

Konzept zur datengetriebenen Analyse und Modellierung des preisbeeinflussten Verhaltens

Zur Erlangung des akademischen Grades eines
DOKTORS DER INGENIEURWISSENSCHAFTEN (Dr.-Ing.)
bei der KIT-Fakultät für Maschinenbau des
Karlsruher Instituts für Technologie (KIT)

genehmigte
Dissertation
von

Dipl.-Ing. Simon Waczowicz

Tag der mündlichen Prüfung: 5. März 2018

Referent: apl. Prof. Dr.-Ing. Ralf Mikut

Korreferenten: Prof. Dr.-Ing. Peter Bretschneider,
Prof. Dr.-Ing. Veit Hagenmeyer

Kurzfassung

Der verstärkte Ausbau von Erneuerbare-Energien-Anlagen in Deutschland führt dazu, dass der Anteil der erneuerbaren Energien am gesamten Bruttostromverbrauch im Jahr 2016 bei ca. 32% lag. Die Erzeugungscharakteristik von Photovoltaik- und Windkraftanlagen macht jedoch nachfrage- und erzeugungsseitige Anpassungen im deutschen Elektrizitätsmarkt notwendig. Im Energiesystem der Zukunft (*Smart Grid*) wird die Bereitstellung von lastseitigen Flexibilitäten eine bedeutende Rolle einnehmen.

Die Veränderung der Verbraucherlast durch Preisanreize steht dabei ausdrücklich nicht in Konkurrenz zum verstärkten Ausbau von Energiespeichern, Energieübertragungs- und -verteilnetzen. Die Verbraucherbeeinflussung mithilfe von Steuer- (*Demand Side Management, DSM*) und Preissignalen (*Demand Response, DR*) setzt allerdings voraus, dass die Auswirkungen von Preissignalen für Energieverbraucher auf deren Verbrauchsverhalten untersucht, mathematisch modelliert und validiert werden. Die vorliegende Doktorarbeit beschäftigt sich aus diesem Grund mit der datengetriebenen Analyse und Modellierung des Verbrauchsverhaltens als Reaktion auf variable Stromtarife.

Dazu wurde ein Gesamtkonzept entwickelt, das aus verschiedenen Vorverarbeitungs-, Analyse- und Modellierungsmethoden besteht. Das Konzept behandelt die gesamte Prozesskette von der Erfassung der *Smart-Meter*-Rohdaten bis hin zur Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens. Das Zeitreihen-Clustering als wichtiger Bestandteil des neuen Konzeptes erlaubt, Aussagen über saisonale, wochentagsbedingte, tarifbedingte Unterschiede für eine *Demand-Response*-Maßnahme treffen zu können. Des Weiteren erlaubt das neue Konzept, dass haushaltsindividuelle Unterschiede in der Verbrauchsreaktion (Responder-, Semi-Responder, Non-Responder-Haushalte) identifiziert werden können. Zudem wurde eine neue *Demand-Response*-Modellklasse (Virtuelle-Speicher-Modelle) entwickelt, deren Modelle aus einem System von Differenzgleichungen bestehen und das Verbrauchsverhalten von Haushaltstromkunden als Reaktion auf verschiedene Preissignale beschreiben.

Abstract

The increased expansion of renewable energy plants in Germany leads to a share of renewable energies in the total gross electricity consumption of around 32% (2016). However, the production characteristics of photovoltaic and wind power plants makes it necessary to adjust the demand and production side of the German electricity market. In the energy system of the future (Smart Grid), the provision of demand side flexibility will play an important role.

The change of the consumer load via variable electricity prices does explicitly not compete with the increased expansion of energy storages, transmission grids and distribution grids. However, the use of control (Demand Side Management, DSM) and price signals (Demand Response, DR) to change the consumer load assumes that the effect of these signals has to be investigated, mathematically modeled and validated. Therefore, the present work deals with the data-driven analysis and modeling of the consumption behavior of household customers in response to different price signals.

