur Diskretisierung zu entnehmen, wie z. B. das *Coarse Coding* [116]. Auch wenn mit Hilfe der Diskretisierung das Nutzen von Algorithmen, welche *diskrete* Zustands- sowie Aktionsräume vorsehen, auf Problemstellungen mit *stetigen* ermöglicht wird, so sind die Problemstellungen auf kleine Zustands- sowie Aktionsräume begrenzt. Die Autoren in [46] legen anhand einer Beispielsrechnung dar, dass die Diskretisierung für Problemstellungen im hochdimensionalen Raum nicht praktikabel ist und begründen dies mit einem erhöhten numerischen Aufwand und der damit einhergehenden Rechenzeit.

2. Näherungsverfahren: Eine weitere Option ist durch die Approximation gegeben. Das hierdurch beschriebene Näherungsverfahren sorgt für eine generalisierte Betrachtung der Problemstellung. Mit Hilfe der Parametrisierung wird das zugrunde liegende Modell durch Parametrieren des Parametervektors θ angenähert. Im einfachsten Fall kann dies mit Hilfe einer linearen Funktion erfolgen. Für eine generalisierte Betrachtung, welche eine nichtlineare Approximation einschließt, erfolgt bspw. der Einsatz von KNN. Diese weisen eine einfache Struktur auf und sind in der Lage, eine beliebige, kontinuierliche Funktion zu approximieren. Letzteres wurde in [126] mathematisch bewiesen. Der Autor weist darauf hin, dass lediglich eine (stückweise) kontinuierliche Aktivierungsfunktion vonnöten sei, wie z. B. der Tangens hyperbolicus oder die Sigmoidfunktion. Aus Gründen der Performanz finden im Rahmen des RL darüber hinaus die sog. *Rectified Linear Unit* (engl., frei übersetzt: gleichgerichteten linearen Einheiten; ReLU) Anwendung [127].

Während die unter Punkt 1 beschriebene Diskretisierung einer weiteren Limitierung unterliegt, zeigt die Veröffentlichung nach [47] vielversprechende Ergebnisse hinsichtlich eines generalisierten Lösungsansatzes unter Verwendung von DNN im Zuge der Approximation der Problemstellung. Die Autoren präsentieren einen Algorithmus, welcher in der Lage ist Videospiele auf dem Niveau eines Menschen zu spielen. Das hierdurch ausgelöste mediale Echo führte zu einer Renaissance des RL und letztlich zur neuen Forschungsdisziplin des DRL. Aufbauend auf den Erkenntnissen nach [47] veröffentlichten Lillicrap et al. in [49]

einen Ansatz, welcher zudem kontinuierliche Aktionsräume explizit berücksichtigt. Beide genannten Publikationen nutzen den Ansatz des tiefen Lernens um den Zustandsraum \mathcal{S} in [47] bzw. den Zustands- und Aktionsraum \mathcal{S} respektive \mathcal{A} in [49] zu approximieren. Das *tief* in DRL verweist an dieser Stelle auf das tiefe Lernen, welches zum Lösen der Problemstellung ein KNN mit mehreren verdeckten Schichten nutzt. Beim DRL wird ein solches (Multilayer) KNN genutzt um Fragestellungen des RL zu lösen.

Im Allgemeinen stellen Verfahren des DRL einen vergleichsweise neuen Ansatz dar. Im Rahmenwerk des (D)RL wird grundlegend zwischen Ansätzen unterschieden, welche zum einen die Approximation einer Wertfunktion vorsehen, wie z. B. des Zustand-Aktion-Tupels (als wertbasiert bezeichnet) mit

$$Q(s_t, a_t | \theta^Q) \approx Q(s_t, a_t), \quad (4.22)$$

zum anderen eine direkte Näherung der Kontrollstrategie (als kontrollstrategiebasiert bezeichnet) mit

$$\pi(a|s, \theta) \approx \pi(a|s). \quad (4.23)$$

Algorithmen, welche die Optimierung basierend auf der Kontrollstrategie vorsehen, wie bspw. der REINFORCE Algorithmus nach [128], werden im Rahmenwerk des RL als *Policy Gradient* (engl., frei übersetzt: Kontrollstrategiegradient; PG) zusammengefasst. PG stellen insbesondere bei kontinuierlichen Domänen, welche die Lösung von Problemstellungen mit stetigen Zustands- und Aktionsräumen beschreiben, einen weitverbreiteten Ansatz dar [49].

Für beide Ansätze, wert- und kontrollstrategiebasiert, können im Kontext des DRL KNN für die nichtlineare Approximation vorgesehen werden, womit eine Parametrisierung mit Hilfe des Parametervektors θ beschrieben wird (vgl. Gleichung 4.22 sowie 4.23). Das Parametrieren erfolgt i.d.R. mit Hilfe von Gradientenverfahren.

Der Abbildung 4.4 sind die zuvor beschriebenen tabellarischen Verfahren (vgl. Abbildung 4.4 linke Seite) sowie weitere Algorithmen, gruppiert unter der Bezeichnung der Näherungsverfahren, zu entnehmen. Zudem ist der Einfärbung die Unterscheidung zwischen Algorithmen, welche einen wertbasierten Ansatz (gelb) bzw. kontrollstrategiebasierten Ansatz (rot) verfolgen, zu entnehmen. Eine Kombination beider Ansätze, welche in der einschlägigen Literatur als *Actor-Critic*-Methoden geführt werden ist grün eingefärbt. Hierbei beschreibt der *Critic* (engl., frei übersetzt: Kritiker) einen Part des Algorithmus, welcher i.d.R. einen wertbasierten Ansatz dargestellt und eine Grundlinie für den *Actor* (engl., frei übersetzt: Akteur) vorgibt. Dieser stellt einen PG dar, welcher basierend auf der Kontrollstrategie die Aktion A_t im jeweiligen Zeitpunkt t ausführt. Demnach erfolgt im Allgemeinen durch den *Critic* bei *Actor-Critic*-Methoden eine Bewertung der Güte der Aktion des *Actors* im jeweiligen Zeitpunkt.

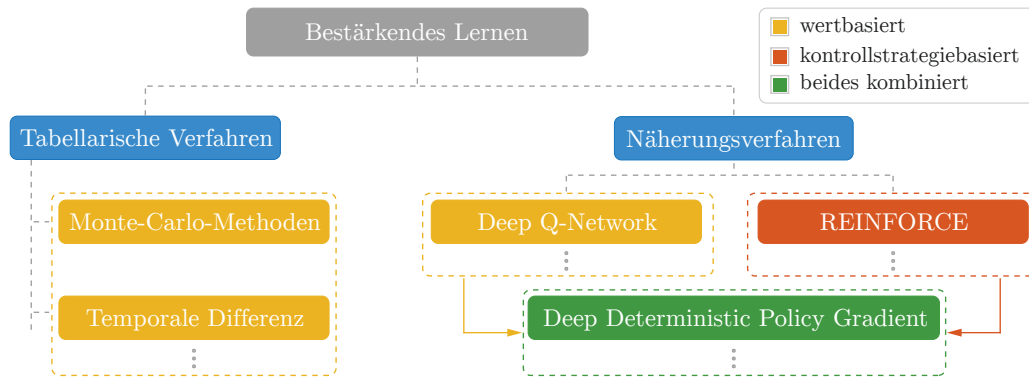


Abbildung 4.4.: Schematische Darstellung unterschiedlicher Algorithmen des RL. Die Gruppen der tabellarischen Verfahren und Näherungsverfahren sowie die Einordnung der unterschiedlichen Algorithmen, welche im Rahmen der vorliegenden Arbeit genannt werden, sind der Darstellung zu entnehmen. Zudem ist farblich hervorgehoben, ob der Aufbau des Agenten eine wert- oder kontrollstrategiebasierte Optimierung, gelb respektive rot, bzw. beides kombiniert (grün), vorsieht. Agenten, welche beide Optimierungen verfolgen, sind in der einschlägigen Literatur unter *Actor-Critic* klassifiziert.

Nachfolgend werden die einzelnen Methoden der verschiedenen Gruppen der Näherungsverfahren nach Abbildung 4.4 (rechte Seite) mit Hilfe der Veröffentlichungen nach [47], [49] und [128] näher beschrieben. Insbesondere die Verbesserungen innerhalb der *Lernphase* sollen hervorgehoben und diskutiert werden.

Deep Q-Network

Anwendungen der realen Welt weisen häufig einen hohen Grad an Komplexität auf. Sollen diese Problemstellungen mit Hilfe des Rahmenwerks des RL gelöst werden, so müssen die Algorithmen und Lösungsansätze in der Lage sein, hochdimensionale Informationen mit dem Ziel zu verarbeiten, dass historische Informationen auf neue Situationen adaptiert werden können [47]. Zudem stehen Wissenschaftler im Kontext des RL vor der Herausforderung, einen einzelnen Algorithmus zu entwickeln, welcher in der Lage ist selbstständig Lösungsansätze für verschiedenste Problemstellungen zu erarbeiten. Dieser als generativer Algorithmus bezeichnete Ansatz ist nach wie vor Stand der Wissenschaft und führte in der jüngsten Vergangenheit zu einer Vielzahl an wissenschaftlichen Publikationen im Rahmen des (D)RL [110] [114]. Der Veröffentlichung nach [47] ist ein Algorithmus zu entnehmen, welcher die genannten Anforderungen erfüllt. Die Autoren betiteln diesen mit Deep Q-Network (DQN).

Der DQN ist ein modellfreier RL-Agent, welcher auf Grundlage von DNN hochdimensionale Sensordaten in Form von Pixelinformationen verarbeitet. Dem Ansatz zugrunde liegend ist das in Abschnitt 4.1.4.2 beschriebene Verfahren des SARSAmax. Die Wertfunktion des Zustand-Aktion-Tupels

$$Q(s, a | \theta^Q) = \mathbb{E}_{R_t, S_t \sim \mathcal{E}, A_t \sim \pi} [G_t | S_t, A_t] \quad (4.24)$$

wird in der Arbeit nach [47] mit Hilfe von sog. *faltenden neuronalen Netzwerken* (aus dem engl. convolutional neural network), einer speziellen Form der KNN, welche insbesondere

bei der Verarbeitung von Bildinformationen zum Einsatz kommen, approximiert. Der von den Autoren beschriebene Ansatz resultiert in einer durch Parametervektor θ und mittels faltender neuronaler Netzwerke approximierten Wertfunktion des Zustand-Aktion-Tupels (vgl. Gleichung 4.22), auch als Q-Netzwerk bezeichnet.

Im Allgemeinen gilt das Rahmenwerk des RL in Kombination mit nichtlinearer Funktionsapproximation zur Abbildung einer Wertfunktion als instabil und neigt zur Divergenz [47]. Mnih et al. weisen in [47] darauf hin, dass Lösungsansätze zur Stabilisierung der *Lernphase* existieren, diese jedoch nicht praktikabel sind, insbesondere nicht mit großen KNN. Um diese Lücke zu schließen, werden in [47] zwei essentielle Verbesserungen, welche die Parametrierung des Parametervektors θ im Kontext des Rahmenwerks des RL adressieren, vorgestellt:

1. *Erfahrungs-Wiedergabepuffer* (aus dem engl. experience replay buffer; ERB): Der ERB, in der einschlägigen Literatur in kurz als *Wiedergabepuffer* bezeichnet, ist ein finiter Pufferspeicher \mathcal{D} mit Speichergröße M . Im Zuge der *Lernphase* werden während der Interaktion des Agenten mit der Umgebung \mathcal{E} für jeden Zeitpunkt t einzelne Beobachtungen in Form eines Tupels $e_t = (s_t, a_t, r_t, s_{t+1})$ im ERB $\mathcal{D} = \{e_1, \dots, e_M\}$, basierend auf der aktuell gültigen Kontrollstrategie π , gespeichert. Anschließend wird einer Gleichverteilung $\mathcal{U}(\mathcal{D})$ folgend, zufällig ein *kleiner Stapel* X (aus dem engl. mini-batch) an Übergängen $(s_i, a_i, r_i, s_{i+1}) \sim \mathcal{U}(\mathcal{D})$ aus dem ERB gezogen. Abschließend wird diese gleichmäßig nach dem Zufallsprinzip gezogene Stichprobe X von Übergängen e verwendet, um den Parametervektor θ^Q des Q-Netzwerks zu aktualisieren. Diese Herangehensweise sorgt für eine Dekorrelation der Beobachtungen und letztlich zu einer Stabilisierung des Parametrierprozesses des Agenten [47].
2. *Zielnetzwerk* (aus dem engl. target network; TN): Das TN stellt eine Kopie des Q-Netzwerks θ^Q mit $\theta^{Q'}$ dar und wird periodisch alle C Schritte mit $\theta^{Q'} = \theta^Q$ aktualisiert. Zwischen diesem sog. *harten Aktualisieren* werden die Parameter $\theta^{Q'}$ konstant gehalten. Hierdurch wird gewährleistet, dass das Aktualisieren der Parameter gegen ein fixiertes Q-Netzwerk erfolgt. Dies resultiert in einer reduzierten Korrelation der Parameter des Q-Netzwerks während der *Lernphase* [47].

Abweichend zur Aktualisierung der Einträge der Q-Tabelle nach Abschnitt 4.1.4.2 mit Hilfe der Methode der TD, wird der Parametervektor θ^Q des Q-Netzwerks $Q(s, a|\theta^Q)$ je Übergang $e_i = (s_i, a_i, r_i, s_{i+1})$ der Stichprobe X durch Minimieren der Zielfunktion

$$\mathcal{L}(\theta^Q) = \mathbb{E}_{(s_i, a_i, r_i, s_{i+1}) \sim \mathcal{U}(\mathcal{D})} \left[\left(y_i - Q(s_i, a_i | \theta^Q) \right)^2 \right] \quad (4.25)$$

mit

$$y_i = r_i + \gamma \max_{a'} Q'(s_{i+1}, a' | \theta^{Q'}) \quad (4.26)$$

aktualisiert. Der Gleichung 4.26 ist das TN $Q'(s_{i+1}, \pi'(s_{i+1}) | \theta^{Q'})$, der korrespondierende

Parametervektor θ^Q sowie der *Diskontierungsfaktor* $\gamma \in [0, 1]$ zu entnehmen. Wie bei den oben beschriebenen tabellarischen Verfahren nach Abschnitt 4.1.4.2 wird auch beim DQN das *Exploration-Exploitation*-Dilemma mit Hilfe der beschriebenen ϵ -greedy Kontrollstrategie adressiert. Hierbei wählten die Autoren in [47] ein ϵ , welches während der *Lernphase* linear reduziert und anschließend fixiert wurde.

Die Lösung der Zielfunktion nach Gleichung 4.25 erfolgt durch Bildung des Gradienten

$$\nabla_{\theta^Q} \mathcal{L}_i(\theta^Q) = \mathbb{E}_{(s_i, a_i, r_i, s_{i+1}) \sim \mathcal{U}(\mathcal{D})} \left[\left(y_i - Q(s_i, a_i | \theta^Q) \right) \nabla_{\theta^Q} Q(s_i, a_i | \theta^Q) \right] \quad (4.27)$$

unter Berücksichtigung der Kettenregel [116]. Der Gradient nach Gleichung 4.27 wird aus numerischen Gründen im Kontext des Rahmenwerks des RL durch ein stochastisches Gradientenabstiegsverfahren bestimmt [47]. Die Parametrierung der KNN erfolgt mit Hilfe der sog. *Fehlerrückführung* (aus dem engl. backpropagation), welche ein Gradientenabstiegsverfahren beschreibt. Eine Erweiterung hierzu, welche mit Hilfe eines *Trägheitsmomentes* die Wahrscheinlichkeit für das Verweilen in lokalen Minima reduzieren soll, ist durch den adaptiven Momentum Schätzer (ADAM) nach [129] gegeben. Die Lösung mit Hilfe des ADAM ist im Kontext des DRL weit verbreitet und findet in [47] Anwendung. Dem Algorithmus 4 des Anhangs B.3 sind die einzeln beschriebenen Schritte der *Lernphase* des DQN zu entnehmen.

Mit dem Algorithmus des DQN nach [47] und dessen essentiellen Verbesserungen mit Bezug auf die *Lernphase* des Agenten ist es den Autoren gelungen einen einzigen, modellfreien Algorithmus des Rahmenwerks des RL zu entwickeln, der in der Lage ist Aufgaben unterschiedlicher Komplexität unter Berücksichtigung eines kontinuierlichen Zustandsraums \mathcal{S} und diskreten Aktionsraums \mathcal{A} zu lösen. Zudem führte die erfolgreich vorgestellte Kombination des RL mit DNN des DQN zur neuen Forschungsdisziplin des DRL [47].

Deep Deterministic Policy Gradient

Wie oben beschrieben ist ein zentrales Ziel des Rahmenwerks des (D)RL der Umgang mit komplexen Aufgabenstellungen der *realen Welt*. Diese, welche insbesondere durch unverarbeitete, hochdimensionale Sensordaten geprägt sind, werden nach [49] als eine der Kernaufgaben des Anwendungsfeldes des RL beschrieben. Folglich sehen aktuelle Entwicklungen einen generativen Algorithmus vor, der in der Lage ist, Aufgaben direkt aus den unverarbeiteten, hochdimensionalen sensorischen Daten zu lösen [49]. Mit dem zuvor vorgestellten DQN haben die Autoren in [47] bewiesen, dass mit Hilfe eines einzelnen Algorithmus ein breites Spektrum an komplexen Aufgaben gelöst werden kann. Jedoch ist der vorgestellte Ansatz auf einen diskreten Aktionsraum begrenzt. Die Notwendigkeit eines entsprechenden Ansatzes, welcher hochdimensionale, kontinuierliche Zustands- und Aktionsräume berücksichtigt, ist insbesondere in Anwendungen der realen Welt begründet. Als Beispiel sei an dieser Stelle ein Industrieroboter genannt, der auf Grundlage von Bildinformationen Gegenstände greift. Die Bildinformationen als Pixeldaten sowie die aktuelle Position eines

mehrachsigem Industrieroboter stellen den Zustandsraum dar, während die Geschwindigkeit und Position jeder einzelnen Achse den kontinuierlichen Aktionsraum beschreiben. Auch im Energieversorgungssystem sind Entscheidungsprozesse und Regelungsaufgaben zu finden, welche einen kontinuierlichen Aktionsraum \mathcal{A} vorsehen, wie bspw. das CVR oder das Blindleistungsmanagement.

Der DQN basiert auf der Idee des SARSAmax nach [123], beschrieben in Abschnitt 4.1.4.2. Ein essentieller Part des Ansatzes ist eine Wahl der Aktion, welche die Wertfunktion des Zustand-Aktion-Tupels $Q_\pi(s, a)$ maximiert. Im Fall eines kontinuierlichen Aktionsraums würde diese, in Gleichung 4.14 formulierte $\arg \max$ -Operation eine globale Optimierung erfordern. Mit Hilfe der oben beschriebenen Diskretisierung könnte der Aktionsraum in eine endliche, diskrete Teilmenge unterteilt werden, sodass der DQN für Fragestellungen im kontinuierlichen Aktionsraum Anwendung findet. Hierdurch würde der Ansatz jedoch der oben beschriebenen Limitierung eines *kleinen* Aktionsraums unterliegen [49]. Zwar existieren Kontrollstrategien und Algorithmen des Rahmenwerks des RL, welche einen kontinuierlichen Aktionsraum verarbeiten können, wie bspw. der REINFORCE Algorithmus nach [128]. Diese Gruppe der sog. *stochastischen* PG finden häufig Anwendung bei Problemstellungen mit kontinuierlichem Zustands- und Aktionsraum. Jedoch tendieren diese auf KNN basierenden Ansätze im Allgemeinen zum Divergieren [130].

Um die Lücke eines generativen Algorithmus zu schließen, der vergleichbar zu der Leistungsfähigkeit des DQN ist und eine kontinuierliche Domäne berücksichtigt, wurde in [49] der Deep Deterministic Policy Gradient (engl., frei übersetzt: tiefer deterministischer Kontrollstrategiegradient; DDPG) vorgestellt. Der DDPG ist, wie der DQN, ein modellfreier Algorithmus, welcher DNN zur Funktionsapproximation nutzt. Der Algorithmus basiert auf einer einfachen Struktur, die durch eine *Actor-Critic*-Architektur beschrieben wird (illustriert in Abbildung 4.5).

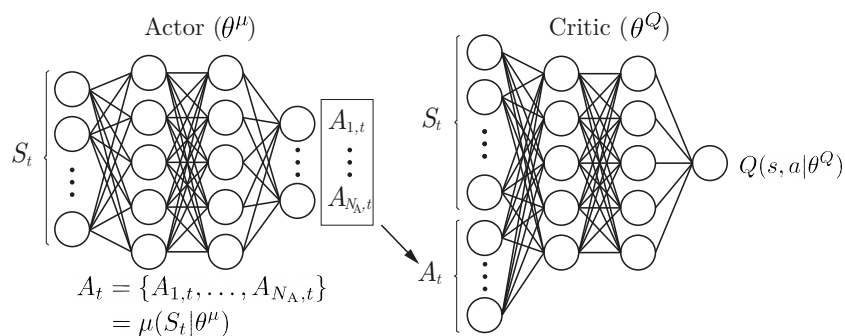


Abbildung 4.5.: Darstellung einer *Actor-Critic*-Architektur, die das Arbeitsprinzip des DDPG veranschaulicht (eigene Darstellung nach [133]).

Die *Actor-Critic*-Architektur basiert auf dem sog. *Policy Gradient Theorem* (engl., frei übersetzt: Kontrollstrategie-Gradienten-Theorem) in [116] und besteht im Wesentlichen aus zwei Komponenten. Der *Critic*, der durch eine Wertfunktion des Zustand-Aktion-Tupels $Q(s, a | \theta^Q)$ beschrieben wird. Die Parametrierung erfolgt mit Hilfe der sog. *Mean Squared Bellman Equation* (engl., frei übersetzt: mittlere quadratische Bellman Abwei-

chung; MSBE). Wie beim DQN erfolgt die Optimierung aus Gründen der Performanz durch ein stochastisches Gradientenabstiegsverfahren.

Die zweite Komponente der *Actor-Critic*-Architektur ist der *Actor*, der üblicherweise durch einen *stochastischen* PG beschrieben wird. Anstelle einer Wertfunktion approximieren Algorithmen der Gruppe der PG die stochastische Kontrollstrategie $\pi_\theta(s)$ direkt. Begründet durch die Zieldefinition des RL, den erwarteten Gewinn $\mathbb{E}[G_t]$ zu maximieren, wird der Parametervektor θ bei der Methode der PG mit Hilfe des stochastischen Gradientenaufstiegsverfahrens angepasst [130]. Aufbauend auf der Idee des *stochastischen* PG haben die Autoren in [130] den mathematischen Beweis für die Existenz eines *deterministischen* PG (DPG) mit einer Kontrollstrategie $\mu(s|\theta^\pi)$ erbracht, welcher die Leistungsfähigkeit des *stochastischen* PG signifikant übertrifft und die Basis für den in diesem Abschnitt beschriebenen Ansatz des DDPG bildet.

Die Aktualisierung der Parametervektoren innerhalb der *Lernphase* stützt auf einem, einer Gleichverteilung $\mathcal{U}(\mathcal{D})$ folgenden, zufällig gezogenen Stichprobe X mit B Tupeln an Übergängen $(s_i, a_i, r_i, s_{i+1}) \sim \mathcal{U}(\mathcal{D})$. Die Parametrierung des Parametervektors θ^μ des *Critic* erfolgt unter Anwendung eines stochastischen Gradientenabstiegsverfahrens und basiert auf der Minimierung des MSBE

$$\mathcal{L}(\theta^Q) = \mathbb{E} \left[(y_i - Q(s, a | \theta^Q))^2 \right], \quad (4.28)$$

mit $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s | \theta^{\mu'}) | \theta^{Q'})$. Parallel hierzu wird basierend auf derselben Stichprobe X der Parametervektor θ^Q des *Actors* aktualisiert. Hierbei wird das Ziel verfolgt, die Aktion $a_i = \mu(s_i | \theta^\mu)$ zu finden, welche den durch die Ausgabe des *Critics* $Q(s_i, a_i | \theta^Q)$ repräsentierten, zu erwarteten Gewinn maximiert. Dies erfolgt unter Anwendung des Gradienten

$$\nabla_{\theta^\mu} \mathcal{J}(\theta^\mu) \approx \mathbb{E} \left[\nabla_{\theta^\mu} Q(s_i, \mu(s_i | \theta^\mu) | \theta^Q) \right] \quad (4.29)$$

mit Hilfe eines stochastischen Gradientenaufstiegsverfahrens. Die beschriebenen einzelnen Schritte können dem Algorithmus 5 des Anhangs B.4 entnommen werden.

Zusammengefasst basiert der DDPG auf der Idee des DPG nach [130], kombiniert mit den zuvor beschriebenen Verbesserungen der *Lernphase* des DQN in [47]. Ergänzend hierzu wurden die nachfolgenden Anpassungen, mit dem Ziel der Stabilisierung der Parametrierung der Parametervektoren während der *Lernphase*, vorgenommen. So wurde das periodische, harte Aktualisieren der TN im Ansatz des DDPG hin zu einem *weichen Aktualisieren* mit

$$\theta' \leftarrow \tau\theta + (1 - \tau)\theta' \quad (4.30)$$

modifiziert. Es gilt $\tau \ll 1$ für das Actor- sowie das Critic-Netzwerk der TN, $\mu'(s | \theta^{\mu'})$ respektive $Q'(s, a | \theta^{Q'})$. Dieser durch τ parametrisierte Aktualisierungsschritt resultiert in einer Stabilisierung der *Lernphase*, begründet durch eine im Vergleich zum *harten Aktua-*

lisieren des DQN langsameren Änderung der Parametervektoren θ^Q und θ^μ des TN [49].

Eine zweite Ergänzung zum Algorithmus des DQN ist die sog. *Stapel-Normalisierung* (aus dem engl. batch normalisation) nach [131]. Mit Hilfe dieser wird die Tatsache berücksichtigt, dass der Zustandsraum \mathcal{S} einer beliebigen Umgebung \mathcal{E} unterschiedliche Informationen, physikalische Einheiten sowie Größenordnung einzelner Observationen aufweisen kann. Durch die in [131] beschriebene Normalisierung kann die Problemstellung verallgemeinert und die *Lernphase* folglich stabilisiert werden [46].

Ein weiterer essentieller Unterschied zwischen den Ansätzen des DQN und DDPG ist im Umgang mit dem *Exploration-Exploitation*-Dilemma zu finden. So wird dieses bei Algorithmen, welche unter die Gruppe der PG fallen, nicht durch eine ϵ -greedy Kontrollstrategie adressiert, wie es bspw. bei den oben beschriebenen tabellarischen Verfahren sowie dem DQN der Fall ist. Stattdessen erfolgt dies für kontinuierliche Domänen typischerweise durch das additive Hinzufügen eines zusätzlichen Rauschprozesses (vgl. Zeile 11 des Algorithmus 5 des Anhangs B.4) mit

$$a_t = \mu(s_t|\theta^\mu) + \omega_t . \quad (4.31)$$

Der Rauschprozess $\omega_t \sim \mathcal{W}$ ist in Abhängigkeit der jeweiligen Problemstellung zu wählen und kann bspw. durch ein Gauß'sches Rauschen beschrieben werden. Eine weitere Option ist durch den Ornstein-Uhlenbeck-Prozess nach [132] gegeben, welcher u.a. in [46] zum Einsatz kommt.

Der DDPG weist im Allgemeinen eine gute Leistungsfähigkeit auf und kommt der Beschreibung eines generalisierten Algorithmus nach. So bewerten die Autoren in [49] unterschiedlichste Ansätze zur Lösung von Problemstellungen mit Hilfe des Rahmenwerks des RL. Im Fokus der Analyse steht die Skalierbarkeit und das Potenzial zur Generalisierung der Ansätze. Darüber hinaus vergleichen die Autoren den DDPG und den DQN mit Hilfe einer identischen Problemstellung mit dem Resultat, dass der DDPG um einen Faktor zwanzig weniger Episoden zur adäquaten Lösung der Problemstellung benötigt.

Im Allgemeinen kommen die Autoren der Publikation zu folgendem Schluss:

„*It seems that deep function approximators are the most promising approach for scaling reinforcement learning to large, high-dimensional domains.*“ [49] (engl., frei übersetzt: Tiefe Funktionsapproximationen scheinen die vielversprechendsten Ansätze für die Skalierung des bestärkenden Lernens auf große, hochdimensionale Bereiche darzustellen.)

Die Analyse hat folglich ergeben, dass Verfahren, welche DNN als Funktionsapproximation nutzen, besser skalieren und im Allgemeinen dem Anspruch eines generalisierten Ansatzes am nächsten kommen. Darüber hinaus lässt die Bewertung der Autoren den Schluss zu, dass der vorgestellte Ansatz des DDPG eine im Vergleich wesentlich größere Bandbreite an komplexen Problemstellungen lösen kann. Basierend auf dieser Bewertung und den zuvor

genannten Punkten findet der DDPG im Kontext der vorliegenden Arbeit Anwendung.

4.2. Rahmenwerk für den Einsatz im elektr. Energieversorgungssystem

Wie einleitend beschrieben soll das in Abbildung 1.1 des Kapitels 1 illustrierte Rahmenwerk in der vorliegenden Arbeit angewandt werden. Die einzelnen Module sowie der damit einhergehende Arbeitsmodus des Agenten sind im nachfolgenden Abschnitt beschrieben.

4.2.1. Lernmodul

Das *Lernmodul* stellt eine typische Implementierung der *Lernphase* des Rahmenwerks des RL dar. Hierbei wird insbesondere das zuvor thematisierte *Exploration-Exploitation*-Dilemma adressiert. Hierdurch wird gewährleistet, dass der Agent zu Beginn zufällig Systemzustände über einen bestimmten Zeitraum *erkundet*. Mit Fortschreiten des Experimentes und einer damit einhergehenden zunehmenden Anzahl an Episoden und besuchten Zustand-Aktion-Tupeln (s, a) fließt das *erarbeitete Wissen* über die Problemstellung kontinuierlich in die Entscheidungsfindung ein. Dies gewährleistet, dass der Agent gegen eine optimale Kontrollstrategie π_{opt} konvergiert. Hierbei muss sichergestellt sein, dass der Agent möglichst viele Systemzustände mehrfach *besucht*, um die Systemantwort in Form der Belohnung adäquat zu bewerten. Dies wird in der vorliegenden Arbeit durch die folgenden drei Punkte sichergestellt:

1. Detaillierte Simulationsumgebung

Wie bereits erläutert, sieht das Rahmenwerk des RL für Regelungsaufgaben und Entscheidungsprozesse bei Fragestellungen mit Bezug auf das Energieversorgungssystem aus Sicherheitsgründen eine Trennung in zwei Phasen vor. Dies ist mit der beschriebenen, zufälligen Erkundung des Zustandsraums auf der Suche nach der optimalen Kontrollstrategie π_{opt} des Agenten begründet [28]. Um systemkritische Zustände in der realen Umgebung $\mathcal{E}_{\text{real}}$ zu vermeiden, erfolgt das Erlernen und Ableiten einer optimalen Kontrollstrategie π_{opt} unter Zuhilfenahme einer Simulationsumgebung \mathcal{E}_{sim} . Diese Anforderung legt eine Modellierungstiefe der Simulationsumgebung \mathcal{E}_{sim} zugrunde, welche die zu lösende Problemstellung möglichst realitätsnah wiedergibt. Zudem sollten bei Fragestellungen mit Bezug auf das Energieversorgungssystem Unsicherheiten Berücksichtigung finden. Insbesondere da exogene Größen, wie bspw. meteorologische Größen, den elektr. Verteilungsprozess heutzutage enorm beeinflussen.

Neben exogenen Größen können auch endogene Informationen fehlerhaft vorliegen. Demnach können auch physikalische Parameter des elektr. Verteilungsprozesses mit Unsicherheiten behaftet sein, wie z. B. Informationen über Netzteilnehmer, Endkunden und Netzparameter. Letztlich kann ein fehlerhaft zugrundeliegender Informationsstand zur Ableitung einer fehlerhaften oder sogar systemkritischen Kontrollstrategie führen. Im Allgemeinen resultiert die Berücksichtigung von Unsicherheiten im Zuge des Ableitens der Kontrollstra-

ategie innerhalb des *Lernmodules* in einem stabileren Verhalten und einer besseren Leistungsfähigkeit des Agenten [42].

Im Rahmen der vorliegenden Arbeit wird die Simulationsumgebung \mathcal{E}_{sim} auf Basis einer Netzsimulation umgesetzt. Die Modellierung von Verbraucherverhalten sowie exogener Größen erfolgt durch WDF. Im Rahmen einer Monte-Carlo-Simulation werden diese stochastischen Prozesse durch zeitvariante ZV abgebildet. Dezentrale Erzeugungsanlagen in Form von PVA werden detailliert unter Berücksichtigung spezifischer Anlagenparameter und in Abhängigkeit exogener Größen modelliert. Darüber hinaus wird die Spannungsabhängigkeit einzelner Knotenleistungen explizit modelliert und im Rahmen einer Leistungsflussberechnung berücksichtigt. Eine umfangreiche Beschreibung der einzelnen Last- und Einspeiseprozesse der Simulationsumgebung sowie die Modellierung spannungsabhängiger Prozesse im elektr. Verteilnetz ist dem Kapitel 2 zu entnehmen.

2. Szenarien

Mit Hilfe von definierten Szenarien können spezifische Systemzustände innerhalb der Simulationsumgebung \mathcal{E}_{sim} forciert werden. Hierdurch kann der Suchraum begrenzt sowie das Auftreten einzelner Systemzustände erzwungen werden. Dies gewährleistet, dass der Agent im Zuge der Interaktion mit der Simulationsumgebung und dem daraus resultierenden *Belohnungssignal* die optimale Kontrollstrategie, unter Einsatz einer adäquaten Rechenzeit, ableiten kann. Zudem kann mit Hilfe der Definition von Szenarien die Problemstellung, z. B. auf einen Normalbetrieb des elektr. Verteilnetzes eingegrenzt werden.

3. Rückführung von Beobachtungen

Die Rückführung von Beobachtungen, in der einschlägigen engl. Literatur als *feedback loop* (engl. für Rückkoppelungsschleife) bezeichnet, beschreibt eine Anreicherung der Simulationsumgebung mit Observationen des realen Systems, wie in Abbildung 1.1 des Kapitels 1 angedeutet. Dieser Wissenstransfer aus dem *realen* System $\mathcal{E}_{\text{real}}$ in die Simulationsumgebung \mathcal{E}_{sim} kann in einer verbesserten Kontrollstrategie resultieren und hilft neue Gegebenheiten und Bedingungen innerhalb des *realen* Systems kontinuierlich im Zuge der *Lernphase* zu berücksichtigen. Diese periodische Rückführung von Beobachtungen führt zu einer kontinuierlichen Anpassung an neue Dynamiken, was einem adaptiven System gleicht. Dies ist insbesondere bei komplexen Systemen notwendig, da nicht davon ausgegangen werden kann, dass die Simulationsumgebung alle möglichen Systemzustände erfasst und die Systemdynamik ganzheitlich widerspiegelt [28].

4.2.2. Ausführendes Modul

Im *ausführenden Modul* nutzt der Agent die im *Lernmodul* erarbeitete Kontrollstrategie zur Lösung der Regelungsaufgabe (vgl. Abbildung 1.1 des Kapitels 1). In dieser Phase wählt der Agent für jeden einzelnen Systemzustand S_t zum Zeitpunkt t die Aktion A_t , welche die erwartete Summe der diskontierten Belohnungen über zukünftige Zeitschritte mit Bezug auf die zu lösende Regelungsaufgabe maximiert. Wie in den vorangegangenen Abschnitten

detailliert ausgeführt, wird diese Arbeitsweise des Agenten in der einschlägigen Literatur als *gierig* bezeichnet.

4.3. Weitbereichsregelung im elektr. Niederspannungsnetz

Zur Bewertung des Rahmenwerks des DRL im Zuge einer aktiv herbeigeführten Spannungsanpassung als Handlungsoption im intelligenten Netzbetrieb soll eine Weitbereichsregelung im elektr. Verteilnetz auf Niederspannungsebene adressiert werden. Im Fokus des Untersuchungsszenarios liegt die Auffindung optimaler Positionen des Stufenschalters eines rONT im Tagesverlauf, unter Anwendung des in Abschnitt 4.1.4.2 vorgestellten DDPG. Hierbei ist das Ziel, die Versorgungsspannung innerhalb der vorgegebenen Spannungsgrenzen zu halten. Dies erfolgt unter der praxisnahen Prämisse, die Anzahl an Schaltoperationen des Stufenschalters auf ein Minimum zu reduzieren. Letztlich verkürzt eine erhöhte Anzahl an Schaltoperationen das Intervall zur nächsten Revision des rONT, was wiederum in erhöhten Wartungskosten resultiert. Das Adressieren dieses Sachverhalts erfolgt im Rahmen der vorliegenden Arbeit erstmals mit einer berechneten Steigung über ein definiertes Intervall an zurückliegenden Observationen, welche im Zustandsraum als auch in der Belohnungsfunktion Berücksichtigung findet.

Als Grundlage für das beschriebene Szenario dienen die im Forschungsprojekt nach [36] untersuchten Netze sowie die erhobenen Messdaten. Das zugrunde liegende Netzgebiet ist am Niederrhein verortet und weist eine ländliche Struktur sowie eine hohe Durchdringung an PVA auf (vgl. Abbildung A.6 des Anhangs A.7). Das Niederspannungsnetz verfügt über einen rONT mit einer Scheinleistung von $S_{r, \text{rONT}} = 400 \text{ kVA}$, 114 Haushalte sowie 24 PVA. Die Netztopologie besteht aus zwei Abgängen mit einer Gesamtlänge der Kabel von $l = 3515 \text{ m}$ sowie einer projektierten, generatorischen Einspeiseleistung von $P_{\text{PVA}} = 398 \text{ kW}_p$ der PVA. Zudem sind der Abbildung A.6 in Summe sechs Messsysteme an unterschiedlichen Netzknoten zu entnehmen. Neben der Messung am rONT sind diese in Kabelverteilerschränken positioniert.

Zur Ableitung einer Kontrollstrategie mit den zuvor definierten Zielen muss der DDPG im Zuge der *Lernphase* in die Lage versetzt werden, durch Interaktion mit der Simulationsumgebung \mathcal{E}_{sim} die (optimale) Position des Stufenschalters zu jedem Zeitpunkt $t \in \mathcal{T}$ des Tages $\mathcal{T} = [1, 1440] \text{ min}$ zu erarbeiten. Dies muss durch sorgfältig definierte Signale des Markow-Entscheidungsprozesses sichergestellt sein.

4.3.1. Signale des Markow-Entscheidungsprozesses

Ziel des RL ist die Ableitung einer optimalen Kontrollstrategie π_{opt} durch Interaktion mit einer Umgebung \mathcal{E} . Für Fragestellungen, die Regelungsaufgaben und Entscheidungsprozesse mit Bezug auf das Energieversorgungssystem adressieren, erfolgt dies typischerweise im *Lernmodul* mit Hilfe einer Simulationsumgebung \mathcal{E}_{sim} (vgl. Abbildung 1.1 des Kapitels 1). Um den DDPG in die Lage zu versetzen die beste Aktion $A_t \in \mathbb{R}^{N_A}$ zu erar-

beiten, muss die Netzsimulation dahingehend erweitert werden, dass diese einen Markow-Entscheidungsprozess mit dem Tupel $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$ beschreibt. Die Definition der Signale der Aktion, des Zustandes sowie der Belohnung müssen die Problemstellung vollumfänglich, unter Einhaltung der beschriebenen Markow-Eigenschaft, beschreiben. Für ein Adressieren der Weitbereichsregelung mit Hilfe eines rONT seien die Signale sowie die Episode wie folgt definiert:

1) Aktion: Der definierte Markow-Entscheidungsprozess ist über einen kontinuierlichen Aktionsraum \mathcal{A} beschrieben, welcher die Position des Stufenschalters Tap_t , zwischen den physikalischen Grenzen $Tap_{limit,\pm}$ des rONT innerhalb der Umgebung \mathcal{E} repräsentiert. Folglich ist die einzelne Aktion des Agenten durch den Wechsel der Position $\pm\Delta tap_t$ des Stufenschalters definiert, wodurch Einfluss auf die Netzversorgungsspannung um $\pm\Delta U_{rel}$ je Schaltstufe genommen wird. Die Aktion $A_t \in \mathbb{R}^{N_A}$ mit $N_A = 1$ ist als reelle Zahl $A_t \in [-1, 1]$ definiert und wird innerhalb der Umgebung \mathcal{E} in $N_{Tap,total}$ mögliche Schaltstufen des rONT diskretisiert.

2) Zustand: Der Zustand $S_t \in \mathcal{S}$ des Markow-Entscheidungsprozesses ist als Vektor von Tupeln $(U, f_m)_{n,t}$ definiert. Die einzelnen Tupel $(U, f_m)_{n,t}$ beschreiben die gemessene Spannung U am Knotenpunkt $n \in \mathcal{N}$ zum Zeitpunkt t für N in Summe gemessener Knotenpunkte. Ergänzt wird die Spannungsmessung zudem durch die Information, welchen Trend $m_{n,t}$ die Spannungsmessung U am Knotenpunkt n über einen zurückliegenden Zeitraum von $\mathcal{K}_t = \{k_t, k_{t-1}, \dots, k_{t-K}\}$ mit $K \subset \mathcal{T}$ zum Zeitpunkt t hat. Zur besseren numerischen Verarbeitung wird die Steigung $m_{n,t}$ mit einem Parameter h skaliert, sodass durch

$$f_m = h \cdot m_{n,t} \quad (4.32)$$

die Information des Tupels vervollständigt wird. Diese zusätzliche Information ist dahingehend notwendig, da i.d.R. die Wahl der Aktion $+\Delta tap_t$ bzw. $-\Delta tap_t$ direkt mit einer positiven bzw. negativen Steigung $f_m > 0$ respektive $f_m < 0$ korreliert. Diese Annahme ist analog zu den Erkenntnissen in [47] zu sehen und resultiert im Allgemeinen in einem vollständig observierbaren Markow-Entscheidungsprozess.