For this purpose, an overall concept was developed that consists of different preprocessing, analysis and modeling methods. The concept covers the entire process chain from the acquisition of the smart meter raw data to the analysis and modeling of the consumption behavior. The time series clustering as an important part of the new concept allows statements to be made about seasonal, weekly, tariff-related differences for a demand response measure. Furthermore, the new concept allows that household-specific DR differences (responder, semi-responder, non-responder households) can be identified. In addition, a new demand response model class (virtual storage model) was developed whose models consist of a system of difference equations and describe the consumption behavior of household customers in response to different price signals.

Danksagung

Die vorliegende Doktorarbeit entstand während meiner Zeit als wissenschaftlicher Mitarbeiter am Institut für Angewandte Informatik (IAI) des Karlsruher Instituts für Informatik (KIT). Bedanken möchte ich mich deshalb zuallererst bei Prof. Dr.-Ing. habil. Georg Bretthauer und Prof. Dr.-Ing. Veit Hagenmeyer, die mir die Möglichkeit gaben, als Doktorand im IAI an einem äußerst zukunftssträchtigen Forschungsthema zu arbeiten. In diesem Zusammenhang möchte ich mich bei der Deutschen Forschungsgemeinschaft für die Förderung des Forschungsprojektes (MI 1315/2-1) bedanken. Ein ganz besonderer Dank geht an meinen Betreuer apl. Prof. Dr. Ralf Mikut, der durch seine hilfreichen Ratschläge, kreativen Ideen, ständige Diskussionsbereitschaft und durch seine konstruktive Kritik eine große Unterstützung für mich war. Mein herzlicher Dank gilt auch Prof. Dr.-Ing. Peter Bretschneider und Prof. Dr.-Ing. Veit Hagenmeyer für die Übernahme des Korreferats.

Ein weiterer Dank geht an alle Kooperationspartner des DFG-Forschungsprojektes: Prof. Dr.-Ing. Dirk Westermann und Irina Konotop von der Technischen Universität Ilmenau sowie Prof. Dr.-Ing. Peter Bretschneider und Stefan Klaiber vom Fraunhofer IOSB. Ein besonderer Dank geht dabei an Stefan Klaiber zur Bereitstellung von Teilen eines Werkzeugs zur Validierung der entwickelten Methoden.

Zudem möchte ich mich bei meinen Kolleginnen und Kollegen innerhalb der Arbeitsgruppe für wertvolle Diskussionen und für die familiäre Atmosphäre bedanken: Andreas Bartschat, Wolfgang Doneit, Jorge Ángel González Ordiano, Nicole Ludwig, Markus Reischl, Benjamin Schott, Johannes Stegmaier und Michele René Tuga. Danken möchte ich außerdem allen anderen Mitarbeiterinnen und Mitarbeitern des Instituts für die Zusammenarbeit.

Ganz besonders möchte ich mich bei meinen Eltern Elke und Valentin und bei meiner Familie bedanken, auf deren Rückhalt ich jederzeit bauen konnte. Mein innigster Dank gilt meiner Freundin Juliane für ihre permanente Unterstützung und ihr Verständnis für entgangene Feierabende und verkürzte Wochenenden gegen Ende der Promotion. Zu guter Letzt danke ich unserem Sohn Leo, dessen Geburt den nötigen Motivationsschub zur Beendigung der Promotion geliefert hat.