Neben dem beschriebenen Tupel $(U, f_m)_{n,t}$ je gemessenem Knotenpunkt n findet die aktuell zum Zeitpunkt t vorliegende Position des Stufenschalters Tap_t des rONT Berücksichtigung im Zustandsraum. Demzufolge ist der Systemzustand S_t zum Zeitpunkt $t \in \mathcal{T}$ definiert durch

$$S_t = ((U, f_m)_{1,t}, \dots, (U, f_m)_{N,t}, Tap_t). \quad (4.33)$$

3) Belohnung: Bei der Anwendung des Rahmenwerks des RL auf eine beliebige Problemstellung muss die Zieldefinition des Entscheidungsprozesses durch eine zugehörige Beloh-

nungsfunktion \mathcal{R} mit Belohnung $R_t = \mathcal{R}(S_t, A_t, S_{t+1})$ zum Zeitpunkt t beschrieben werden. Die vorliegende Untersuchung adressiert die aktive Spannungshaltung im stationären Fall. Der Fokus liegt hierbei auf der Beseitigung von kritischen Spannungen mit Bezug auf die Spannungsbandgrenzen bei minimalen Kosten. Als aktives Element ist der rONT definiert, welcher Einfluss auf die jeweils gemessene Spannung U am Knotenpunkt n nimmt. Um diese definierte Anforderung zu adressieren, besteht die Belohnungsfunktion im Wesentlichen aus zwei Teilen. Mit Hilfe des ersten Teils können Spannungsgrenzen, analog zu den Spannungsqualitätsmerkmalen (reglementiert in der DIN EN 50160), definiert werden. Hierdurch ist gegeben, dass der Agent zwischen diesen vorgegebenen Spannungsgrenzen bleibt. Zu jedem Zeitpunkt t am Knotenpunkt $n \in \mathcal{N}$ wird eine Belohnung

$$R_{n,t}(U, b) = \frac{100b}{\left(\frac{1}{b}|U_0 - U|^3 + 50b\right)} - 1 \quad (4.34)$$

berechnet. Per Definition resultiert die Belohnung des Agenten mit Parameter $b = 2V^{\frac{3}{2}}$ in

$$R_{n,t}(U \approx U_0, b = 2V^{\frac{3}{2}}) = 1 \quad (4.35)$$

für Knotenspannungen $U_{n,t} \approx U_0$ nahe der Nennspannung U_0 . Knotenspannungen mit einer Abweichung von etwa $\pm 3\%$ zur Nennspannung U_0 werden durch eine Belohnung

$$R_{n,t}(U \pm 0.03 \cdot U_0, b = 2V^{\frac{3}{2}}) < 0 \quad (4.36)$$

negativ bestärkt.

Der zweite Teil der Belohnungsfunktion adressiert eine Reduzierung der Schalthäufigkeit des rONT und wird durch eine mit Parameter h sowie der Anzahl an Schaltoperationen Δtap_t skalierte Steigung $m_{n,t}$

$$f_{r_{n,t}} = |h \cdot m_{n,t} \cdot \Delta tap_t| \quad (4.37)$$

an Knotenpunkt n zum Zeitpunkt t beschrieben. Die Steigung $m_{n,t}$ der linearen Gleichung basiert auf der Berücksichtigung eines gleitenden Fensters von Spannungsmessungen $U_{n,t}$ über $\mathcal{K} \subset \mathcal{T}$ zurückliegende Schritte.

Da die Belohnung $R_t \in \mathbb{R}$ im Rahmen des Markow-Entscheidungsprozesses durch eine skalare Größe beschrieben wird, kann durch die Summation beider Teile der Belohnungsfunktion eine gesamt Belohnung

$$R_t = \sum_{n=0}^N R_{n,t}(U, b) - f_{r_{n,t}} \quad (4.38)$$

über die Summe aller gemessenen Knotenpunkte $n \in N$ zum Zeitpunkt t berechnet werden.

4) **Episode:** Die Problemstellung ist als abgeschlossene Episode definiert. Eine Episode repräsentiert eine Tagessimulation und ist für jeden Zeitschritt $t \in \mathcal{T}$, beschrieben durch die Minuten des Tages mit $\mathcal{T} = [1, 1440]$ min, als stationärer Systemzustand definiert. Hierbei liegt im Fokus der vorliegenden Arbeit der Normalbetrieb des elektr. Niederspannungsnetzes. Folglich sind keine Störfälle, wie bspw. der Ausfall eines Betriebsmittels, innerhalb einer Episode definiert und demnach nicht Teil des Untersuchungsszenarios.

4.3.2. Ableiten einer Kontrollstrategie innerhalb des Lernmoduls

Wie in Abschnitt 4.2 beschrieben, sieht das *Lernmodul* die Berücksichtigung konkreter Szenarien sowie die Anreicherung der Simulationsumgebung \mathcal{E}_{sim} mit Observationen aus dem *realen* System $\mathcal{E}_{\text{real}}$ im Kontext der *Rückkoppelungsschleife* vor. Um zu gewährleisten, dass der Agent im Zuge des Ableitens einer (optimalen) Kontrollstrategie möglichst *viele* Systemzustände mehrmals observiert, werden folgende Szenarien definiert:

- Extremszenario a) nach Abschnitt 2.1.1.1 Max. Leistungsbedarf bei min. Rückspeisung.
- Extremszenario b) nach Abschnitt 2.1.1.1 Min. Leistungsbedarf bei max. Rückspeisung.
- Im Intervall zwischen den beiden Extremszenarien soll im Szenario des *Normalbetriebs* der Leistungsbedarf der Haushalts-Endkunden sowie der Leistungseinspeisung, erzeugt durch PVA, zufällig mit Hilfe einer Monte-Carlo-Simulation generiert werden. Hierbei werden Ereignisse berücksichtigt, die eine direkte Korrelation zu der definierten Problemstellung haben, wie bspw. die Globalstrahlung.
- Um den Einfluss des überlagerten Umspannwerks auf die Versorgungsspannung am Knotenpunkt des rONT (Unterspannungsseite) zu berücksichtigen, werden bis zu zehn zufällig positionierte Spannungsänderungen $\pm \Delta U_{\text{rel}}$ pro Episode simuliert.
- Im Zuge der *Rückkoppelungsschleife* werden Messreihen des Netzes am Referenzknoten als Versorgungsspannung angenommen.

Die ersten beiden Aufzählungspunkte definieren die Randwerte des Zustandsraums. Auch wenn diese Szenarien eine geringe Auftrittswahrscheinlichkeit haben [63], ist deren Auftreten unter Normalbetrieb möglich, sodass diese Berücksichtigung im Untersuchungsszenario finden.

Der dritte und vierte Aufzählungspunkt adressiert explizit Unsicherheiten im System. Das Wetter als *exogene* Unsicherheit, die im Rahmen der vorgestellten Problemstellung und Netztopologie nach Abbildung A.6 des Anhangs A.7 auf die Globalstrahlung reduziert werden kann, wird in Punkt drei durch einen zufälligen Bewölkungsgrad $f_{t,\text{BG}} \in (0, 1]$ je Zeitpunkt t berücksichtigt. Dieser skaliert die Leistungseinspeisung der PVA im Versorgungsgebiet gleichermaßen.

Die sprunghaften Spannungsänderungen als *endogene* Unsicherheit wird im vierten Aufzählungspunkt berücksichtigt. Diese können im elektr. Energieversorgungssystem auf Basis unterschiedlichster Ereignisse eintreten. Im Kontext der vorliegenden Problemstellung ist die sprunghafte Spannungsänderung am Referenzknoten durch ein zusätzliches $\pm\Delta U_{\text{rel}}$ modelliert.

Das letzte beschriebene Szenario der Aufzählung adressiert den Fakt, dass nicht sichergestellt ist, dass mit Hilfe einer Simulationsumgebung, welche ein Abbild des realen Systems darstellen soll, alle Systemzustände erfasst werden. Die beschriebene *Rückkoppelungsschleife* hilft diese Lücke zu schließen [28]. Hierfür werden im Rahmen der vorliegenden Arbeit Spannungsmessungen der ONS, aufgenommen im Forschungsprojekt nach [36], herangezogen und innerhalb der Simulationsumgebung am Referenzknoten nachgefahren. Dies ermöglicht, dass der Agent während der Lernphase zusätzliche Systemzustände erfährt, welche den Vorgängen im realen System entsprechen. Eine Ableitung einer robusten Kontrollstrategie für den Einsatz im realen System wird hierdurch begünstigt.

Neben den definierten Szenarien muss zur Sicherstellung der Konvergenz gegen eine (*gute*) Kontrollstrategie das skizzierte *Exploration-Exploitation*-Dilemma adressiert werden (vgl. Abschnitt 4.1.4.2). Wie bereits erwähnt, findet bei Ansätzen des RL mit diskretem Aktionsraum \mathcal{A} die ϵ -greedy Kontrollstrategie Anwendung. In kontinuierlichen Domänen hingegen wird das *Exploration-Exploitation*-Dilemma durch Überlagern der gewählten Aktion A_t mit einem zusätzlichen Rauschterm ω_t realisiert [49]. Im Fall des gewählten Ansatzes, des DDPG mit einer deterministischen Kontrollstrategie μ , resultiert dies in

$$\mu'(s_t) = \mu(s_t | \theta^\mu) + \omega_t \quad (4.39)$$

mit $\omega_t \sim \mathcal{W}$ einem Rauschprozess \mathcal{W} folgend, wie bspw. dem Gauß'schen Rauschen, welches Anwendung im vorliegenden Untersuchungsszenario findet.

Das Ableiten der Kontrollstrategie innerhalb des *Lernmoduls* erfolgt durch Interaktion mit der Simulationsumgebung \mathcal{E}_{sim} , unter Berücksichtigung der oben genannten Punkte. Hierbei werden die drei in Abbildung 4.6 visualisierten Schritte wiederholt, bis das Abbruchkriterium erfüllt ist. Nach dem Initialisieren der Simulationsumgebung \mathcal{E}_{sim} und Laden des aktuellen Szenarios sowie der Netztopologie nach Abbildung A.6 des Anhangs A.7 wird eine Leistungsflussberechnung durchgeführt, um den initialen Systemzustand S_0 zu bestimmen. Anschließend wählt der Agent, basierend auf der aktuell gültigen Kontrollstrategie μ , eine Aktion $a = \mu(S_t = s | \theta^\mu)$. Entsprechend Schritt ③ der Abbildung 4.6, wird das beobachtete Tupel $e_t = (S_t, A_t, R_t, S_{t+1})$ des Übergangs zum Zeitpunkt t im ERB \mathcal{D} abgelegt. Die Simulationsumgebung \mathcal{E}_{sim} nimmt dem Schritt ① der Abbildung folgend die gewählte Aktion $A_t = a$ entgegen und bestimmt auf Grundlage einer weiteren Leistungsflussberechnung und unter Berücksichtigung der gewählten Aktion die resultierende Belohnung R_t sowie den neuen Systemzustand $S_{t+1} = s'$. Das Parametrieren der Parametervektoren des DDPG, markiert als Schritt ② in Abbildung 4.6, wird so lange übersprungen,

bis der ERB \mathcal{D} mindestens B Elemente aufweist und eine Stichprobe X an Übergängen $(S_i, A_i, R_i, S_{i+1}) \sim \mathcal{U}(\mathcal{D})$ einer Gleichverteilung \mathcal{U} folgend zufällig gezogen werden kann. Anschließend werden alle drei skizzierten Schritte je Zeitpunkt t ausgeführt.

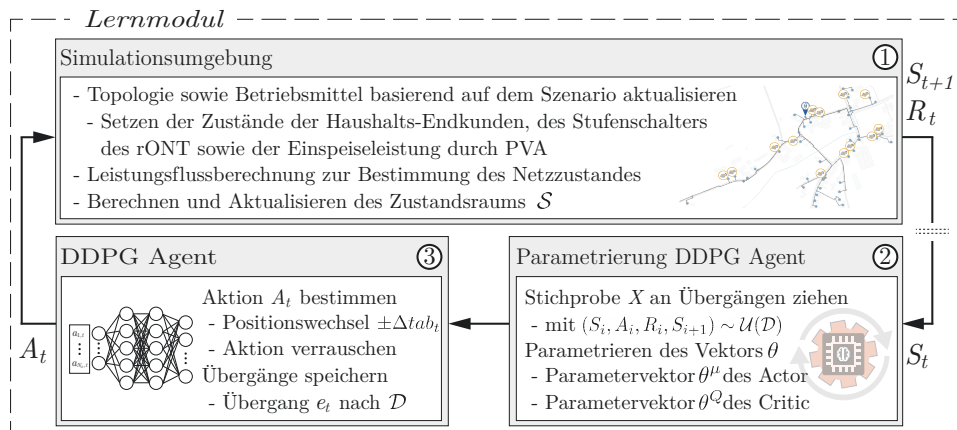


Abbildung 4.6.: Einzelne Schritte zur Ableitung einer Kontrollstrategie μ innerhalb des *Lernmoduls*.

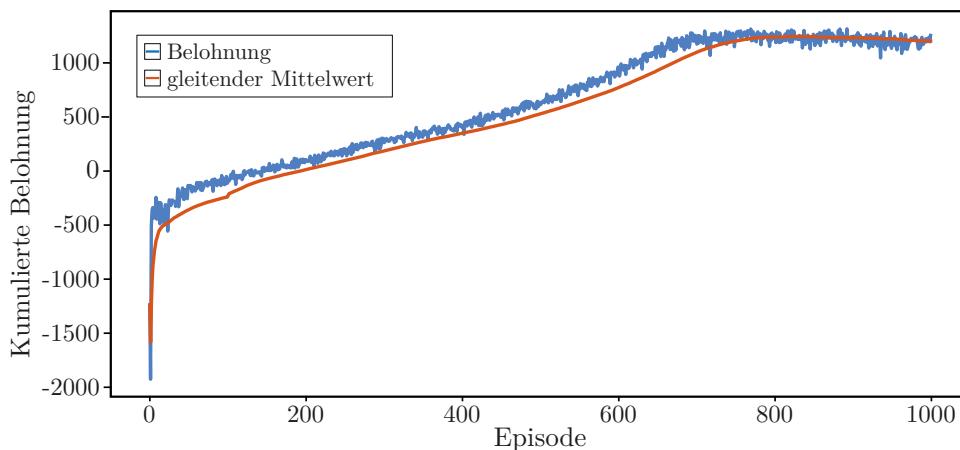


Abbildung 4.7.: Entwicklung der kumulierten Belohnung G_t über der Episode (blau) sowie der gleitende Mittelwert von G_t über ein Intervall von 100 Episoden (rot).

Zur Visualisierung des Lernfortschrittes des DDPG ist in Abbildung 4.7 die Entwicklung der kumulierten Belohnung G_t über der Episode (blau) dargestellt. Die rote Kennlinie zeigt einen gleitenden Mittelwert über ein Intervall von 100 Episoden. Es ist deutlich zu erkennen, dass der Agent eine *gute* Kontrollstrategie im Kontext der Problemstellung erarbeitet. Basierend auf dem *Exploration-Exploitation*-Dilemma erkundet der Agent zu Beginn der *Lernphase* durch zufällige Aktionen den Zustandsraum. Mit fortschreitender Anzahl an Episoden nutzt der Agent zunehmend die gesammelten Informationen über die Problemstellung im Zuge der Entscheidungsfindung. Demnach sind die Aktionen des Agenten zum Ende des Lernprozesses *erfahrungsbasiert* und der Prozess konvergiert in einer Kontrollstrategie.

4.3.3. Verifizierung der Kontrollstrategie

Im Rahmen der vorliegenden Arbeit wird das *ausführende Modul* zur Verifizierung des vorgestellten Ansatzes genutzt. Für diesen Zweck werden zwei unterschiedliche Agenten definiert und deren abgeleitete Kontrollstrategie mit Hilfe von Messdaten, aufgenommen im Forschungsprojekt nach [36], bewertet. Für den Vergleich werden die Agenten nach erfolgreicher Ableitung einer Kontrollstrategie in einen *gierigen* Arbeitsmodus versetzt. Folglich wählen die jeweiligen Agenten die Aktion A_t zum Zeitpunkt t , welche den zu erwarteten diskontierten Gewinn $\mathbb{E}[G_t]$ mit Bezug auf die jeweils abgeleitete Kontrollstrategie μ maximiert.

Des Weiteren sieht das *ausführende Modul* eine Interaktion mit dem *realen* System $\mathcal{E}_{\text{real}}$ vor (vgl. Abbildung 1.1 des Kapitels 1). Das *reale* System $\mathcal{E}_{\text{real}}$ wird im Rahmen der vorliegenden Arbeit mit Hilfe der vorgestellten Modellierung spannungsabhängiger Last- und Einspeiseprozesse, einer Netzsimulation sowie den im Forschungsprojekt erhobenen Informationen modelliert. Hierbei erfolgt die Verifizierung der Kontrollstrategie anhand von 41 vollständig gemessenen Tagen an Spannungswerten. Im Zuge einer Vorverarbeitung werden die Spannungswerte dieser Validierungsdaten dahingehend angepasst, dass Schaltaktionen des im Niederspannungsnetz verbauten rONT herausgefiltert werden. Die Validierungsdaten wurden während der *Lernphase*, innerhalb der Rückkopplungsschleife nicht genutzt. Die Spannungsmessungen werden am Referenzknoten nachgefahren, wodurch der zum Zeitpunkt der Messung vorliegende Netzzustand möglichst nachgebildet wird. Wie der Abbildung 1.1 des Kapitels 1 zu entnehmen ist, erfolgt innerhalb des *ausführenden Moduls* keine Belohnung des Agent. Es findet lediglich das ausführen einer Aktion A_t auf Basis des aktuellen Systemzustandes S_t statt. Wie in Abschnitt 4.3.1 beschrieben, sieht der beschriebene Markow-Entscheidungsprozess einen Wechsel der Stufenschalterposition $\pm\Delta tap_t$ als Aktion vor, wodurch neben dem Nachfahren der Spannungsmessung am Referenzknoten eine zusätzliche Anpassung der Netzversorgungsspannung um $\pm\Delta U_{\text{rel}}$ je Schaltstufe erfolgt.

Im Zuge der Verifizierung und Bewertung soll der Fokus zum einen auf den Beitrag des RL im Kontext einer Weitbereichsregelung im Niederspannungsnetz sowie dem autonomen Führen elektr. Verteilnetze gelegt werden. Zum anderen soll eine Bewertung der Berücksichtigung der beschriebenen Steigung, welche den Trend der Spannungsmessung beschreibt, erfolgen. Hierfür werden zwei Agenten definiert, mit (Agent A1) und ohne (Agent A2) Berücksichtigung des Trends der Spannungsmessung innerhalb des Zustandsraums und der Belohnungsfunktion nach Abschnitt 4.3.1. Dies resultiert in abweichenden Signalen innerhalb des *Lern-* und *ausführenden Moduls* für Agent A2. So wird Gleichung 4.32 sowie 4.37 durch $f_m = h \cdot \Delta tap_t$ bzw. $f_{r_{n,t}} = |h \cdot \Delta tap_t|$ ersetzt.

Das Verhalten beider Agenten im Vergleich zur gemessenen Nennspannung U_0 als Referenz ist der Abbildung 4.8 zu entnehmen. Die Abbildung weist zwei Diagramme auf, welche zum einen die Nennspannung U_0 am Referenzknoten visualisiert (oberes Diagramm), zum anderen die Position des Stufenschalters Tap_t im Tagesverlauf (unteres Diagramm). Auf

dem oberen der beiden Diagramme ist deutlich zu erkennen, dass beide Agenten in der Lage sind, die Spannung innerhalb der definierten Spannungsbandgrenzen (obere und untere rote Markierung) nahe der Nennspannung U_0 zu halten (mittlere rote Markierung). Das untere Diagramm weist bei einer Gegenüberstellung weniger Schalthandlungen des Stufenschalters für den Agenten A1 auf, welcher den Trend der Spannungsmessung explizit berücksichtigt.