Karlsruhe, im Dezember 2017

Simon Waczowicz

Inhaltsverzeichnis

Kurzfassung	i
Abstract	iii
Danksagung	v
Inhaltsverzeichnis	vii
1 Einleitung	1
1.1 Einordnung der Arbeit	1
1.2 Darstellung des Entwicklungsstandes	4
1.2.1 Versorgungssicherheit Strom	5
1.2.2 Intelligente Stromnetze - <i>Smart Grids</i>	6
1.2.3 <i>Demand Side Management</i>	6
1.2.4 <i>Demand Response</i>	8
1.2.5 Automatisches <i>Demand Response</i>	12
1.2.6 Datenvorverarbeitung bei <i>Smart-Meter</i> -Datensätzen	13
1.2.7 Verfahren zur Prognose der Verbraucherlast	14
1.2.8 <i>Demand-Response</i> -Modellierung	18
1.3 Ungelöste Probleme und offene Fragen	21
1.4 Ziele und Aufgaben	23
2 Neues Datenanalyse- und Modellierungskonzept	25
2.1 Preisbeeinflusster Verbrauchsprozess mit Systemrückkopplung	27
2.2 Anforderungsanalyse	31
2.3 Datenstruktur und Benchmarkdatensätze	36
2.3.1 Datenstruktur	36
2.3.2 Generierung von Benchmarkdatensätzen - Benchmarkgenerator	37
2.3.3 Benchmarkdatensätze	44
2.4 Verfahrensübergreifende Methoden der Pipeline	47
2.4.1 Datenvorverarbeitung	47
2.4.2 Zeitreihenzerlegung	53
2.4.3 Zeitreihennormierung	54
2.4.4 Zeitreihenaggregation	57
2.4.5 Zusammenfassung und Diskussion	59
2.5 Analyseorientierte Methoden der Pipeline	61
2.5.1 Clustering zur Ermittlung von preisbeeinflussten Typtagen	62

2.5.2	Bewertung der Clusterlösung	68
2.5.3	Clusterzuordnungen	69
2.5.4	Sensitivitätsanalysen	74
2.5.5	Analyse der Tarifeinflüsse auf die Last	86
2.5.6	Zusammenfassung und Diskussion	87
2.6	Prognoseorientierte Methoden der Pipeline	91
2.6.1	Integration von <i>a-priori</i> -Wissen bei der Entwicklung des neuen DR-Modells	92
2.6.2	Virtueller Speicher als adaptierbares Verbrauchermodell	93
2.6.3	Strategie zur Berechnung des beeinflussten Anteils der Verbraucherlast	99
2.6.4	Modellentwurf	101
2.6.5	Modellanwendung	103
2.6.6	Zusammenfassung und Diskussion	105
3	Implementierung	109
3.1	<i>Open-Source</i> MATLAB-Toolbox SciXMiner	109
3.2	SciXMiner-Erweiterung <i>Load Benchmark Generator</i>	110
3.3	SciXMiner-Erweiterung <i>Energy</i>	111
3.3.1	Automatisierung der verfahrensübergreifenden Methoden der Pipeline .	112
3.3.2	Automatisierung der analyseorientierten Methoden der Pipeline	113
3.3.3	Automatisierung der prognoseorientierten Methoden der Pipeline . . .	114
3.4	Zusammenfassung	114
4	Anwendungen	117
4.1	<i>Olympic-Peninsula-Project</i> -Datensatz	117
4.1.1	Allgemeines zum Datensatz	117
4.1.2	Optimale Clusterparameter	119
4.1.3	Identifikation von saisonalen und wochentagsbedingten Unterschieden	122
4.1.4	Identifikation von tarifbedingten Unterschieden	128
4.1.5	Identifikation von (Non-) Responder-Haushalten	130
4.1.6	Prognose des preisbeeinflussten Verbrauchsverhaltens	132
4.2	RESIDENS-Datensatz	138
4.2.1	Allgemeines zum Datensatz	138
4.2.2	Optimale Clusterparameter	139
4.2.3	Identifikation von saisonalen und wochentagsbedingten Unterschieden	139
4.2.4	Identifikation von (Non-) Responder-Haushalten	143
4.3	Zusammenfassung und Diskussion	144
5	Zusammenfassung und Ausblick	147
A	Verzeichnis der verwendeten Abkürzungen und Formelzeichen	151
B	Anhang	163
B.1	Einleitung	163
B.1.1	Smart-Grid-Projekte	163
B.1.2	Wichtige Formeln für ökonomische <i>Demand-Response</i> -Modelle	175