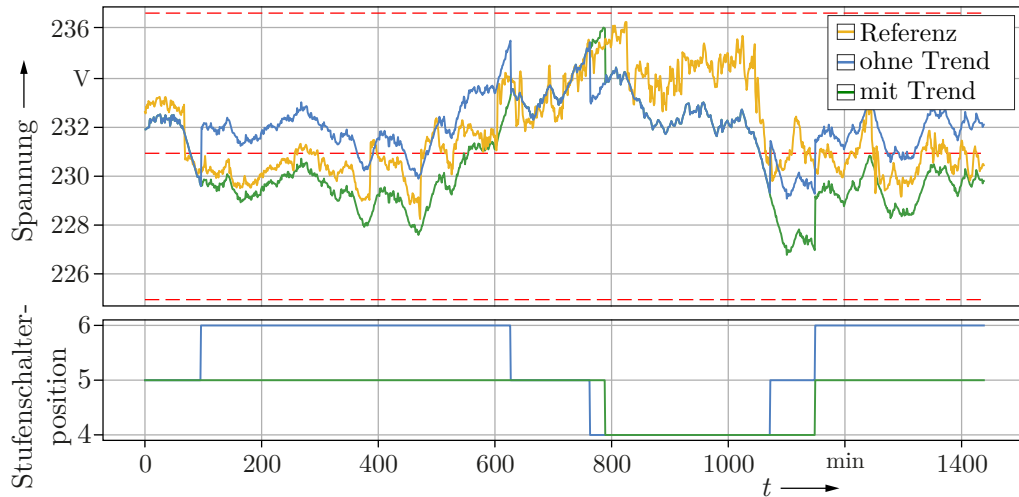


Abbildung 4.8.: Darstellung der Nennspannung U_0 (oben) sowie der Position des Stufenschalters Tap_t über der Minute des Tages $t \in \mathcal{T}$ zur Bewertung und Vergleich der Agenten mit (Agent A1 in grün) und ohne (Agent A2 in blau) Berücksichtigung des Trends der Spannungsmessung innerhalb des Zustandsraums und der Belohnungsfunktion. Als Referenz ist dem oberen Diagramm die unveränderte Spannungsmessung zu entnehmen.

Im Allgemeinen ist in Abbildung 4.8 das in Abschnitt 4.3.1 definierte Verhalten, beschrieben durch das Belohnungssignal, zu entnehmen. Dieses wurde als Summe von zwei Teilen definiert. Der erste Teil der beschriebenen Belohnungsfunktion soll sicherstellen, dass der Agent Aktionen A_t ausführt, welche die Versorgungsspannung des Netzes innerhalb der fest definierten Spannungsbandgrenzen, möglichst nahe an der vorgegebenen Nennspannung $U_0 \approx 230.94$ V, hält. Um das richtige Verhalten der Agenten mit Bezug auf den ersten Teil der Belohnungsfunktion zu bewerten sind in Abbildung 4.9 acht zufällig gewählte Tage der Validierungsdaten mit Hilfe von Box-Plots visualisiert. Die einzelnen Box-Plots der Abbildung verdeutlichen den Schwerpunkt der im Tagesverlauf aufgetretenen Spannungswerte. In der Gesamtheit sind der Abbildung 4.9 drei Datensätze je Tag zu entnehmen:

1. Die Spannungsmessung am Netzknoten des rONT als Referenz (gelb).
2. Die resultierenden Spannungen nachdem Agent A1 Aktion $A_{t,A1} = \mu_{A1}(s|\theta_{A1}^Q)$ der Kontrollstrategie μ_{A1} folgend ausgeführt hat, dargestellt in grün.
3. In blau die resultierenden Spannungen nachdem Agent A2 Aktion $A_{t,A2} = \mu_{A2}(s|\theta_{A2}^Q)$ der Kontrollstrategie μ_{A2} folgend ausgeführt hat.

Der Abbildung ist im Allgemeinen eine erhöhte Referenzspannung (gelb) im Vergleich zur vorgegebenen Nennspannung $U_0 \approx 230.94$ V der beschriebenen Problemdefinition zu entnehmen. Hierdurch ist sichergestellt, dass der Einfluss der Agenten deutlich hervorsticht. So

liegt der Median der Box-Plots der Agenten A1 (grün) und A2 (blau) an jedem der acht visualisierten Tage näher an der definierten Nennspannung U_0 im Vergleich zur Spannungsmessung (gelb).

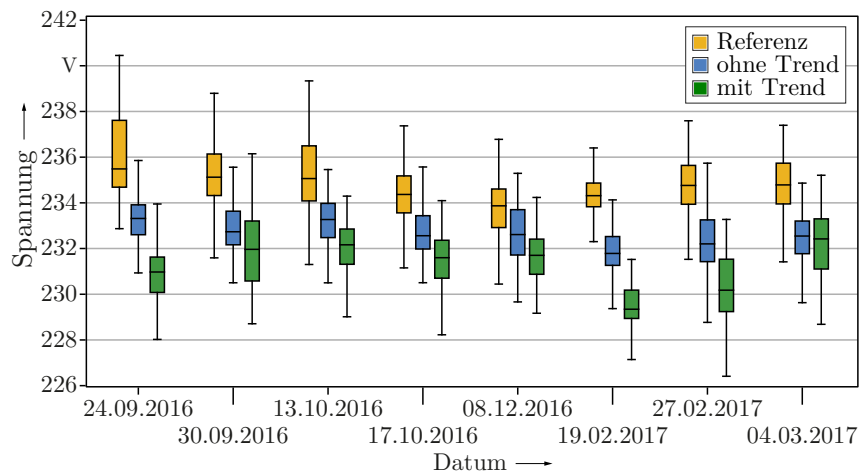


Abbildung 4.9.: Box-Plots zur Visualisierung des Schwerpunkts der im Tagesverlauf aufgetretenen Spannungswerte an acht verschiedenen Tagen. Als Referenz ist die Spannungsmessung am Netzknoten des rONT dargestellt (gelb) sowie die resultierenden Spannungen nachdem die Agenten A1 (grün) und A2 (blau) die jeweiligen Aktionen ausgeführt haben.

Der zweite Teil, der mit Hilfe der definierten Belohnungsfunktion adressiert wurde, zielt auf eine Reduzierung der Schalthäufigkeit ab. Um die reduzierte Schalthäufigkeit durch den beschriebenen Ansatz zu bewerten, wird die Anzahl der Schalthandlungen der Agenten A1 und A2, welche jeweils an den 41 Tagen des Validierungsdatensatzes aufgetreten sind, im Histogramm nach Abbildung 4.10 verglichen. Der Abbildung ist zu entnehmen, dass der Agent A1 (grün) im Allgemeinen weniger Schalthandlungen durchführt als Agent A2 (blau).

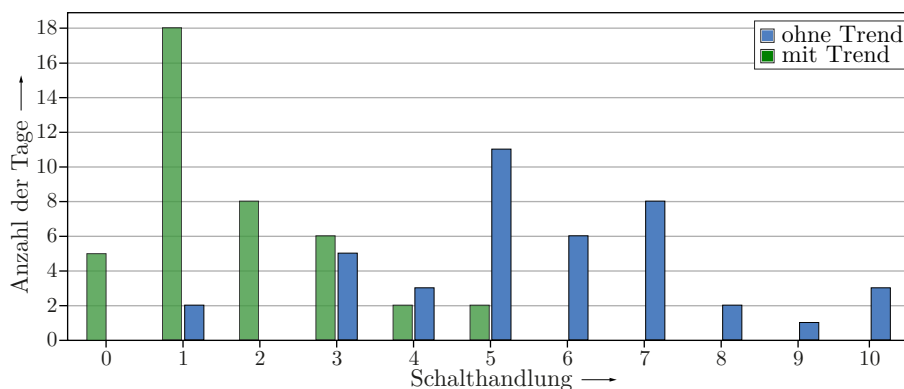


Abbildung 4.10.: Vergleich der Anzahl der Schalthandlungen je Tag der 41 Tage des Validierungsdatensatzes der Agenten A1 (grün) und A2 (blau).

4.3.4. Resümee

Das beschriebene Untersuchungsszenario einer Weitbereichsregelung im elektr. Niederspannungsnetz zeigt einen möglichen Ansatz, einen Algorithmus des ML die Aufgabenstellung selbstständig zu erarbeiten und eine Kontrollstrategie zur Lösung der Weitbereichsregelung abzuleiten.

Der vorgestellte Vergleich im Zuge der Verifizierung der Kontrollstrategie führt zu dem Schluss, dass der beschriebene Aufbau sowie die sorgfältig definierten Signale in einem Ansatz resultieren, welcher mit Hilfe eines rONT eine Weitbereichsregelung im elektr. Niederspannungsnetz erfolgreich adressiert. Hierbei konnte das vorrangige Ziel, die Schaltheufigkeit des Stufenschalters zu reduzieren, erfolgreich umgesetzt werden. Zudem führen die gewählte Beschreibung und Definition der Problemstellung zu einer allgemeinen Lösung, die eine beliebige Anzahl von gemessenen Knotenpunkten berücksichtigen kann.

Grundlegend können dem beschriebenen Untersuchungsszenario der vorliegenden Arbeit und dem Einsatz des Rahmenwerks des RL folgende Punkte zugeschrieben werden:

- Das Rahmenwerk bietet eine simple Struktur zur Beschreibung von Problemstellungen mit Hilfe weniger Signale (vgl. Abschnitt 4.3.1).
- Eine schnelle und einfache Ergänzung der Zieldefinition ist durch die additive Struktur der Belohnungsfunktion gegeben.
- Die Integration bestehender Simulationsumgebungen in das Rahmenwerk des RL ist durch klare Schnittstellen schnell umgesetzt.

Darüber hinaus bietet der beschriebene Ansatz die Möglichkeit, exogene sowie endogene Unsicherheiten mit Hilfe der vorgestellten Definition unterschiedlicher Szenarien direkt in die Erarbeitung der Kontrollstrategie, innerhalb der *Lernphase*, zu berücksichtigen. Das beschriebene Anreichern der Szenarien des *Lernmoduls* mit Observationen aus dem *realen System* bietet zusätzlich die Möglichkeit weitere Systemzustände dem Agenten zuzuführen, welche ggf. von der Simulationsumgebung nicht wiedergegeben werden. Diese Optionen resultieren in der Ableitung einer robusten Kontrollstrategie, welche die Gegebenheiten des *realen Systems* bestmöglich widerspiegelt.

Zusammenfassung und Ausblick

Zusammenfassung

Vor dem Hintergrund des Transformationsprozesses des elektr. Energieversorgungssystems im Zuge der integrierten Energiewende adressiert die vorliegende Arbeit eine aktiv herbeigeführte Spannungsanpassung als Flexibilitätsoption im intelligenten Netzbetrieb. Bei einer stationären Betrachtung der Vorgänge im elektr. Niederspannungsnetz wird eine Netzführungsstrategie zur Weitbereichsregelung unter Zuhilfenahme des bestärkenden Lernens vorgestellt. Basierend auf Messdaten an neuralgischen Netzknoten im Niederspannungsnetz wird ein zentrales, übergeordnetes Regelkonzept zur Steuerung des regelbaren Ortsnetztransformators erarbeitet. Ziel der Regelung ist die elektr. Spannung innerhalb der Spannungsbänder, vorgegeben von der Europäischen Norm DIN EN 50160, zu halten. Es wird ein praxisorientierter Ansatz verfolgt, welcher auf Basis bestehender Akteure, Infrastrukturdaten sowie erhobener Messdaten beruht. Maßgeblich für diese Bewertung und Datenquelle der vorliegenden Arbeit ist das Forschungsvorhaben *Erfassung der niederspannungsseitigen Netzzustandsgrößen in Echtzeit* (ENERGIE) und die hieraus gewonnenen Erkenntnisse.

Das genutzte Rahmenwerk zur Ableitung der Kontrollstrategie mit Hilfe des bestärkenden Lernens sieht für Problemstellungen im elektr. Energieversorgungssystem den Einsatz einer Simulationsumgebung vor. Diese muss im Kontext der zu lösenden Problemstellung die Prozesse im Zuge der Netzsimulation möglichst realitätsnah wiedergeben, wodurch eine entsprechende Modellierungstiefe der Simulationsumgebung begründet ist. Da die vorliegende Arbeit eine Spannungsanpassung im Niederspannungsnetz adressiert, ist die Modellierung spannungsabhängiger Last- und Einspeiseprozesse unabdingbar. Darüber hinaus berücksichtigen die simulativen Untersuchungen explizit ein energiekonstantes Verbraucherverhalten und den daraus resultierenden Kompensationseffekt.

Da Photovoltaikanlagen in den Referenznetzen des Forschungsvorhabens ENERGIE eine vergleichsweise hohe Durchdringung sowie wirksame elektr. Leistung aufweisen, wird der Fokus auf die Modellierung des Einspeiseprozesses mittels Photovoltaikanlagen gelegt. Demnach erfolgt in der vorliegenden Arbeit die Berücksichtigung dezentraler Erzeugungsanlagen zur Generierung synthetischer Einspeiseprofile ausschließlich mit Hilfe von Photovoltaikanlagen. Das beschriebene Modell wird wiederum mit Messdaten, aufgenommen im Forschungsvorhaben ENERGIE, validiert. Das Einspeisemodell wird im Wesentlichen durch eine Wirkungskette beschrieben, die detaillierte Anlageninformationen berücksich-

tigt. Das synthetische Einspeiseprofil variiert in Abhängigkeit des Aufstellungsorts, der Ausrichtung und Neigung sowie der Modultemperatur. Zur einfachen Berücksichtigung einer stark wechselnden Bewölkung, einer Wolkenreflexion oder einer systematischen Verschattung wurde die Wirkungskette zur Beschreibung der generatorischen Einspeiseleistung einer Photovoltaikanlage um einen Faktor ergänzt, der einen direkten Einfluss auf das Einspeiseprofil nimmt. Hierdurch werden exogene Unsicherheiten im Rahmen der Netzsimulation unkompliziert berücksichtigt. Zudem ermöglicht der vorgestellte Ansatz die Gegebenheiten vor Ort mit Hilfe einer Messung der globalen Bestrahlungsstärke nachzufahren.

Der Verifizierung des Modells der Photovoltaikanlage ist die Notwendigkeit einer detaillierten Modellierung zur Ableitung und Bewertung von Netzführungs- und Entscheidungsprozessen im elektr. Netz zu entnehmen. So zeigt die Simulationsrechnung unter Anwendung des beschriebenen Ansatzes einen deutlichen Kapazitätsgewinn an übertragbarer Wirkleistung unter Berücksichtigung der detaillierten Datenbasis. Dabei wird im Allgemeinen auf eine detaillierte Modellierung meteorologischer Vorgänge verzichtet. Vielmehr ist davon auszugehen, dass unter Berücksichtigung der Durchdringung der Photovoltaikanlagen im elektr. Mittel- und Niederspannungsnetz eine Einbindung der gemessenen globalen Bestrahlungsstärke, innerhalb oder im direkten Umfeld des jeweiligen Netzgebietes, unabdingbar ist. Die Modellierung der Spannungsabhängigkeit erfolgt mit Hilfe des ZIP-Modells als leistungskonstanter Einspeiseprozess.

Mit Bezug auf die Verbraucherlastmodellierung ist der vorliegenden Arbeit ein Vergleich zwischen verschiedenen Ansätzen zu entnehmen sowie eine Gruppierung in klassische und probabilistische Ansätze. Insbesondere die Notwendigkeit der Beschreibung von Endkunden, speziell Haushalts-Endkunden und deren individuelles, stochastisches Verhalten mit Hilfe eines probabilistischen Ansatzes, wird hervorgehoben. Das vorgestellte Bottom-Up-Verfahren kann von Ansätzen der einschlägigen Literatur dahingehend unterschieden werden, dass das Konzept zur Nachbildung eines Profils mit Hilfe einer definierten *Zeitschablone* erfolgt. Zudem liegt dem vorgestellten Verfahren ein *energietreuer* Ansatz zugrunde.

Bedingt durch eine bessere allgemeine Verfügbarkeit von Koeffizienten einzelner elektr. Geräte findet vorrangig das beschriebene Bottom-Up-Verfahren in Kombination mit einer Modellierung spannungsabhängiger Prozesse mittels ZIP-Modell Anwendung. Die Validierung des Verfahrens mit Hilfe aufgenommener Messdaten im Forschungsvorhaben ENERGIE zeigt, dass die Modellierung die individuelle statistische Systematik einzelner Haushalts-Endkunden mit ausreichender Genauigkeit wiedergibt. Hierdurch wird ermöglicht die erhöhte Komplexität und Dynamik des Niederspannungsnetzes synthetisch nachzubilden. Die Berücksichtigung der Spannungsabhängigkeit mit Hilfe des ZIP-Modells erfolgt durch Zuordnung eines Satzes von ZIP-Koeffizienten zu Gerätegruppen. Darüber hinaus können Gerätegruppen als energiekonstant definiert werden, wodurch gewährleistet wird, dass einzelne Gerätegruppen unabhängig von Spannungsschwankungen einen konstanten Energiebezug aufweisen.

Begründet durch die Tatsache, dass das energiekonstante Verhalten einzelner Verbraucher i.d.R. vernachlässigt wird und eine umfangreiche Bewertung energiekonstanter Prozesse im Niederspannungsnetz der einschlägigen Literatur nicht zu entnehmen ist, wird die Notwendigkeit zur Berücksichtigung des energiekonstanten Verbraucherverhaltens und dem daraus resultierenden Kompensationseffekt im Rahmen der vorliegenden Arbeit untersucht. Hierfür wird im Zuge von zwei umfangreichen simulativen Untersuchungen die Systemantwort für ein passives sowie aktives Netz beleuchtet. Zur Bewertung des Einflusses auf die elektr. Vorgänge im Niederspannungsnetz werden im Rahmen einer Monte-Carlo-Simulation eine Vielzahl von Leistungsflussberechnungen im Zusammenhang einer aktiv herbeigeführten Spannungsanpassung durchgeführt. Hierbei werden die Auswirkungen auf die Spitzenlast, die Energieaufnahme und die Verlustenergie des gesamten Netzes mit Hilfe eines Optimierungsproblems beschrieben und im Kontext potenzieller Flexibilitätsoptionen untersucht. Die wichtigsten wiss. Erkenntnisse sind nachfolgend stichpunktartig aufgeführt:

- Bei der Berücksichtigung der Spannungsanpassung als Flexibilitätsoption ist zwischen einem passiven und aktiven Niederspannungsnetz in Abhängigkeit der zu optimierenden Größe zu unterscheiden.
- Eine multivariate Optimierung mit dem Ziel der Reduzierung der Spitzenlast, der Energieaufnahme sowie der Verlustenergie des gesamten Netzes ist für ein aktives Niederspannungsnetz, mit dem Übergang vom passiven zum aktiven Netz zur Mittagszeit, möglich.
- Der kompensierende Effekt reduziert im Allgemeinen die positiven Auswirkungen einer Spannungsabsenkung auf die Energieaufnahme sowie die Verlustenergie des Netzes.
- Der Kompensationseffekt resultiert in einer Zeitabhängigkeit der Netzführungsstrategie, die insbesondere bei der Spitzenlastreduktion zum Tragen kommt.

Grundsätzlich bietet eine aktiv herbeigeführte Spannungsanpassung das Potenzial, einen Baustein zur Flexibilisierung des Energieversorgungssystems darzustellen und ist als sinnvolle Ergänzung zu sehen. Hierbei ist zu berücksichtigen, dass eine aktiv herbeigeführte Spannungsanpassung in Abhängigkeit der Netztopologie und der jeweiligen Verbraucher- und Erzeugerstruktur eine unterschiedliche Systemantwort hervorruft. Zudem hat die zu optimierende elektr. Größe einen signifikanten Einfluss auf den Zeitpunkt und die Richtung der Spannungsanpassung. Im Allgemeinen zeigen die Ergebnisse der vorliegenden Arbeit, dass die positiven Auswirkungen auf die Reduzierung der Spitzenlast, der Energieaufnahme des Gesamtnetzes sowie der Gesamtverlustenergie im Rahmen einer Netzführungsstrategie geringer ausfallen können. Dies ist in Abhängigkeit der energiekonstanten Prozesse im Netzgebiet zu bewerten und durch den Kompensationseffekt begründet. Hierbei bleibt abschließend festzuhalten, dass das Einbeziehen des energiekonstanten Verbraucherverhaltens zwingend erforderlich ist. So wirkt der beschriebene Kompensationseffekt zwar entgegen der Spannungsänderung, nimmt aber gleichzeitig einen erheblichen Einfluss auf die Prozesse im Niederspannungsnetz. Demnach wird das energiekonstante Verbraucherverhalten innerhalb der Simulationsumgebung explizit berücksichtigt.