B.1.3	Heiz- und Klimatechniksysteme	176
B.2	Methodik	180
B.2.1	Anforderungsanalyse	180
B.2.2	Benchmarkgenerator	181
B.2.3	Parametereinstellungen des Benchmarkgenerators	182
B.2.4	Datenvorverarbeitung	185
B.2.5	Zeitreihennormierung	196
B.2.6	Zeitreihenaggregation	197
B.2.7	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern	198
B.2.8	Sensitivität der Clusterergebnisse bezüglich der Wahl der Startcluster	203
B.3	Implementierung	205
B.4	Anwendungen	209
B.4.1	<i>Olympic-Peninsula-Project</i> -Datensatz	209
B.4.2	RESIDENS-Datensatz	225
C	Abbildungsverzeichnis	233
D	Tabellenverzeichnis	237
E	Listings	239
F	Literaturverzeichnis	241
G	Stichwortverzeichnis	269

1 Einleitung

1.1 Einordnung der Arbeit

In Deutschland findet, spätestens mit dem Energiekonzept der Bundesregierung aus dem Jahr 2010 [1], ein gesellschaftliches und politisches Umdenken bezüglich der strategischen Ausrichtung der Energieversorgung statt. Der Transformationsprozess, der unter dem Begriff „Energiewende“ zusammengefasst ist, macht weitreichende Veränderungen im gesamten deutschen Energieversorgungssystem notwendig [2]. Die Ziele der Energiewende lassen sich in politische Ziele, Kernziele und Steuerungsziele einteilen [3]. Die politischen Ziele legen auf oberster Ebene fest, wie die Energieversorgung in Deutschland umgebaut werden soll. Neben den Klimazielen (z.B. Reduktion der Treibhausgasemissionen) ist hier der Ausstieg aus der Nutzung der Kernenergie bis 2022 und die Sicherstellung der Wettbewerbsfähigkeit und der Versorgungssicherheit zu nennen [3]. Die Kernziele der Energiewende beinhalten die Strategien zum Erreichen der genannten politischen Ziele. Als erstes Kernziel ist die Steigerung des Anteils der erneuerbaren Energien (EE) am Energieverbrauch zu nennen. Das zweite Kernziel der Energiewende ist die Reduktion des Primärenergieverbrauchs bzw. die Erhöhung der Energieeffizienz [3]. Im Strombereich lassen sich die Kernziele in folgende Steuerungsziele konkretisieren: die Erhöhung des Anteils erneuerbarer Energien am Bruttostromverbrauch und die Reduktion des Bruttostromverbrauchs. Im Wärmebereich ergeben sich analoge Steuerungsziele: die Erhöhung des Anteils erneuerbarer Energien am Wärme-/Kälteverbrauch und die Reduktion des Endenergieverbrauchs für Wärme. Die Erhöhung des Anteils erneuerbarer Energien und die Reduktion des Endenergieverbrauchs sind die entsprechenden Steuerungsziele im Verkehrsbereich. Als weiteres Steuerungsziel ist die Erhöhung der Endenergieproduktivität zu nennen. Die Steuerungsziele und die damit verbundenen Maßnahmen (z.B. Erneuerbare-Energien-Gesetz, Erneuerbare-Energien-Wärme-gesetz, Marktanreizprogramm) bestimmen die konkrete Umsetzung der Ziele der Energiewende.

In der öffentlichen Wahrnehmung sind die Energiewende und deren Ziele meist auf den Strombereich beschränkt, da in diesem Bereich die größten Erfolge erzielt wurden (siehe Abb. 1.1a). Der Anteil der erneuerbaren Energien am Bruttostromverbrauch¹ lag im Jahr 2016 bei 31,7%, wobei sich der EE-Anteil an der Stromerzeugung ungefähr zu zwei Fünfteln aus Windkraft und jeweils einem Fünftel aus Photovoltaik und Biomasse zusammensetzt (siehe Abb. 1.2b).

¹Der Bruttostromverbrauch eines Landes entspricht der inländisch produzierten Strommenge (Wind, Wasser, Sonne, Kohle, Öl usw.), zuzüglich der Stromimporte und abzüglich der Stromexporte. Die Bruttostromerzeugung ist dagegen die inländisch produzierte Strommenge. Der EE-Anteil am Bruttostromverbrauch lag im Jahr 2016 deshalb um 2,7 Prozentpunkte (2016: 31,7% in Abb. 1.1a im Vergleich zu 29% in Abb. 1.2a) höher als der EE-Anteil an der Bruttostromerzeugung, da in Deutschland seit 2003 mehr Strom produziert als verbraucht wird.