Nach erfolgreicher Bewertung und Validierung einzelner Bestandteile der Simulationsumgebung folgt die Ableitung einer Netzführungsstrategie, basierend auf einem Ansatz der künstlichen Intelligenz. Der vorgestellte Ansatz, der mit Hilfe des bestärkenden Lernens eine Netzführungsstrategie ableitet, nutzt die zuvor beschriebene Simulationsumgebung zur Lösung einer Weitbereichsregelung im elektr. Niederspannungsnetz. Hiermit wird eine Spannungsanpassung im intelligenten Netzbetrieb zur Einhaltung der Spannungsbandgrenzen adressiert und bewertet. Mit der vorliegenden Arbeit ist der Einsatz eines Agenten des bestärkenden Lernens zur Haltung der Spannung innerhalb eines vorgegebenen Spannungsbandes, nahe eines definierten Referenzwertes gelungen. Dies erfolgt unter der Prämisse, die Schalthäufigkeit des regelbaren Ortsnetztransformators zu minimieren. Für diesen Zweck werden Spannungsmessungen an unterschiedlichen Netzknoten hinzugezogen und eine zusätzliche Steigung über einen definierten Zeitraum an zurückliegenden Beobachtungen bestimmt. Dieser über ein gleitendes Intervall berechnete Trend findet im Zustandsraum sowie der Belohnungsfunktion Anwendung. Zudem wird eine explizite Berücksichtigung von endogenen und exogenen Unsicherheiten diskutiert und adressiert. Die Definition unterschiedlicher Szenarien zur Überführung der Simulationsumgebung in bestimmte Systemzustände sowie eine Rückkopplungsschleife zur Einbindung realer Messdaten wird vorgestellt und angewandt. Hierdurch wird die Ableitung einer robusten Kontrollstrategie begünstigt. Bei der Wahl des eingesetzten Algorithmus steht eine gute Skalierung und Generalisierung im Vordergrund. Die beschriebene Methodik ermöglicht eine schnelle Adaption auf weitere Problemstellungen. Zu beachten bleibt, dass alle modellfreien, rein datengetriebenen Ansätze eine hohe Anzahl an Episoden während der *Lernphase* benötigen.

Die Verifizierung des vorgestellten Verfahrens erfolgt wiederum mit Hilfe der im Forschungsvorhaben ENERGIE erhobenen Stamm- und Bewegungsdaten. Hierfür leiten zwei identisch definierte Agenten jeweils eine Kontrollstrategie ab. Zu unterscheiden sind die Ansätze lediglich durch einen unterschiedlichen Zustandsraum sowie Belohnungsfunktion innerhalb der *Lernphase*. Das Untersuchungsszenario basiert auf einer realen Netztopologie mit zugehörigen Messdaten an unterschiedlichen Netzknoten. Den Ergebnissen ist zu entnehmen, dass:

- beide Agenten selbstständig eine Kontrollstrategie erarbeiten.
- die unterschiedlich definierten Zustandsräume und Belohnungsfunktionen für die beiden Agenten jeweils in einer Kontrollstrategie resultieren, wodurch die Referenzspannung innerhalb des vorgegebenen Spannungsbandes, nahe eines Referenzwertes, gehalten wird.
- ein Agent wesentlich träger agiert und seltener Schalthandlungen durchführt.
- die Anpassung einer resultierenden Kontrollstrategie durch eine geringfügige Anpassung der Signale gegeben ist.

Der Vergleich beider Agenten hebt hervor, dass das vorgestellte Verfahren zur Erarbeitung einer Netzführungsstrategie erfolgreich umgesetzt ist. Beide Agenten sind in der Lage eine erfolgreiche Kontrollstrategie im Kontext der Problemstellung zu erarbeiten. Zudem zeigt

die Untersuchung, dass die erstmals in der vorliegenden Arbeit vorgestellte Herangehensweise die Schalthäufigkeit des regelbaren Ortsnetztransformators reduziert. Dies bestärkt die Annahme einer schnellen Anpassung des Verfahrens.

Mit Hilfe der simulativen Untersuchung verdeutlicht die vorliegende Arbeit das Potenzial des bestärkenden Lernens mit Bezug auf einen intelligenten Netzbetrieb. Anhand der vorgestellten Lösung einer Weitbereichsregelung im elektr. Niederspannungsnetz konnte erfolgreich die Ableitung einer Netzführungsstrategie aufgezeigt werden, welche eine Spannungsanpassung als Handlungsoption im intelligenten Netzbetrieb ermöglicht. Die Zuhilfenahme des maschinellen Lernens ermöglicht hierbei eine sinnvolle, maschinelle Berücksichtigung und Verarbeitung des kontinuierlich zunehmenden Datenstroms und bietet eine sinnvolle Alternative zu bestehenden Ansätzen.

Ausblick

Die Notwendigkeit die Flexibilität des elektr. Energieversorgungssystems zu erhöhen und sektorenübergreifend netzdienliche Vorgänge zu etablieren, ist Ziel der integrierten Energie- wende. Wie der vorliegenden Arbeit zu entnehmen ist, arbeiten Verbände und Gesetzgeber an einem normativen und gesetzlichen Rahmen, der die Einbindung weiterer Akteure ermöglicht und zugleich verpflichtend macht. Demnach kann davon ausgegangen werden, dass das zukünftige Energieversorgungssystem in seiner Komplexität stetig steigt. Eine weitere Zunahme dezentraler Erzeugungsanlagen, des allgemeinen Elektrifizierungsgrades und neuer Verbrauchertypen sowie der verstärkte Einsatz von Messsystemen und eine Erweiterung sowie Verbesserung der vorhandenen Informations- und Kommunikationstechnik über alle Spannungsebenen hinweg, schließen die Transformation des Energieversorgungssystems ab. Diesem Anstieg an Komplexität des Gesamtsystems muss mit neuen Lösungsansätzen entgegengewirkt werden. Aufbauend hierauf liefert die vorliegende Arbeit einen Beitrag zur Bewertung und Umsetzung neuer Netzführungsstrategien. Hierbei bieten modellfreie, rein datengetriebene Ansätze eine interessante Alternative zu klassischen Ansätzen. Insbesondere vor dem Hintergrund der beschriebenen, stetig zunehmend Datenbasis.

Aufbauend auf diesen Erkenntnissen bedarf es weiterer Untersuchungen, mit dem Ziel, einer Spannungsebenen übergreifenden, ganzheitlichen Untersuchung. Bestehende Simulationswerkzeuge können schnell für den Einsatz mit dem beschriebenen Rahmenwerk des bestärkenden Lernens verknüpft werden. Der in der vorliegenden Arbeit adressierte Normalbetrieb des elektr. Netzes kann um Szenarien zur Erkennung von Störfällen im Netzgebiet erweitert werden. Zudem können Netzführungsstrategien zur Spitzenlastreduktion oder Reduzierung der Netzverluste mit Hilfe des bestärkenden Lernens adressiert werden. Der auf einem kontinuierlichen Aktionsraum basierende Ansatz ist in der Lage, weitere Akteure im Netz in den Entscheidungsprozess zu integrieren. Hierbei ist keine Anpassung des Verfahrens notwendig, lediglich die Architektur und ggf. die definierten Hyperparameter sind anzupassen. Dies ermöglicht in nachfolgenden Arbeiten die Integration weiterer Flexibilitäten.

Auch wenn den Verfahren der künstlichen Intelligenz ein vielversprechendes Potenzial zugesprochen wird [44], so stehen diese Ansätze auch zunehmend in der Kritik. Insbesondere im Zusammenhang mit Transparenz, Diskriminierung und ethischen Bedenken. Auch der Einsatz in Systemen der kritischen Infrastruktur findet Platz im gesellschaftlichen Diskurs. Zudem wird der ökologische Fußabdruck der Technologie thematisiert [134]. Der zunehmende Einsatz von Großrechnern, die Teils zum Berechnen genauerer Klimaprognosen oder der effizienten Orchestrierung von Akteuren im elektrischen Netz notwendig sind, benötigen wiederum Energie. Der hieraus resultierende CO₂-Fußabdruck, welcher mit der Bedeutung der künstlichen Intelligenz kontinuierlich zunimmt, sollte kritisch bewertet werden. In [135] bewerten die Autoren die Entwicklung der Computerlinguistik im Rahmen von Großrechner gestützter Sprachverarbeitung mit Hilfe von tiefen neuronalen Netzen in einem ökonomischen Kontext. Neben den Kosten für die Hardware bewerten die Autoren den Energieverbrauch sowie den CO₂ Ausstoß. Demnach darf der hohe Energieverbrauch dieser Großrechner gestützten Technologien nicht vernachlässigt werden. Daher wird die Entwicklung energieeffizienter Ansätze in Zukunft eine wichtige Rolle spielen.

Dessen ungeachtet zählen Algorithmen und Ansätze der künstlichen Intelligenz nach wie vor zu den Zukunftstreibern über alle Branchen in Deutschland hinweg. Ende 2018 veröffentlichte die Bundesregierung in gemeinsamer Federführung mit den Ministerien für Bildung und Forschung, Wirtschaft und Klimaschutz (damals noch Wirtschaft und Energie) sowie Arbeit und Soziales die Nationale Strategie für künstliche Intelligenz [136]. Unter dem Slogan „AI made in Germany“ bildet die Bundesregierung eine ganzheitliche, bundesweite Strategie für diese Schlüsseltechnologie ab. Im Fokus steht zwar die Wettbewerbsfähigkeit des Standortes Deutschland und Europa, jedoch wird dies mit dem „Nutzen für Mensch und Umwelt“ begleitet um „eine verantwortungsvolle und gemeinwohlorientierte Entwicklung und Nutzung“ von künstlicher Intelligenz zu gewährleisten [136]. So ist es nur naheliegend, dass die Forschungstätigkeiten im Bereich des bestärkenden Lernens im Kontext des Energieversorgungssystems weiterhin zunehmen werden. Zudem ist zu erwarten, dass der Wissenstransfer und die Anwendung des bestärkenden Lernens und weiteren Ansätzen des maschinellen Lernens in konkreten und realen Anwendungen kontinuierlich zunehmen werden.

Knotenmodellierung

A.1. Einfluss des Gleichzeitigkeitsgrades auf die anzunehmende Spitzenlast von Haushalts-Endkunden

Im Zuge von Planungsmaßnahmen und der Bewertung vorhandener Netzkapazitäten erfolgt die Abschätzung der Spitzenlast $p_s(N)$ für die einzelnen, am Abgang eines Niederspannungsnetzes angeschlossenen Haushalts-Endkunden i.d.R. mit Hilfe des Gleichzeitigkeitsgrades nach Gleichung 2.1 des Abschnitts 2.1.1.1. Der Abbildung A.1 ist der Verlauf der Spitzenlast $p_s(N)$ in Abhängigkeit vollelektrifizierter Haushalts-Endkunden N zu entnehmen. Hierdurch soll der abnehmende Einfluss eines einzelnen Haushalts-Endkunden bedingt durch den Gleichzeitigkeitsgrad verdeutlicht werden.

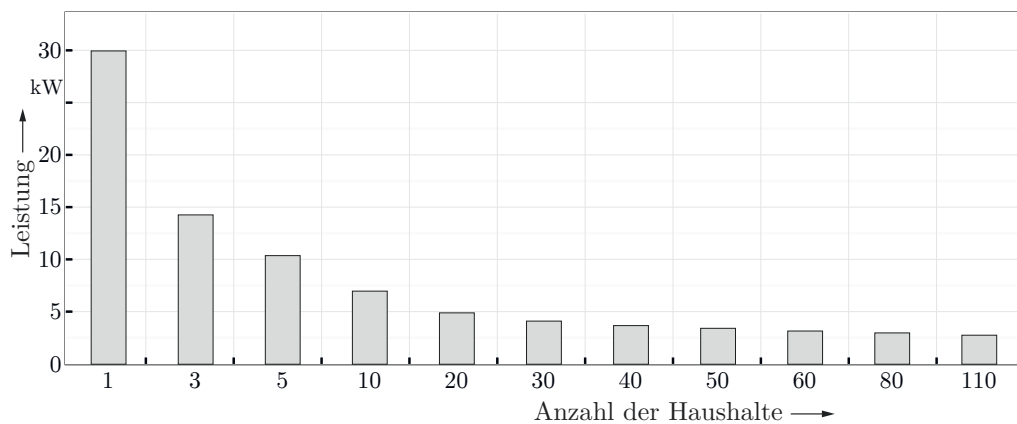


Abbildung A.1.: Einfluss des Gleichzeitigkeitsgrads in Abhängigkeit der Anzahl an Haushalts-Endkunden N auf die anzunehmende Spitzenlast $p_s(N)$ je Haushalts-Endkunden.

A.2. Darstellung des Standardlastprofils nach Profiltyp H0

Darstellung repräsentativer Lastprofile für Haushalts-Endkunden, Privatverbrauch sowie geringfügiger gewerblicher Bedarf (Profiltyp H0, vgl. Tabelle 2.2 des Abschnitts 2.1.1.2) der Jahreszone *Winter*. Der Abbildung A.2 sind exemplarisch die Zeitreihen eines *Werktages*, *Samstags* und *Sonntags* zu entnehmen, welche jeweils auf einen JEB von 1000 kWh/a normiert sind.

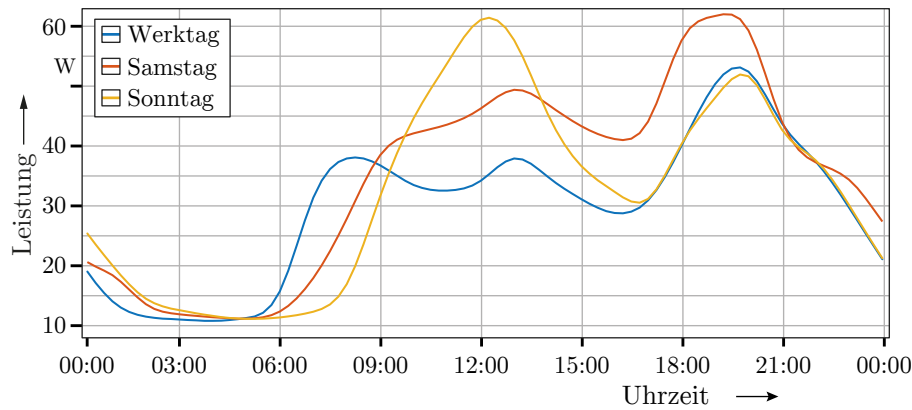


Abbildung A.2.: Verlauf des Standardlastprofils für den Profiltyp H0 (Haushalts-Endkunden, Privatverbrauch sowie geringfügiger gewerblicher Bedarf) eines *Werktages*, *Samstags* und *Sonntags* im *Winter*.

A.3. Definition der Standardabweichung des Top-Down-Verfahrens

Zur Ableitung einer sinnvollen Größenordnung des Streumaßes respektive der für das TDV notwendigen rel. Standardabweichung $\sigma_{\mathcal{P},rel}$ werden die in [76] und [77] erhobenen Messdaten zugrunde gelegt. Die Abbildung A.3 visualisiert einen Ausschnitt $\Delta t = |14:00 - 16:00|$ Uhr der resultierenden rel. Standardabweichung $\sigma_{\mathcal{P},rel}$, basierend auf einer Wirkleistungsmessung von $N \in [10, 50, 150, 300]$ kumulierten Haushalts-Endkunden im Tagesverlauf. Erwartungsgemäß ist den einzelnen Zeitreihen eine kontinuierliche Reduzierung der Volatilität mit steigender Anzahl an kumulierten Haushalts-Endkunden zu entnehmen.

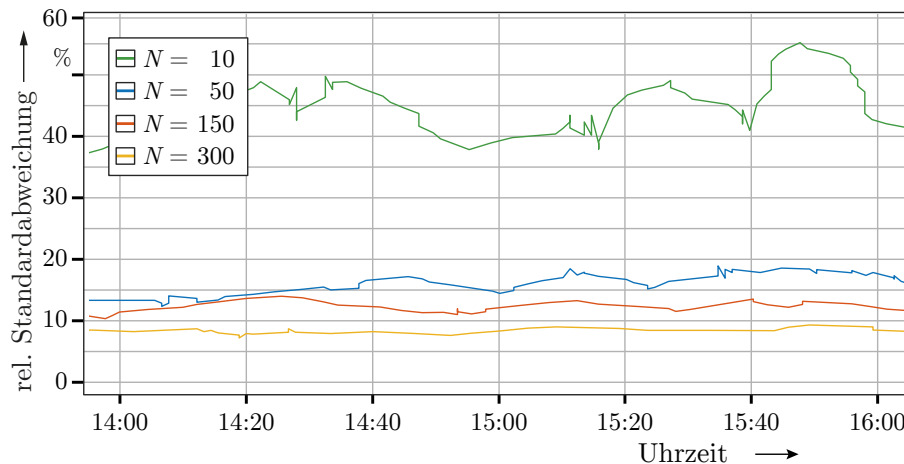


Abbildung A.3.: Ausschnitt $\Delta t = |14:00 - 16:00|$ Uhr der statistischen Auswertung der Messdatenerhebung nach [76] sowie [77]. Zu entnehmen ist die resultierende rel. Standardabweichung $\sigma_{\mathcal{P},rel}$, basierend auf einer Wirkleistungsmessung von $N \in [10, 50, 150, 300]$ kumulierten Haushalts-Endkunden im ländlichen Raum.

Der schematischen Darstellung in Abbildung 2.4 des Abschnitts 2.1.2 ist zu entnehmen, dass das vorgestellte TDV die synthetische Generierung einer Verbraucherlast ermöglicht, wodurch eine beliebige Anzahl an kumulierten Haushalts-Endkunden $N \geq 1$ repräsentiert wird. Hierfür werden die Daten der statistischen Auswertung der Messdatenerhe-

bung nach Abbildung A.3 im Querschnitt bewertet und eine analytische Vorschrift abgeleitet. Der Abbildung A.4 ist exemplarisch eine resultierende Vorschrift über das Intervall $\Delta t = |14:00 - 16:00|$ Uhr zu entnehmen. Basierend auf dem Mittelwert der rel. Standardabweichung $\sigma_{\mathcal{P},\text{rel}}$ des Intervalls Δt werden die diskreten Datenpunkte mit Hilfe einer Interpolationsfunktion im Querschnitt beschrieben. Im Beispiel nach Abbildung A.4 resultiert dies in einer analytischen Vorschrift zur Beschreibung der rel. Standardabweichung $\sigma_{\mathcal{P},\text{rel}}(N) = 147,77N^{-0,489}$ in Abhängigkeit der kumulierten Haushalts-Endkunden N im ländlichen Raum.



Abbildung A.4.: Auswertung der Kennlinienverläufe nach Abbildung A.3 im Querschnitt, exemplarisch für den Mittelwert über das Intervall $\Delta t = |14:00 - 16:00|$ Uhr. Dies resultiert in einer analytischen Vorschrift zur Beschreibung der rel. Standardabweichung $\sigma_{\mathcal{P},\text{rel}}(N) = 147,77N^{-0,489}$ in Abhängigkeit der kumulierten Haushalts-Endkunden N im ländlichen Raum.

A.6. Netztopologie zur Validierung des Bottom-Up-Verfahrens

Die Validierung des vorgestellten BUV erfolgt auf Basis von Stamm- und Bewegungsdaten, aufgenommen im Forschungsvorhaben ENERGIE [36]. Das verwendete Niederspannungsnetz ist am Niederrhein verortet und weist eine ländliche Struktur auf. Der schematischen Darstellung in Abbildung A.5 kann die Netzstruktur entnommen werden. Es sind fünf Abgänge zu entnehmen, wovon Abgang ② als Reserve ausgeführt ist. Die Netztopologie weist Kabel vom Typ NAYY-J 4x150 mm² sowie NAYY-J 4x35 mm² auf.

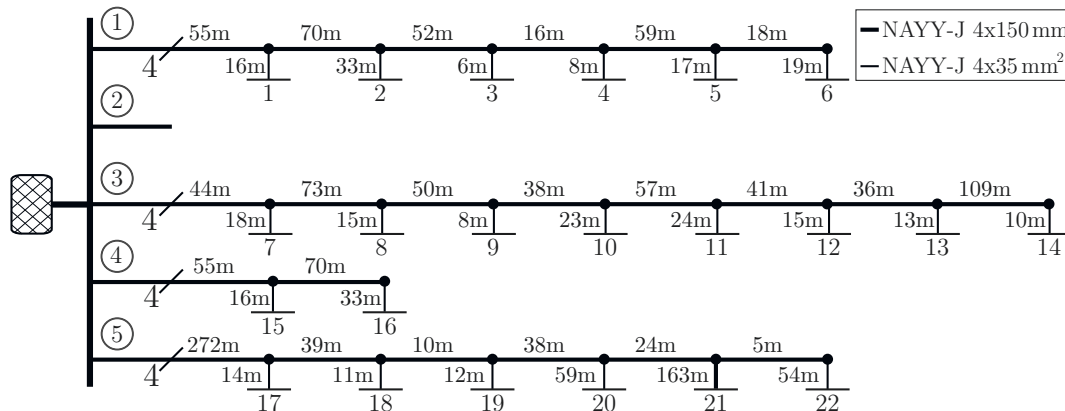


Abbildung A.5.: Topologie eines realen Niederspannungsnetzes. Das schematisch dargestellte elektr. Netz ist am Niederrhein verortet und Teil des Forschungsvorhabens ENERGIE [36]. Kabel vom Typ NAYY-J 4x150 mm² sowie NAYY-J 4x35 mm² wurden verlegt.

Hinzuzufügen ist, dass das Netz über PVA an den Abgängen ①, ③ und ⑤ verfügt. Die Position sowie installierten Leistungen der PVA ist dem Auszug der Datenbank in Tabelle A.1 zu entnehmen. Darüber hinaus ist der JEB der Haushalts-Endkunden aufgelistet.

Tabelle A.1.: Dem Auszug des Datenbankeintrages der Topologie nach Abbildung A.5 sind einzelne Informationen der Netzknoten zu entnehmen. Der kumulierte JEB sowie die installierte Einspeiseleitung der PVA je Netzknoten sind vermerkt.

Netzknoten	Anzahl der Haushalts-Endkunden	W kWh	P_{PVA} kW _p	Abgangs- nummer
1	1	9663		1
2	4	18766		1
3			24	1
4			29	1
5	2	16938		1
6	1	8120		1
7	2	10015		3
8	1	7721		3
9	2	7923		3
10	1	5684		3
11	8	34435		3

12			29	3
			30	3
13	1	3545		3
14	8	36524	29	3
15	1	3823		4
16	2	20291		4
17	1	3601		5
18	1	5252		5
19	2	2604		5
20	1	5229		5
21			30	5
			45	5
			54	5
22	8	36948		5

A.7. Niederspannungsnetz mit hoher Durchdringung an Photovoltaikanlagen

Der Abbildung A.6 ist eines der im Rahmen des Forschungsvorhabens ENERGIE [36] untersuchten Niederspannungsnetze zu entnehmen. Das am Niederrhein verortete elektr. Netz weist eine ländliche Struktur auf und besteht aus zwei Abgängen mit einer Gesamtlänge $l = 3515$ m. Nach Modernisierungsmaßnahmen sind im Niederspannungsnetz 72% der Kabel vom Typ NAYY-J 4x150 mm² verbaut (vgl. Abbildung 3.6 (b) des Abschnitts 3.2.2). Die Netztopologie verfügt über einen rONT mit einer Scheinleistung von $S_{r, \text{rONT}} = 400$ kVA, 114 Haushalte sowie 24 PVA mit einer projektierten, generatorischen Einspeiseleistung von $P_{\text{PVA}} = 398$ kW_p. Zudem sind die Positionen der sechs Messsysteme hervorgehoben.

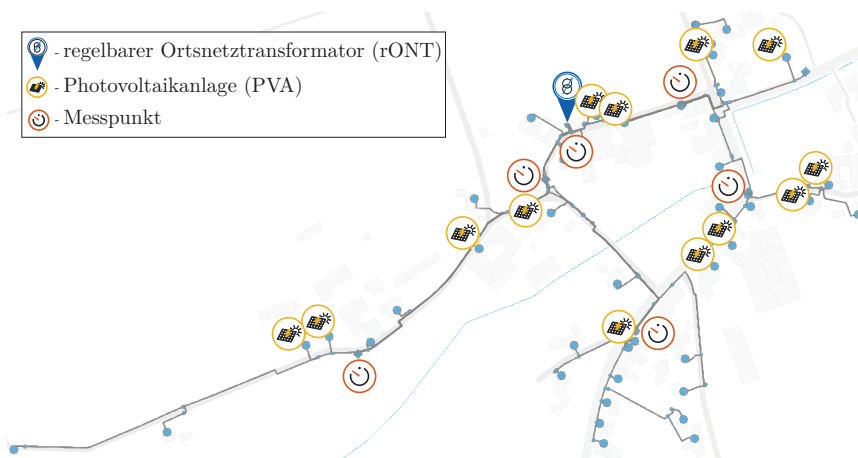


Abbildung A.6.: Dargestellt ist ein reales Niederspannungsnetz (Bestandteil des Forschungsvorhabens nach [36]). Die Positionen der einzelnen PVA, des rONT sowie der Messpunkte sind hervorgehoben.

Komponenten des bestärkenden Lernens

B.1. Monte-Carlo-Vorhersage beim erstmaligen Besuch

Die Monte-Carlo-Vorhersage stellt einen essentiellen Schritt des EVK dar und ist den Monte-Carlo-Methoden zuzuordnen. Wie Abschnitt 4.1.4.2 zu entnehmen ist, werden während der *Lernphase* eine Vielzahl an Zustand-Aktion-Tupel (s, a) besucht. Hierauf aufbauend wird eine Q-Tabelle mit Hilfe der Wertfunktion des Zustand-Aktion-Tupels $Q_\pi(s, a)$ nach Gleichung 4.13 des Kapitels 4 abgeleitet. Hierbei kann es vorkommen, dass während des numerischen Experiments die Zustand-Aktion-Tupel (s, a) innerhalb einer Episode mehrfach besucht werden. Einen möglichen Umgang mit dieser Situation stellt die sogenannte Monte-Carlo-Vorhersage beim erstmaligen Besuch dar, dargestellt mit Hilfe des Pseudocodes nach Algorithmus 3. Dem Pseudocode ist zu entnehmen, dass innerhalb einer Episode nur der erste Besuch eines Zustand-Aktion-Tupels (s, a) in der Q-Tabelle berücksichtigt wird (vgl. achte Zeilennummer).

Algorithmus 3: Monte-Carlo-Vorhersage beim erstmaligen Besuch [116].

Eingabe: Kontrollstrategie π

Eingabe: Anzahl der Episoden $num_episodes$

Ausgabe: Q-Tabelle Q

```

Beginn 1
  // Initialisierungsprozess 2
  Initialize  $N(s, a) = 0$  for all  $s \in \mathcal{S}, a \in \mathcal{A}(s)$  3
  Initialize  $returns\_sum(s, a) = 0$  for all  $s \in \mathcal{S}, a \in \mathcal{A}(s)$  4
  for  $i \leftarrow 1$  to  $num\_episodes$  do 5
    Generate an episode  $S_0, A_0, R_1, \dots, S_T$  using  $\pi$  6
    for  $t \leftarrow 0$  to  $T - 1$  do 7
      if  $(S_t, A_t)$  is a first visit (with return  $G_t$ ) then 8
        // Berechne die kumulierte Belohnung bei Erstbesuch 9
         $N(S_t, A_t) \leftarrow N(S_t, A_t) + 1$  10
         $returns\_sum(S_t, A_t) \leftarrow returns\_sum(S_t, A_t) + G_t$  11
      end 12
    end 13
  end 14
   $Q(s, a) \leftarrow returns\_sum(s, a) / N(s, a)$  for all  $s \in \mathcal{S}, a \in \mathcal{A}(s)$  15
  return Q 16
Ende 17

```

B.2. Optimierungsansätze der Monte-Carlo-Methoden

Inkrementeller Mittelwert

Im Ursprung sehen die Ansätze der Monte-Carlo-Methoden vor, dass eine Vielzahl von Episoden durchlaufen werden, bis bspw. die Schätzung der Q-Tabelle konvergiert. Anschließend erfolgt ein Vergleich der abgeleiteten Kontrollstrategien $\pi' \geq \pi$ für alle π hinsichtlich einer höherwertigen Kontrollstrategie π' . Die Modifikation, welche hinter dem *inkrementellen Mittelwert* (aus dem engl. incremental mean) steckt, ist durch das Aktualisieren der Q-Tabelle nach jeder Episode im Evaluierungsschritt des EVK beschrieben. Durch

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \frac{1}{N(S_t, A_t)}(G_t - Q(S_t, A_t)) \quad (\text{B.1})$$

kann dieses kontinuierliche Aktualisieren der Q-Tabelle gewährleistet werden. Hierbei wird die Differenz $\delta_t = (G_t - Q(S_t, A_t))$ als Fehler- bzw. Differenzterm interpretiert [116]. In Worten beschreibt δ_t die Abweichung zwischen dem erwarteten, modellierten Wert $Q(S_t, A_t)$ und jenem Wert G_t , welchen die Umgebung im Zuge der Systemantwort zurückgibt. Hierdurch sind 3 Optionen gegeben:

- $\delta_t > 0$: Die Schätzung für den Wert $Q(S_t, A_t)$ positiv um $\frac{1}{N(S_t, A_t)}$ anpassen.
- $\delta_t < 0$: Die Schätzung für den Wert $Q(S_t, A_t)$ negativ um $\frac{1}{N(S_t, A_t)}$ anpassen.
- $\delta_t = 0$: Keine Anpassung nötig. Der Wert entspricht der Erwartung.

Der Faktor $\frac{1}{N(S_t, A_t)}$ beschreibt wie häufig das Zustand-Aktion-Tupel (s, a) besucht wurde. Demnach fallen die ersten Besuche des Zustand-Aktion-Tupels (s, a) stärker ins Gewicht. Mit Fortschreiten des Experimentes nimmt der Einfluss kontinuierlich ab.

Konstante-Alpha-Methode

Eine weitere Optimierungsoption im Evaluierungsschritt des EVK ist in der einschlägigen Literatur durch die *Konstante-Alpha-Methode* (aus dem engl. constant alpha) beschrieben [116]. Um den Freiheitsgrad zu erhöhen wird die Division $\frac{1}{N(S_t, A_t)}$ des oben beschriebenen inkrementellen Mittelwerts nach Gleichung B.1 durch eine konstante Schrittweite α ersetzt. Gleichung B.1 wird folglich zu

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha(G_t - Q(S_t, A_t)) \quad (\text{B.2})$$

umgeschrieben. Durch Ausklammern von $Q(S_t, A_t)$ kann die Schreibweise in

$$Q(S_t, A_t) \leftarrow (1 - \alpha)Q(S_t, A_t) + \alpha G_t \quad (\text{B.3})$$

überführt werden. Hierdurch ist der Einfluss der Gleichung B.3 des Hyperparameters $\alpha \in [0, 1]$ deutlich zu entnehmen. Mit den Randwerten

- $\alpha = 1$ gilt $Q(S_t, A_t) \leftarrow G_t$: Die Schätzung des Gewinns G_t zum Zeitpunkt t wird als Wert angenommen. Demnach findet nur die aktuelle Systemantwort Berücksichtigung.
- $\alpha = 0$ gilt $Q(S_t, A_t) \leftarrow Q(S_t, A_t)$: Die Systemantwort zum Zeitpunkt t wird nicht berücksichtigt. Lediglich die zuvor gesammelten Informationen nehmen Einfluss auf die Kontrollstrategie π .

Den beiden beschriebenen Randwerten des Hyperparameters α ist zu entnehmen, dass α eine Gewichtung zwischen der aktuellen und vergangenen Schätzung ermöglicht [116]. Allgemein kann der Hyperparameter α als Lernrate bzw. Schrittweite interpretiert werden [122]. Kleine Werte für α gewährleisten, dass weit zurückliegende Informationen des Experiments Berücksichtigung finden. Mit einem zunehmenden Zahlenwert für α wird der Fokus kontinuierlich auf die aktuelle Systemantwort und den aktuellen Schätzwert des Gewinns G_t gelegt.

Die Ergänzung durch die Konstante-Alpha-Methode in Verbindung mit dem inkrementellen Mittelwert erhöhen den Freiheitsgrad der MCM. Insbesondere mit Bezug auf ein schnelles Auffinden einer optimalen Kontrollstrategie π_{opt} . Ein adäquater Wert für α ist hierbei stark von der Problemstellung abhängig und kann nur empirisch ermittelt oder geschätzt werden. Grundsätzlich gilt, dass ein hoher Wert für α unter Umständen in einer nicht optimalen Kontrollstrategie $\pi \neq \pi_{\text{opt}}$ resultiert [116].

B.3. Algorithmus des *Deep Q-Network*

Dem Algorithmus 4 sind die einzelnen Schritte der *Lernphase* des tiefen Q-Netzwerks nach [47] zu entnehmen. Der Einsatz der ϵ -greedy Kontrollstrategie zur Adressierung des *Exploration-Exploitation*-Dilemmas sowie der des Erfahrungs-Wiedergabepuffers sind dem Pseudocode zu entnehmen. Das Parametrieren des Zielnetzwerks erfolgt zyklisch mit Hilfe eines *harten Aktualisierens* (vgl. Abschnitt 4.1.4.2).

Algorithmus 4: *Lernprozess* des Deep Q-Network nach [47].

Eingabe: Anzahl der Episoden $num_episodes$
Eingabe: Anzahl der Zeitschritte \mathcal{T} je Episode
Eingabe: Speichergröße M des Erfahrungs-Wiedergabepuffer (ERB)
Eingabe: Anzahl der Elemente B des Stapels X
Eingabe: Parameter Update alle C Schritte

Beginn 1
 // **Initialisierungsprozess** 2
 Initialisiere ERB \mathcal{D} mit Speichergröße M 3
 Initialisiere Wertfunktion des Zustand-Aktion-Tupels Q 4
 mit zufälligen Parametern θ^Q
 Initialisiere Wertfunktion des Zustand-Aktion-Tupels des TN Q' 5
 mit $\theta^{Q'} \leftarrow \theta^Q$
for $episode \leftarrow 1$ **to** $num_episodes$ **do** 6
 Initialisiere Umgebung und erhalte initialen Zustand s_1 7
for $t \leftarrow 1$ **to** \mathcal{T} **do** 8
 // // **Adressierung des Exploration-Exploitation-Dilemmas** 9
via ϵ -greedy 10
 Führe mit der Wahrscheinlichkeit von ϵ eine zufällige Aktion a_t aus, 11
 andernfalls wähle Aktion $a_t = \operatorname{argmax}_a Q(s_t, a | \theta^Q)$ 12
 // **Aktion an Umgebung übermitteln** 13
 Die gewählte Aktion a_t ausführen und Belohnung r_t sowie neuen Zustand 14
 s_{t+1} observieren
 // **Tupel des Übergangs im ERB puffern** 15
 Beobachteten Übergang $e_t \leftarrow (s_t, a_t, r_t, s_{t+1})$ in ERB $\mathcal{D}_t \leftarrow e_t$ speichern 16
 // **Aktualisierung der Parameter θ^Q mittels stochastischen Gradienten-** 17
abstiegsverfahrens
 Ziehe einer Gleichverteilung $\mathcal{U}(\mathcal{D})$ folgend zufällig B Elemente aus dem 18
 ERB \mathcal{D} und setze Stapel $X \leftarrow (s_i, a_i, r_i, s_{i+1})$
 Setze $y_i \leftarrow \begin{cases} r_i & \text{Wenn Zeitschritt } i + 1 \notin \mathcal{T} \\ r_i + \gamma \max_{a'} Q'(s_{i+1}, a' | \theta^{Q'}) & \text{andernfalls} \end{cases}$ 19
 Aktualisiere Q durch lösen der Zielfunktion $(y_i - Q(s_i, a_i | \theta^Q))^2$ mit 20
 Bezug auf θ^Q unter Anwendung des SGD
 // **Hartes Update der Parameter alle C Schritte** 21
if t ein Vielfaches von C **then** 22
 | $Q' \leftarrow Q$ 23
end 24
end 25
end 26
Ende 27

B.4. Algorithmus des *Deep Deterministic Policy Gradient*

Der Pseudocode nach Algorithmus 4 veranschaulicht die *Lernphase* des DDPG [49]. Das Parametrieren des Actors und Critics mittels stochastischen Gradientenaufstiegsverfahrens respektive -abstiegsverfahrens ist hervorgehoben. Das Parametrieren der jeweiligen Zielnetzwerke erfolgt zyklisch mit Hilfe eines *weichen Aktualisierens* (vgl. Abschnitt 4.1.4.2).

Algorithmus 5: *Lernprozess* des Deep Deterministic Policy Gradient nach [49].

Eingabe: Anzahl der Episoden $num_episodes$
Eingabe: Anzahl der Zeitschritte \mathcal{T} je Episode
Eingabe: Speichergröße M des Erfahrungs-Wiedergabepuffer (ERB)
Eingabe: Anzahl der Elemente B des Stapels X
Eingabe: Parameter Update alle C Schritte
Eingabe: Schrittweite des *weichen Updates* τ

Beginn 1

```

// Initialisierungsprozess 2
Initialisiere Actor  $Q(s, a|\theta^Q)$  und Critic Netzwerk  $\mu(s|\theta^\mu)$  mit zufälligen 3
  Parametern  $\theta^Q$  sowie  $\theta^\mu$ 
Initialisiere das TN  $Q'$  und  $\mu'$  mit den Parametern  $\theta^{Q'} \leftarrow \theta^Q$  sowie  $\theta^{\mu'} \leftarrow \theta^\mu$  4
Initialisiere ERB  $\mathcal{D}$  mit Speichergröße  $M$  5
for  $episode \leftarrow 1$  to  $num\_episodes$  do 6
  | Initialisiere einen Rauschprozess  $\mathcal{W}$  7
  | Initialisiere Umgebung und erhalte initialen Zustand  $s_1$  8
  | for  $t \leftarrow 1$  to  $\mathcal{T}$  do 9
  | | // Adressierung des Exploration-Exploitation-Dilemmas 10
  | | Wähle eine durch  $\mathcal{W}$  verrauschte Aktion  $a_t = \mu(s_t|\theta^\mu) + \omega_t$  der aktuellen 11
  | | Kontrollstrategie  $\mu$  folgend
  | | // Aktion an Umgebung übermitteln 12
  | | Die gewählte Aktion  $a_t$  ausführen und Belohnung  $r_t$  sowie neuen Zustand 13
  | |  $s_{t+1}$  observieren
  | | // Tupel des Übergangs im ERB puffern 14
  | | Beobachteten Übergang  $e_t \leftarrow (s_t, a_t, r_t, s_{t+1})$  in ERB  $\mathcal{D}_t \leftarrow e_t$  speichern 15
  | | // Parametrieren der Netzwerke  $\theta$  16
  | | Ziehe einer Gleichverteilung  $\mathcal{U}(\mathcal{D})$  folgend zufällig  $B$  Elemente aus dem 17
  | | ERB  $\mathcal{D}$  und setze Stapel  $X \leftarrow (s_i, a_i, r_i, s_{i+1})$ 
  | | // Parametrieren des Critics 18
  | | Setze  $y_i \leftarrow r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}|\theta^{\mu'})|\theta^{Q'})$  19
  | | Minimiere Zielfunktion  $\mathcal{L} = \frac{1}{B} \sum_i (y_i - Q(s_i, a_i|\theta^Q))^2$  20
  | | // Parametrieren des Actors 21
  | | Berechne Gradienten  $\nabla_{\theta^\mu} J \approx \frac{1}{B} \sum_i \nabla_a Q(s_i, \mu(s_i)|\theta^Q) \nabla_{\theta^\mu} \mu(s_i|\theta^\mu)$  22
  | | // Weiches aktualisieren der Parameter des Zielnetzwerks alle  $C$  Schritte 23
  | | if  $t$  ein Vielfaches von  $C$  then 24
  | | |  $\theta^{Q'} \leftarrow \tau \theta^Q + (1 - \tau) \theta^{Q'}$  25
  | | |  $\theta^{\mu'} \leftarrow \tau \theta^\mu + (1 - \tau) \theta^{\mu'}$  26
  | | end 27
  | end 28
end 29
Ende 30

```

Publikationen des Autors

- [19] Kopczynski, L.; Schallenburger, M.; Huppertz, P. und Zeise R., „Berücksichtigung spannungsabhängiger Verbraucherleistung als Handlungsoption im intelligenten Netzbetrieb,“ „Considering voltage-dependent loads as course of action for intelligent grid operation,“ in *Germany: VDE-Congress - Internet of Things*, Mannheim, 2016.
- [33] Schallenburger, M.; Kopczynski, L.; Huppertz, P. and Zeise, R., „Optimal Reconfiguration of Real Low-Voltage Grids Based on Probabilistic Simulation.“ *In Proceedings of the Ninth International Conference on Future Energy Systems (e-Energy '18). Association for Computing Machinery*, New York, NY, USA, 2018, 593–598. DOI: 10.1145/3208903.3210277.
- [50] Schallenburger, M.; Kopczynski, L.; Huppertz, P. and Zeise, R., „Acquisition of Low-Voltage Grid States in Real-Time,“ in *2018 3rd International Conference on Smart and Sustainable Technologies (SpliTech)*, Split, 1 – 6, 2018.
- [51] Huppertz, P.; Schallenburger, M.; Kopczynski, L. und Zeise, R., „Simulation von zeitvarianten stochastischen Prozessen im 4-Leiter- Niederspannungsnetz.“ in *Frankfurt am Main: VDE-Kongress 2014 Smart Cities*, 2014.
- [52] Huppertz, P.; Kopczynski, L. und Zeise, R. und Kizilcay, M., „Einsatz der Datenbasis zur effizienteren Planung von Ausbaumaßnahmen in elektrischen Niederspannungsnetzen mit dezentralen Erzeugern auf Basis einer probabilistischen Netzsimulation.“ in *Von Smart Grids zu Smart Markets 2015: Zur zukünftigen Wechselwirkung von elektrischen Netz und Energiemarktaktivitäten; Beitrag der ETG-Fachtagung*. Kassel, März 2015.
- [53] Huppertz, P.; Kopczynski, L. und Zeise, R. und Kizilcay, M., „Stabilisierungseffekte durch den Einsatz von Wärmepumpen kleiner Leistung in elektrischen Niederspannungsnetzen.“ in *Von Smart Grids zu Smart Markets 2015: Zur zukünftigen Wechselwirkung von elektrischen Netz und Energiemarktaktivitäten; Beitrag der ETG-Fachtagung*. Kassel, März 2015.
- [82] Huppertz, P. ; Kopczynski, L.; Zeise, R. and Kizilcay, M., „Approaching the diversity of unbalanced residential load in low-voltage grids by probabilistic load-flow simulation of cross-sectional data,“ *2015 IEEE Eindhoven PowerTech*, Eindhoven (Niederlande), 1 – 6, 2015.