Bis zum Jahr 2040 soll der Anteil weiter auf 65 Prozent steigen, bis schließlich im Jahre 2050 ein Anteil von 80 Prozent erreicht werden soll (siehe Abb. 1.1a). Um jedoch den Anteil erneuerbarer Energien am gesamten Bruttoendenergieverbrauch² zu steigern, müssen auch der Wärmesektor (vgl. Abb. 1.1b) und der Verkehrssektor (vgl. Abb. 1.1c) berücksichtigt werden.

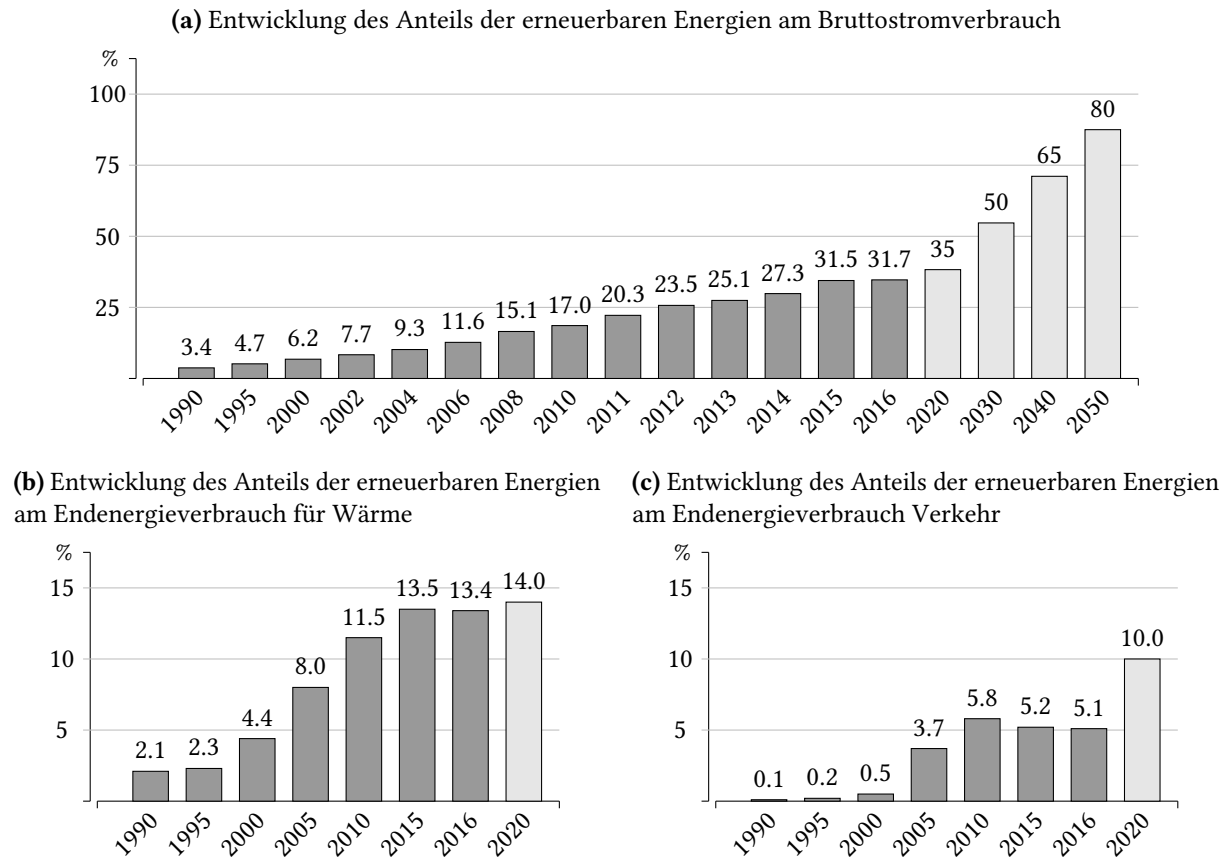


Abbildung 1.1: Anteile der erneuerbaren Energien in den Bereichen (a) Strom, (b) Wärme/Kälte und (c) Verkehr (modifiziert nach [4])