- [107] Kopczynski, L.; Huppertz, P.; Schallenburger, M. and Zeise, R., “Optimal Tap-Operations of a Regulated Distribution Transformer Considering Conservation Voltage Reduction Effects in Active Low-Voltage Grids,” in *2018 Power Systems Computation Conference (PSCC)*, Dublin (Ireland), 1 – 7, 2018, DOI: 10.23919/PSCC.2018.8442803.
- [133] Kopczynski, L.; Wolf, O.; Schallenburger, M.; Zeise, R. and Hirsch, H., *Reinforcement Learning-Based Tap-Operations of a Regulated Distribution Transformer for Autonomous Voltage Control*, International Journal of Simulation: Systems, Science and Technology (**IJSSST**), 2022, 23(1), PP1-8.

Literatur

- [1] BNetzA Bundesnetzagentur, (19.05.2021) *Bundesnetzagentur veröffentlicht Jahresbericht 2020* [Pressemitteilung], URL: https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Allgemeines/Presse/Pressemitteilungen/2021/20210519_Jahresbericht.pdf (abgerufen am 05.12.2021).
- [2] BNetzA Bundesnetzagentur, *Jahresbericht 2020 - Märkte im digitalen Wandel*, April 2021, URL: <https://www.bundesnetzagentur.de/SharedDocs/Mediathek/Jahresberichte/JB2020.pdf> (abgerufen am 05.12.2021).
- [3] BNetzA Bundesnetzagentur, *Bericht zum Zustand und Ausbau der Verteilnetze 2020*, März 2021, URL: https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Sachgebiete/Energie/Unternehmen_Institutionen/NetzentwicklungUndSmartGrid/ZustandAusbauVerteilernetze2020.pdf (abgerufen am 05.12.2021).
- [4] Statista Research Department, (26.05.2023) *Elektromobilität in Deutschland*, zitiert nach de.statista.com, URL: <https://de.statista.com/themen/608/elektromobilitaet/> (abgerufen am 28.06.2023).
- [5] BMWi Bundesministerium für Wirtschaft und Energie, (02.08.2021) *Erstmals rollen eine Million Elektrofahrzeuge auf deutschen Straßen: Mehr als 50 Prozent dieser Elektrofahrzeuge sind rein batteriebetrieben*. [Pressemitteilung], URL: <https://www.bmw.de/Redaktion/DE/Pressemitteilungen/2021/08/20210802-erst-mals-rollen-eine-million-elektrofahrzeuge-auf-deutschen-strassen.html> (abgerufen am 12.12.2021).
- [6] BMWK Bundesministerium für Wirtschaft und Klimaschutz, (05.05.2023) *Photovoltaik-Strategie: Handlungsfelder und Maßnahmen für einen beschleunigten Ausbau der Photovoltaik*, URL: <https://www.bmwk.de/Redaktion/DE/Publikationen/Energie/photovoltaik-strategie-2023.pdf> (abgerufen am 28.06.2023).
- [7] BDEW Bundesverband der Energie- und Wasserwirtschaft e. V. (2021): *Die Energiewende braucht einen PV-Boom. Die Photovoltaik-Strategie des BDEW*, URL: https://www.bdew.de/media/documents/1000_Die_Energiewende_braucht_einen_PV-Boom.pdf (abgerufen am 12.12.2021).
- [8] Steffens, Ph., „Innovative Planungsgrundsätze für ländliche Mittelspannungsnetze,“ Dissertation, Bergische Universität Wuppertal, Neue Energie aus Wuppertal, Band 16, Wuppertal, 2018.

- [9] Mauermann, H. et al., „Verbesserung der Aufnahmefähigkeit und Sicherung der Netzqualität von Verteilnetzen: Abschlussbericht zum Verbundvorhaben Verteilnetz 2020 : Berichtszeitraum: 01.08.2014 bis 28.02.2019, Verbesserung der Aufnahmefähigkeit und Sicherung der Netzqualität von Verteilnetzen,“ in *TU München (Hrsg.)*, Technische Hochschule Nürnberg Georg Simon Ohm, Technische Universität München, A-Eberle GmbH & Co KG, BMZ Batterien-Montage-Zentrum, empuron GmbH, Grass Power Electronics GmbH, IDS GmbH, Infra Fürth Holding GmbH, KACO new energy GmbH, Power Plus Communications AG, München, 2019, URL: <https://www.tib.eu/de/suchen/id/TIBKAT%3A168788191X> (abgerufen am 12.12.2021).
- [10] Wang, Z. and Wang, J., “Review on Implementation and Assessment of Conservation Voltage Reduction,“ in *IEEE Transactions on Power Systems*, 29(3), 1306 – 1315, May 2014, DOI: 10.1109/TPAS.1978.354658.
- [11] dena Deutsche Energie-Agentur (2016): *Roadmap Demand Side Management. Industrielles Lastmanagement für ein zukunftsfähiges Energiesystem*, Berlin, Juni 2016, URL: https://www.dena.de/fileadmin/dena/Dokumente/Pdf/9146_Studie_Roadmap_Demand_Side_Management..pdf (abgerufen am 19.12.2021).
- [12] dena Deutsche Energie-Agentur (2020): *dena-Studie Systemsicherheit 2050*, Berlin, April 2020, URL: https://www.dena.de/fileadmin/dena/Publikationen/PDFs/2020/dena_Systemsicherheit_2050_LANG_WEB.pdf (abgerufen am 19.12.2021).
- [13] “Netzentwicklungsplan Strom 2030, Version 2019,“ in *50Hertz GmbH, TenneT TSO GmbH, TransnetBW GmbH, Amprion GmbH*, URL: https://www.netzentwicklungsplan.de/sites/default/files/paragraphs-files/Umsetzungsbericht_NEP_2030_v2019.pdf (abgerufen am 19.12.2021).
- [14] Eckert, W. (26.12.2021) *Mit Energie zum Atomausstieg: Drei AKW gehen vom Netz*, Tagesschau, URL: <https://www.tagesschau.de/inland/innenpolitik/atomausstieg-133.html> (abgerufen am 28.12.2021).
- [15] dena Deutsche Energie-Agentur (2016): *Beobachtbarkeit und Steuerbarkeit im Energiesystem. Handlungsbedarfsanalyse der dena-Plattform Systemdienstleistungen*, Berlin, Juli 2016, URL: https://www.dena.de/fileadmin/dena/Dokumente/Pdf/9184_Beobachtbarkeit_und_Steuerbarkeit_.pdf (abgerufen am 12.12.2021).
- [16] Preiss, R. F. and Warnock, V. J., “Impact of Voltage Reduction on Energy and Demand,“ in *IEEE Transactions on Power Apparatus and Systems*, PAS-97(5), 1665 – 1671, Sept. 1978, doi: 10.1109/TPAS.1978.354658.
- [17] “Bibliography on load models for power flow and dynamic performance simulation,“ in *IEEE Transactions on Power Systems*, 10(1), 523 – 538, Feb. 1995, doi: 10.1109/59.373979.

- [18] Energieatlas Rheinland-Pfalz (2011): *Smart Country: Netze für die Stromversorgung der Zukunft in Üttfeld*, URL: <https://www.energieatlas.rlp.de/earp/praxisbeispiele/projektsteckbriefe/projekt-steckbriefe/anzeigen/unternehmen/20> (abgerufen am 12.12.2021).
- [19] Kopczynski, L.; Schallenburger, M.; Huppertz, P. und Zeise R.: „Berücksichtigung spannungsabhängiger Verbraucherleistung als Handlungsoption im intelligenten Netzbetrieb,“ „Considering voltage-dependent loads as course of action for intelligent grid operation,“ in *Germany: VDE-Congress - Internet of Things*, Mannheim, 2016.
- [20] Haslbeck, M. et al., „Mehr Netzanschlusskapazität durch regelbare Ortsnetztransformatoren,“ in *etz Elektrotechn. Z.*, 2012(9), 2 – 7, May 2012, URL: <https://www.smart-production.de/etz/news-detailansicht/nsctrl/detail/News/mehr-netzanschlusskapazitaet-durch-regelbare-ortsnetztransformatoren-2013940/> (abgerufen am 02.03.2020).
- [21] Pinney, D., „Costs and Benefits of Conservation VoltageReduction,“ in *National Rural Electric Cooperative Association (NRECA)*, Arlington, Virginia, USA, Rep.DE-OE0000222, Nov. 2013, URL: https://www.energy.gov/sites/prod/files/2016/10/f34/NRECA_DOE_Costs_Benefits_of_CVR_May_2014.pdf (abgerufen am 02.03.2020).
- [22] VDE Verband der Elektrotechnik Elektronik Informationstechnik e.V. (2021): *Energiespeicher für die Energiewende: Speicherungsbedarf und Auswirkungen auf das Übertragungsnetz für Szenarien bis 2050*, Frankfurt am Main, Juni 2012.
- [23] Specht, M. et al., „Speicherung von Bioenergie und erneuerbarem Strom im Erdgasnetz,“ in *Forschungs Verbunds Erneuerbare Energien in Kooperation mit der Agentur für Erneuerbare Energie*, Tagungsband der Jahrestagung, 69 – 78, Berlin, 2009.
- [24] Korotkiewicz, K. et al., „Grid Commander - Vernetzte, teilautarke Steuerung von Mittel- und Niederspannungsnetzen : gemeinsamer Schlussbericht,“ in *Bergische Universität Wuppertal (Hrsg.)*, Bergische Universität Wuppertal , SPIE SAG GmbH - CeGIT Utility Solutions , PHOENIX CONTACT Energy Automation GmbH , Rheinische NETZGesellschaft mbH , Energieversorgung Leverkusen GmbH, Wuppertal, 2019, URL: <https://www.tib.eu/de/suchen/id/TIBKAT%3A1669608859> (abgerufen am 12.12.2021).
- [25] VDE Verband der Elektrotechnik Elektronik Informationstechnik (2014): *Regionale Flexibilitätsmärkte: Marktbasierte Nutzung von regionalen Flexibilitätsoptionen als Baustein zur erfolgreichen Integration von erneuerbaren Energien in die Verteilnetze*, Frankfurt am Main, 2014.
- [26] Specht, M. et al., „Ergebnisse der FNN-Studie zu neuen Verfahren der statischen Spannungshaltung,“ in *Zukünftige Stromnetze für erneuerbare Energien : 2. OTTI Konferenz*, OTTI e.V., Bereich Erneuerbare Energien, 104 – 109, Berlin, Jan. 2015.

- [27] Wintzek, P. et al., „Planungs- und Betriebsgrundsätze für städtische Verteilnetze – Leitfaden zur Ausrichtung der Netze an ihren zukuünftigen Anforderungen,“ in *Zdrallek, M. (Hrsg.), Neue Energie aus Wuppertal*, Band 35, Wuppertal, 2021.
- [28] Glavic M.; Fonteneau, R. and Ernst, D., “Reinforcement Learning for Electric Power System Decision and Control: Past Considerations and Perspectives,“ in *IFAC-PapersOnLine: 20th IFAC World Congress*, 50(1), 6918 – 6927, 2017, DOI: 10.1016/J.IFACOL.2017.08.1217.
- [29] Ekanayake, J.; Liyanage, K.; Wu, J. and Jenkins, N., “Smart Grid: Technology and Applications,“ in *Wiley*, Chichester, 1st Edition, 2012.
- [30] McClellan, S.; Valles, D. and Koutitas, G, “Dynamic Voltage Optimization Based on In-Band Sensors and Machine Learning,“ in *Applied Sciences* 9(14), 2902, 2019, DOI: 10.3390/app9142902.
- [31] Bokhari, A. et al., “Combined Effect of CVR and DG Penetration in the Voltage Profile of Low-Voltage Secondary Distribution Networks,“ in *IEEE Transactions on Power Delivery*, 31(1), 286 – 293, Feb. 2016, DOI: 10.1109/TPWRD.2015.2422308.
- [32] S. Singh, A. K. Thakur and S. P. Singh, “Energy savings in distribution network with smart grid-enabled CVR and distributed generation,“ in *2016 National Power Systems Conference (NPSC)*, Bhubaneswar, 1 – 6, 2016, DOI: 10.1109/NPSC.2016.7858959.
- [33] Schallenburger, M.; Kopczynski, L.; Huppertz, P. and Zeise, R., „Optimal Reconfiguration of Real Low-Voltage Grids Based on Probabilistic Simulation.“ *In Proceedings of the Ninth International Conference on Future Energy Systems (e-Energy '18)*. Association for Computing Machinery, New York, NY, USA, 2018, 593–598. DOI: 10.1145/3208903.3210277.
- [34] Pasha, A. M. et al., “Conservation Voltage Reduction for Autonomous Microgrids Based on V–I Droop Characteristics,“ in *IEEE Transactions on Sustainable Energy*, 8(3), 1076 – 1085, July 2017, DOI: 10.1109/TSTE.2017.2651046.
- [35] ew-Redaktion (29.07.2020) *evm-Gruppe testet Weitbereichsregelung in der Niederspannung*, ew Magazin für die Energiewirtschaft, URL: <https://www.energie.de/ew/news-detailansicht/nsctrl/detail/News/evm-gruppe-testet-weitbereichsregelung-in-der-niederspannung> (abgerufen am 29.12.2021).
- [36] Schild, K. et al., „ENERGIE: Erfassung der niederspannungsseitigen Netzzustandsgrößen in Echtzeit: Schlussbericht: Bewilligungszeitraum 01.09.2014-31.12.2016,“ in *NGN NETZGESELLSCHAFT NIEDERRHEIN MBH (Hrsg.)*, Krefeld, 2017, URL: <https://doi.org/10.2314/GBV:896141616> (abgerufen am 01.10.2021).

- [37] BDEW Bundesverband der Energie- und Wasserwirtschaft e. V. (2019): *Das Messstellenbetriebsgesetz 2016: Anwendungshilfe*, 5. Auflage, Berlin, Okt. 2019 URL: https://www.bdew.de/media/documents/Awh_20191014_MsbG-5Auflage-Kapitel-1-7.pdf (abgerufen am 28.12.2021).
- [38] Carpentier, J., “Contribution to the economic dispatch problem,” in *Bulletin de la Société Française des Électriciens*, 8(3), 431-447, 1962.
- [39] Frank, S.; Steponavice, I. and Rebennack, S., “Optimal power flow: a bibliographic survey I,” in *Energy Systems* 3, 221–258, 2012, DOI: 10.1007/s12667-012-0056-y.
- [40] Robertson, J.; Harrison, G. and Wallace, R., “OPF techniques for real-time active management of distribution networks,” in *2017 IEEE Power & Energy Society General Meeting*, July 2017, DOI: 10.1109/PESGM.2017.8274371.
- [41] Franco, J. F.; Ochoa, L. F. and Romero, R., “AC OPF for Smart Distribution Networks: An Efficient and Robust Quadratic Approach,” in *IEEE Transactions on Smart Grid*, 9(5), 4613 – 4623, Sept. 2018, DOI: 10.1109/TSG.2017.2665559.
- [42] Toubeau, J.-F. et al., “Deep Reinforcement Learning-Based Voltage Control to Deal with Model Uncertainties in Distribution Networks,” in *Energies*, 3928, 13(15), 1996 – 1073, 2020, DOI: 10.3390/en13153928.
- [43] dena Deutsche Energie-Agentur (2020): *Globale Trends der künstlichen Intelligenz und deren Implikationen für die Energiewirtschaft*, Berlin, März 2020, URL: https://www.dena.de/fileadmin/dena/Publikationen/PDFs/2020/dena-ANALYSE_Globale_Trends_der_kuenstlichen_Intelligenz_und_deren_Implikationen_fuer_die_Energiewirtschaft.pdf (abgerufen am 27.12.2021).
- [44] dena Deutsche Energie-Agentur (2019): *Künstliche Intelligenz für die integrierte Energiewende*, Berlin, Sept. 2019, URL: https://www.dena.de/fileadmin/dena/Publikationen/PDFs/2019/dena-ANALYSE_Kuenstliche_Intelligenz_fuer_die_integrierte_Energiewende.pdf (abgerufen am 27.12.2021).
- [45] Custodio, G.; Ochoa, L. F.; Trindade, F. C. L. and Alpcan, T., “Using Q-Learning for OLTC Voltage Regulation in PV-Rich Distribution Networks,” in *2020 International Conference on Smart Grids and Energy Systems (SGES)*, 482 – 487, 2020, DOI: 10.1109/SGES51519.2020.00091.
- [46] Li, C.; Jin, C. and Sharma, R., “Coordination of PV Smart Inverters Using Deep Reinforcement Learning for Grid Voltage Regulation,” in *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, 1930 – 1937, 2019, DOI: 10.1109/ICMLA.2019.00310.
- [47] Mnih, V. et al., “Human-level control through deep reinforcement learning,” in *Nature*, **518**, 529 – 533, Febr. 2015, DOI: 10.1038/NATURE14236.

- [48] Willis, H. L., "Power Distribution Planning Reference Book," in *Marcel Dekker Inc.*, New York, 2nd Edition, 2004.
- [49] Lillicrap, T. P. et al., "Continuous control with deep reinforcement learning," in *arXiv preprint arXiv:1509.02971*, 2015.
- [50] Schallenburger, M.; Kopczynski, L.; Huppertz, P. and Zeise, R., "Acquisition of Low-Voltage Grid States in Real-Time," in *2018 3rd International Conference on Smart and Sustainable Technologies (SpliTech)*, Split, 1 – 6, 2018.
- [51] Huppertz, P.; Schallenburger, M.; Kopczynski, L. und Zeise, R., "Simulation von zeitvarianten stochastischen Prozessen im 4-Leiter- Niederspannungsnetz," in *Frankfurt am Main: VDE-Kongress 2014 Smart Cities*, 2014.
- [52] Huppertz, P.; Kopczynski, L.; Zeise, R. und Kizilcay, M., "Einsatz der Datenbasis zur effizienteren Planung von Ausbaumaßnahmen in elektrischen Niederspannungsnetzen mit dezentralen Erzeugern auf Basis einer probabilistischen Netzsimulation," in *Von Smart Grids zu Smart Markets 2015: Zur zukünftigen Wechselwirkung von elektrischen Netz und Energiemarktaktivitäten* Kassel, März 2015.
- [53] Huppertz, P.; Kopczynski, L.; Zeise, R. und Kizilcay, M., "Stabilisierungseffekte durch den Einsatz von Wärmepumpen kleiner Leistung in elektrischen Niederspannungsnetzen," in *Von Smart Grids zu Smart Markets 2015: Zur zukünftigen Wechselwirkung von elektrischen Netz und Energiemarktaktivitäten* Kassel, März 2015.
- [54] Kopczynski, L., "Modellierung des Leistungsbedarfs in elektrischen Verteilnetzen mit Hilfe synthetisch generierter Zeitreihen," unveröffentlichte Masterthesis, Fachhochschule Düsseldorf, Düsseldorf, Sept. 2014.
- [55] VDEW Vereinigung Deutscher Elektrizitätswerke (1999): *Repräsentative VDEW- Lastprofile*, Frankfurt am Main, 1999, URL: https://www.bdew.de/media/documents/1999_Repraesentative-VDEW-Lastprofile.pdf (abgerufen am 31.12.2021).
- [56] Dickert, J. and Schegner, P., "Residential Load Models for Network Planning Purposes," in *2010 Modern Electric Power Systems*, Wroclaw (Polen), 1 – 6, 2010.
- [57] Esslinger, P. und Witzmann, R., "Entwicklung und Verifikation eines stochastischen Verbraucherlastmodells für Haushalte," in *12. Symposium Energieinnovation*, Graz (Österreich), 2012, URL: https://www.tugraz.at/fileadmin/user_upload/Events/Eninnov2012/files/lf/LF_Esslinger.pdf (abgerufen am 18.06.2023).
- [58] Scheffler, J., "Bestimmung der maximal zulässigen Netzanschlussleistung photovoltaischer Energiewandlungsanlagen in Wohnsiedlungsgebieten," Dissertation, Universität Chemnitz, Chemnitz, 2002, URL: <https://d-nb.info/968781810/34> (abgerufen am 02.01.2022).