Der Anteil erneuerbarer Energien am Bruttoendenergieverbrauch ist in den Jahren 2008 bis 2015 von 9.1% auf 14.9% gestiegen [4]. Bis zum Jahr 2020 soll ein Anteil von 18% erreicht werden. Dies kann jedoch nur erreicht werden, wenn die Energiewende im Strombereich (vgl. Abb. 1.1a) und Wärmebereich (vgl. Abb. 1.1b) weiter fortschreitet und die Energiewende im Verkehrsbereich (vgl. Abb. 1.1c) deutlich forciert wird.

Die angestrebten Zukunftsszenarien machen den zielgerichteten Ausbau des Stromnetzes auf nationaler und europäischer Ebene erforderlich. Der Ausbau der Stromnetze geht einher mit der Entwicklung neuer Leittechnikarchitekturen, da mit dem Netzausbau auch der Automatisierungsgrad in elektrischen Netzen ansteigen wird [8]. Weitere dringende Handlungsfelder für die Umsetzung der Energiewende sind im Einzelnen: der Ausstieg aus der Nutzung der Kernenergie

²Summe aus Bruttostromverbrauch, Endenergieverbrauch für Wärme/Kälte und Endenergieverbrauch Verkehr

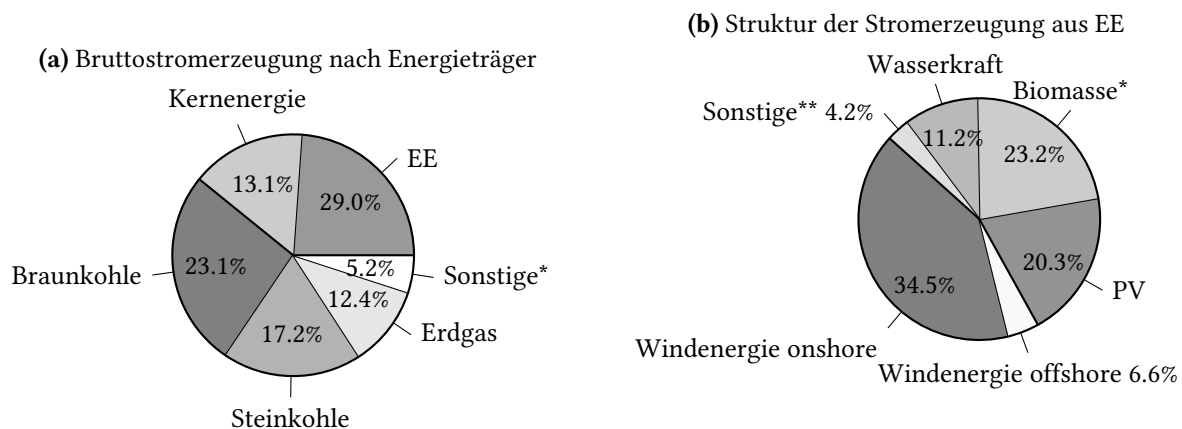


Abbildung 1.2: Bruttostromerzeugung in Deutschland im Jahr 2016: EE - erneuerbare Energien; Sonstige* - Mineralölprodukte, übrige Energieträger; Biomasse* - biogene Festbrennstoffe, biogene flüssige Brennstoffe, Biogas und Biomethan; PV - Photovoltaik; Sonstige** - Geothermie, Deponie-/ Klärgas, biogener Anteil des Abfalls (modifiziert nach [1, 4–7])

bis 2022, das Vorantreiben der Energieeffizienz in privaten Haushalten und in der Industrie, die energetische Gebäudesanierung und das energieeffiziente Bauen neuer Gebäude, der fortwährende Ausbau von erneuerbaren Energiewandlungsanlagen, der Ausbau von Speicherkapazitäten und die Markt- und Systemintegration der erneuerbaren Energien [1]. Im Verkehrsbereich werden folgende zentrale Maßnahmen vorgeschlagen: Ausbau der Elektromobilität und der Ladeinfrastruktur, Einsatz alternativer Kraftstoffe (z.B. Wasserstoff) und die Verlagerung auf umweltfreundliche Verkehrsträger (z.B. Bahn, Rad) [3].