- [59] Wieland, Th.; Reiter, M.; Schmautzer, E.; Fickert, L.; Lagler, M. A. und Eberhart, S., "Gleichzeitigkeitsfaktoren in der elektrischen Energieversorgung – Konventioneller und probabilistischer Ansatz," in *Elektrotechnik und Informationstechnik*, 131(8), 249 – 255, 2014, DOI: 10.1007/s00502-014-0259-0.
- [60] UBA Umweltbundesamt (1999): *Klimaschutz durch Minderung von Leerlaufverlusten bei Elektrogeräten*, Berlin, 1999, URL: <https://www.umweltbundesamt.de/sites/default/files/medien/publikation/short/k1789.pdf> (abgerufen am 02.01.2022).
- [61] Quaschnig, V., "Regenerative Energiesysteme: Technologie - Berechnung - Simulation," in *Hanser*, München, 9., aktualisierte und erw. Aufl., 2015.
- [62] Pardatscher, R. et al., "Untersuchung zu den Auswirkungen von Photovoltaikeinspeisung auf das Nieder- und Mittelspannungsnetz," in *Internationaler ETG-Kongress 2011* Würzburg, Nov. 2011.
- [63] Wiest, M. und Finkel, M., "Lastflussberechnung unter Verwendung statistischer Methoden zur Nachbildung probabilistischer Last- und Einspeiseprofile," in *VDE Kongress 2012*, Stuttgart, 2012.
- [64] VDEW Vereinigung Deutscher Elektrizitätswerke (2000): *Anwendung der Repräsentativen VDEW-Lastprofile - step-by-step*, Frankfurt am Main, Jan. 2000, URL: https://www.bdew.de/media/documents/2000131_Anwendungsrepraesentativen_Lastprofile-Step-by-step.pdf (abgerufen am 02.01.2022).
- [65] VDEW Vereinigung Deutscher Elektrizitätswerke (2000): *Umsetzung des Analytischen Lastprofilverfahren - Step-by-step*, Frankfurt am Main, Juni 2000, URL: <https://www.bdew.de/media/documents/Umsetzung-der-Analytischen-Lastprofilverfahren-Step-by-step.pdf> (abgerufen am 02.01.2022).
- [66] EP Europäisches Parlament (2021): "Energiebinnenmarkt: Kurzdarstellungen über die Europäische Union - 2021," Okt. 2021, URL: https://www.europarl.europa.eu/ftu/pdf/de/FTU_2.1.9.pdf (abgerufen am 02.01.2022).
- [67] VDEW Vereinigung Deutscher Elektrizitätswerke (1985): *Ermittlung der Lastganglinien bei der Benutzung elektrischer Energie durch die bundesdeutschen Haushalte während eines Jahres*, Frankfurt am Main, 1985.
- [68] VDEW Vereinigung Deutscher Elektrizitätswerke (2000): *Zuordnung der VDEW-Lastprofile zum Kundengruppenschlüssel*, Frankfurt am Main, Aug. 2000, URL: https://www.bdew.de/media/documents/Zuordnung_der_VDEW-Lastprofile_zum_Kundengruppenschluessel.pdf (abgerufen am 02.01.2022).
- [69] Dörnemann, Ch. and Handschin, E., "Betriebsmittelbezogene Lastmodellierung in Verteilungsnetzen," in *Springer - Archiv f. Elektrotechnik*, 76, 33 – 40, April 1992, DOI: 10.1007/BF01451984.

- [70] Hinterstocker, M.; von Roon, S. und Rau, M., „Bewertung der aktuellen Standardlastprofile Österreichs und analyse zukünftiger Anpassungsmöglichkeiten im Strommarkt,“ in *13. Symposium Energieinnovation*, Graz (Österreich), 2014.
- [71] Kerber, G., „Aufnahmefähigkeit von Niederspannungsverteilnetzen für die Einspeisung aus Photovoltaikkleinanlagen,“ Dissertation, TU München, München, März 2011.
- [72] Dickert, J., „Synthese von Zeitreihen elektrischer Lasten basierend auf technischen und sozialen Kennzahlen - Grundlage für Planung, Betrieb und Simulation von aktiven Verteilungsnetzen,“ Dissertation, Technische Universität Dresden, Dresden, Okt. 2015.
- [73] Karner, Th., „Synthetische Haushaltslastprofile,“ Bachelorarbeit, Technische Universität Wien, Wien (Österreich), Nov. 2014.
- [74] Piller, W., „Der elektrische Leistungsbedarf von Haushalten,“ in *Der Leistungsbedarf und seine Deckung. Analysen und Strategien. VDI/VDE/GFPE-Tagung*, Schliersee, 1979.
- [75] Wieland, Th., „Eine neuartige probabilistische Methode zur Betriebsmitteldimensionierung in aktiven urbanen Niederspannungsnetzen,“ Dissertation, Technischen Universität Graz, Graz (Österreich), Dez. 2016.
- [76] Cagni, A.; Carpaneto, E.; Chicco, G. and Napoli, R., „Characterisation of the aggregated load patterns for extraurban residential customer groups,“ in *Proceedings of the 12th IEEE Mediterranean Electrotechnical Conference (IEEE Cat. No.04CH37521)*, Dubrovnik (Kroatien), 951 – 954, Mai 2004, DOI: 10.1109/MELCON.2004.1348210.
- [77] Carpaneto, E. and Chicco, G., „Probability distributions of the aggregated residential load,“ in *9th International Conference on Probabilistic Methods Applied to Power Systems*, Stockholm (Schweden), 1 – 6, June 2014, DOI: 10.1109/P-MAPS.2006.360235.
- [78] Papula, L., „Mathematik für Ingenieure und Naturwissenschaftler Band 3: Vektoranalysis, Wahrscheinlichkeitsrechnung, Mathematische Statistik, Fehler- und Ausgleichsrechnung,“ in *Vieweg + Teubner*, Wiesbaden, Band 3, 6. Aufl, 2011.
- [79] Dickert, J. and Schegner, P., „Neue Ansätze der Modellierung synthetischer Lastgänge für Planung und Betrieb von Smart Grids,“ in *Internationaler ETG-Kongress 2011*, Würzburg, Nov. 2011.
- [80] Capasso, A.; Grattieri W.; Lamedica, R. and Prudenzi, A., „A bottom-up approach to residential load modeling,“ in *IEEE Transactions on Power Systems*, 9(2), 957 – 964, May 1994, DOI: 10.1109/59.317650.
- [81] Walker, C. F. and Pokoski J. L., „Residential Load Shape Modeling Based on Customer Behavior,“ in *IEEE Power Engineering Review*, PER-5(7), 34 – 34, July 1985, DOI: 10.1109/MPER.1985.5528462.

- [82] Huppertz, P. ; Kopczynski, L.; Zeise, R. and Kizilcay, M., “Approaching the diversity of unbalanced residential load in low-voltage grids by probabilistic load-flow simulation of cross-sectional data,” *2015 IEEE Eindhoven PowerTech*, Eindhoven (Niederlande), 1 – 6, 2015.
- [83] Dickert, J. and Schegner, P., “A time Series Probabilistic Synthetic load Curve Model for Residential Customers,” in *IEEE Trondheim PowerTech*, Trondheim (Norwegen), 2011.
- [84] Michalik, G. and Mielczarski W., “Modelling of energy use patterns in the residential sector using linguistic variables,” in *Proceedings of International Conference on Intelligent System Application to Power Systems*, Orlando, FL (USA), 278 – 282, 1996, DOI: 10.1109/ISAP.1996.501082.
- [85] BDEW Bundesverband der Energie- und Wasserwirtschaft e. V. (2013): *Energie-Info - Stromverbrauch im Haushalt*. Walker, C. and Pokoski, J., “Residential load shape modeling based on customer behavior,” in *IEEE Transactions on Power Apparatus and Systems*, PAS-104(7), 1703 – 1711, Durham, New Hampshire (USA), July 1985, DOI: 10.1109/TPAS.1985.319202.
- [86] Office for Official Publications of the European Communities (2004): „How Europeans spend their time – Everyday life of women and men,“ 2024, URL: <https://ec.europa.eu/eurostat/documents/3930297/5953614/KS-58-04-998-EN.PDF.pdf/c789a2ce-ed5b-4a0c-bcbf-693e699db7d7> (abgerufen am 08.01.2022).
- [87] Tsagarakis, G.; Collin, A. J. and Kiprakis, A. E., “Modelling the electrical loads of UK residential energy users,” in *2012 47th International Universities Power Engineering Conference (UPEC)*, Uxbridge (GB), 1 – 6, 2012, DOI: 10.1109/UPEC.2012.6398593.
- [88] Zeitbudgets - Tabellenband I. 2001/2002. Wiesbaden 2006.
- [89] Huppertz, Ph., „Systematischer Einsatz von Geoinformationen zur probabilistischen Lastapproximation auf Basis niederspannungsseitiger Netzzustandsgrößen in Echtzeit,“ Dissertation, Universität Siegen, Siegen, 2017.
- [90] James Momoh, “Renewable Energy and Storage,” in *Smart Grid: Fundamentals of Design and Analysis*, IEEE, 140 – 159, 2012, DOI: 10.1002/9781118156117.ch7.
- [91] Bird, R. E. and Hulstrom, R. L., “A Simplified Clear Sky Model for Direct and Diffuse Insolation on Horizontal Surfaces,“ in *Solar Energy Research Institute*, 1 – 38, 1981, DOI: 10.2172/6510849.
- [92] DSG Deutsche Gesellschaft für Sonnenenergie (2012, Hrsg.): *Photovoltaische Anlagen: Leitfaden für das Elektro- und Dachdeckerhandwerk, Fachplaner, Architekten, Ingenieure, Bauherren und Weiterbildungsinstitutionen*, Berlin, 5. Aufl., 2012.
- [93] Engerer, N. A., “Minute resolution estimates of the diffuse fraction of global irradiance for southeastern Australia,“ in *Solar Energy*, 116, 215 – 237, 2015, DOI: 10.1016/j.solener.2015.04.012.

- [94] Milano F., “Power system modelling and scripting,” in *Springer Berlin Heidelberg*, 2010, DOI: 10.1007/978-3-642-13669-6.
- [95] Kersting, W. H., “Distribution System Modeling and Analysis,” in *CRC Press*, 3rd edition, 2012, DOI: 10.1201/b11697.
- [96] Collin, A. J.; Tsagarakis, G.; Kiprakis, A. E. and McLaughlin, S., “Development of Low-Voltage Load Models for the Residential Load Sector,” in *IEEE Transactions on Power Systems*, 29(5), 2180 – 2188, Sept. 2014, DOI: 10.1109/TPWRS.2014.2301949.
- [97] PNNL Pacific Northwest National Lab., *Evaluation of Conservation Voltage Reduction (CVR) on a National Level*, Rep. 19596, Richland, WA (USA), July 2010, DOI: 10.2172/990131.
- [98] Bokhari, A. et al., “Experimental Determination of the ZIP Coefficients for Modern Residential, Commercial, and Industrial Loads,” in *IEEE Transactions on Power Delivery*, 29(3), 1372 – 1381, June 2014, DOI: 10.1109/TPWRD.2013.2285096.
- [99] Ohyama, T.; Watanabe, A.; Nishimura, K. and Tsuruta, S., “Voltage Dependence Of Composite Loads In Power Systems,” in *IEEE Transactions on Power Apparatus and Systems*, PAS-104(11), 3064 – 3073, Nov. 1985, DOI: 10.1109/TPAS.1985.318814.
- [100] Hajagos, L. M. and Danai, B., “Laboratory measurements and models of modern loads and their effect on voltage stability studies,” in *IEEE Transactions on Power Systems*, 13(2), 584 – 592, May 1998, DOI: 10.1109/59.667386.
- [101] Hilder, J., „Messung spannungsabhängiger Verbraucher zur automatisierten Ableitung von Modellkoeffizienten mit Hilfe der Regressionsanalyse,“ unveröffentlichte Masterthesis, Hochschule Düsseldorf, Düsseldorf, Aug. 2018.
- [102] Collin, A. J.; Acosta, J. L.; Hayes, B. P. and Djokic, S. Z., “Component-based aggregate load models for combined power flow and harmonic analysis,” in *7th Mediterranean Conference and Exhibition on Power Generation, Transmission, Distribution and Energy Conversion (MedPower 2010)*, Agia Napa (Zypern), 1 – 10, 2010, DOI: 10.1049/cp.2010.0901.
- [103] BDEW Bundesverband der Energie- und Wasserwirtschaft e. V., (05.03.2020) *Stromverbrauch der Haushalte um neun Prozent gesunken: Rund ein Sechstel des Stroms wird für Unterhaltungselektronik und Kommunikationstechnik genutzt, etwa die Hälfte für Wärmezwecke.* [Pressemitteilung], URL: <https://www.bdew.de/presse/presseinformationen/stromverbrauch-der-haushalte-um-neun-prozent-gesunken/> (abgerufen am 16.01.2022).
- [104] Sen, P. K. and Lee, K. H., “Conservation Voltage Reduction Technique: An Application Guideline for Smarter Grid,” in *IEEE Transactions on Industry Applications*, 52(3), 2122 – 2128, 2016, DOI: 10.1109/TIA.2016.2525937.

- [105] Eberhart, R. and Kennedy, J., “A new optimizer using particle swarm theory,” in *MHS’95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, Nagoya (Japan), 39 – 43, 1995, DOI: 10.1109/MHS.1995.494215.
- [106] Pemmasani, G., “dispy: Distributed and Parallel Computing with/for Python,” 2016, URL: <https://dispy.org> (abgerufen am 27.03.2022).
- [107] Kopczynski, L.; Huppertz, P.; Schallenburger, M. and Zeise, R., “Optimal Tap-Operations of a Regulated Distribution Transformer Considering Conservation Voltage Reduction Effects in Active Low-Voltage Grids,” in *2018 Power Systems Computation Conference (PSCC)*, Dublin (Ireland), 1 – 7, 2018, DOI: 10.23919/PSCC.2018.8442803.
- [108] dena Deutsche Energie-Agentur (2019): *Künstliche Intelligenz für die integrierte Energiewende*, Berlin, Sept. 2019, URL: https://www.dena.de/fileadmin/dena/Publikationen/PDFs/2019/dena-ANALYSE_Kuenstliche_Intelligenz_fuer_die_integrierte_Energiewende.pdf (abgerufen am 26.06.2022).
- [109] Fraunhofer IAIS Institut für Intelligente Analyse- und Informationssysteme (2018): *Maschinelles lernen – Kompetenzen, Anwendungen und Forschungsbedarf*, München, 2018, URL: https://www.bigdata.fraunhofer.de/content/dam/bigdata/de/documents/Publikationen/BMBF_Fraunhofer_ML-Ergebnisbericht_Gesamt.pdf (abgerufen am 26.06.2022).
- [110] Elsevier (2018): *Artificial Intelligence: How knowledge is created, transferred, and used: Trends in China, Europe, and the United States*, 2018, URL: https://www.elsevier.com/__data/assets/pdf_file/0011/906779/ACAD-RL-AS-RE-ai-report-WEB.pdf (abgerufen am 02.07.2022).
- [111] McCarthy, J. and Minsky, M. L. and Rochester, N. and Shannon, C. E.: *A proposal for the Dartmouth summer research project on artificial intelligence*, Projektantrag, Aug. 1955, URL: <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html> (abgerufen am 26.06.2022).
- [112] Stonier, T., “The Evolution of Machine Intelligence,” in *Beyond Information: The Natural History of Intelligence*, Springer, London, 107 – 133, 1992, DOI: 10.1007/978-1-4471-1835-0_6.
- [113] Silver, D. et al., “Mastering the game of Go with deep neural networks and tree search,” in *Nature*, **529**, 484 – 489, Jan. 2016, DOI: 10.1038/nature16961.
- [114] Grum, M. et al.,: “Tools des Maschinellen Lernens: Marktstudie, Anwendungsbeispiele & Lösungen der Künstlichen Intelligenz,” *GITO Verlag*, Berlin, 2021, DOI: 10.30844/grum_2020.

- [115] dena Deutsche Energie-Agentur (2020): *Künstliche Intelligenz - vom Hype zur energiewirtschaftlichen Realität: Vertiefte Analyse von KI-Anwendungsfeldern in der Energiewirtschaft*, Berlin, Aug. 2020, URL: [https:// www.dena.de/ fileadmin/ dena/ Publikationen/ PDFs/ 2020/ dena_ANALYSE_Kuenstliche_Intelligenz _ - _vom_Hype_zur_energiewirtschaftlichen_Realitaet.pdf](https://www.dena.de/fileadmin/dena/Publikationen/PDFs/2020/dena_ANALYSE_Kuenstliche_Intelligenz_-_vom_Hype_zur_energiewirtschaftlichen_Realitaet.pdf) (abgerufen am 26.06.2022).
- [116] Sutton, R. S. and Barto, A. G., “Reinforcement Learning: An Introduction,” in *The MIT Press*, 2nd edition, 2018.
- [117] Lapan, M., “Deep Reinforcement Learning - Hands-On,” in *Packt Publishing Ltd.*, Birmingham, UK, 2018, ISBN 978-1-78883-424-7.
- [118] H. Xu, A. D. Domínguez-García and P. W. Sauer, “Optimal Tap Setting of Voltage Regulation Transformers Using Batch Reinforcement Learning,” in *IEEE Transactions on Power Systems*, vol. 35, no. 3, pp. 1990–2001, May 2020, doi: 10.1109/TPWRS.2019.2948132.
- [119] Devidze, R. et al., “Explicable Reward Design for Reinforcement Learning Agents,” in *NeurIPS*, 2021.
- [120] Lovsjö, N., „From Markov chains to Markov decision processes,“ Bachelorthesis, Stockholm University, Stockholm (Stockholm), June 2015.
- [121] Bellman, R., „Dynamic programming,“ in *Science*, **153**, 34 – 37, 1966.
- [122] Koslow, W., „Reinforcement Learning zur Energiemarktanalyse,“ Masterthesis, Universität zu Köln, Köln, Febr. 2020.
- [123] Watkins, C. and Dayan, P., “Q-learning,” in *Machine Learning*, 8(3), 279 – 292, May 1992, DOI: 10.1007/BF00992698.
- [124] van Seijen, H.; van Hasselt, H.; Whiteson, S. and Wiering, M., “A theoretical and empirical analysis of Expected Sarsa,” in *2009 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning*, 177 – 184, 2009, DOI: 10.1109/ADPRL.2009.4927542.
- [125] Even-Dar, E. and Mansour, Y., “Convergence of optimistic and incremental Q-learning,” in *Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic (NIPS’01)*, Vancouver, British Columbia (Canada), 2001.
- [126] Leshno, M.; Lin, V. Ya.; Pinkus, A. and Schocken, S., “Multilayer feedforward networks with a nonpolynomial activation function can approximate any function,” in *Neural Networks*, **6(6)**, 861 – 867, 1993, DOI: 10.1016/S0893-6080(05)80131-5.
- [127] Glorot, X.; Bordes, A. and Bengio, Y., “Deep Sparse Rectifier Neural Networks,” in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics (ICML’14)*, 15, 315 – 323, Fort Lauderdale, Florida (USA), April 2011.

- [128] Williams, R. J., “Simple statistical gradient-following algorithms for connectionist reinforcement learning,” in *Machine Learning*, 8(3-4), 229 – 256, May 1992, DOI: 10.1007/BF00992696.
- [129] Kingma, D. P. and Ba, J., “Adam: A Method for Stochastic Optimization,” in *arXiv preprint arXiv:1412.6980*, 2015, DOI: 10.48550/ARXIV.1412.6980.
- [130] Silver, D. et al., “Deterministic Policy Gradient Algorithms,” in *Proceedings of the 31st International Conference on International Conference on Machine Learning (ICML’14)*, 32, 387 – 395 , Beijing (China), 2014.
- [131] Ioffe, S. and Szegedy, Ch., “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,” in *arXiv preprint arXiv:1502.03167*, 2015, DOI: 10.48550/ARXIV.1502.03167.
- [132] Uhlenbeck, G. E. and Ornstein, L. S., “On the Theory of the Brownian Motion,” in *Phys. Rev.*, 36(5), 823 – 841, Sept. 1930, DOI: 10.1103/PhysRev.36.823.
- [133] Kopczynski, L.; Wolf, O.; Schallenburger, M.; Zeise, R. and Hirsch, H., *Reinforcement Learning-Based Tap-Operations of a Regulated Distribution Transformer for Autonomous Voltage Control*, International Journal of Simulation: Systems, Science and Technology (**IJSSST**), 2022, 23(1), PP1-8.
- [134] Witte, K. und Hungerland, N., „Künstliche Intelligenz allein reicht nicht,“ in *netzpolitik.org*, 23.10.2020, URL: <https://netzpolitik.org/2020/kuenstliche-intelligenz-allein-reicht-nicht/> (abgerufen am 11.09.2022).
- [135] Strubell, E.; Ganesh, A. and McCallum, A., „Energy and Policy Considerations for Deep Learning in NLP,“ in *arXiv preprint arXiv:1906.02243*, 2019, DOI: 10.48550/arXiv.1906.02243.
- [136] Die Bundesregierung, (2018) *Strategie Künstliche Intelligenz der Bundesregierung*, URL: https://www.bmbf.de/bmbf/shareddocs/downloads/files/nationale_ki-strategie.pdf (abgerufen am 11.09.2022).

DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

ub | universitäts
bibliothek

Diese Dissertation wird via DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt und liegt auch als Print-Version vor.

DOI: 10.17185/duepublico/82228

URN: urn:nbn:de:hbz:465-20240821-104216-6

Alle Rechte vorbehalten.