Neben den genannten Maßnahmen in den Sektoren Strom, Wärme und Verkehr erlangt die Sektorkopplung eine wachsende Bedeutung, was sich insbesondere in der Energieforschung zeigt [3, 9, 10]. In zahlreichen Energieforschungsprojekten wird die Integration des Strom-, Wärme- und Verkehrssektors und die Systemintegration innovativer Technologien und Ideen (z.B. EE-Integration und Scheduling unter stochastischer Unsicherheit [11]) zur Erreichung der Ziele der Energiewende erforscht. Beispiele sind die Kopernikus-Projekte³, die Helmholtz-Initiative „Energie System 2050“⁴, das Helmholtz-Energieprogramm „Speicher und vernetzte Infrastrukturen“⁵ [12] und das Energy Lab 2.0⁶ [13] als großskalige Forschungsinfrastruktur zur Erforschung des Zusammenspiels der Komponenten künftiger Energiesysteme.

Der wachsende Anteil der erneuerbaren Energien hat u.a. einen größeren Bedarf an Regelleistung zur Folge, da dieser in erster Linie von der Güte der Leistungsprognose bei Windenergie- und Photovoltaikanlagen abhängt [2]. Der überwiegende Anteil des Regelleistungsbedarfs wird auch zukünftig durch flexible konventionelle Kraftwerke und Energiespeicher gedeckt werden. In diesem Zusammenhang sind Pumpspeicherkraftwerke zu nennen, die vor allem den Spitzenlastbedarf abdecken. Allerdings besteht neben den Pumpspeicherkraftwerken, welche sich

³<https://www.kopernikus-projekte.de/>, Abrufdatum: 20.12.2017

⁴https://www.helmholtz.de/forschung/energie/energie_system_2050/, Abrufdatum: 20.12.2017

⁵<https://www.sci.kit.edu/>, Abrufdatum: 20.12.2017

⁶<https://www.elab2.kit.edu/>, Abrufdatum: 20.12.2017

hauptsächlich in Süddeutschland befinden, ein Defizit an Speicheralternativen im großtechnischen Maßstab. Technisch mögliche Druckluft- und Wasserstoffspeicher kommen derzeit nicht als kommerziell nutzbare Großspeichertechnologien in Betracht [2, 14].

In Ergänzung zu Energiespeichern eignet sich der verbraucherseitig beeinflussbare Leistungsanteil der Nieder- und Mittelspannungsebene zur Bereitstellung der erforderlichen Flexibilität zur Synchronisation von Stromerzeugung und Stromverbrauch. Die ökonomische und ökologische Bewertung verschiedener Maßnahmen zum Ausgleich von Stromerzeugung und -verbrauch ist in Abb. 1.3 veranschaulicht. Es wird deutlich, dass Maßnahmen zur Beeinflussung der Verbraucherlast (Energieeffizienz, *Demand Response*) aus wirtschaftlichen und ökologischen Gesichtspunkten sinnvoll erscheinen. Die Beeinflussung der Verbraucherlast mittels Steuer- (*Demand Side Management*, DSM) und Preissignalen (*Demand Response*, DR) verlangt, dass die Auswirkungen von Preissignalen für Energieverbraucher auf deren Verbrauchsverhalten untersucht, mathematisch modelliert und validiert werden. Die vorliegende Arbeit beschäftigt sich aus diesem Grund mit der Analyse und der mathematischen Modellierung des Verbrauchsverhaltens als Reaktion auf variable Stromtarife.

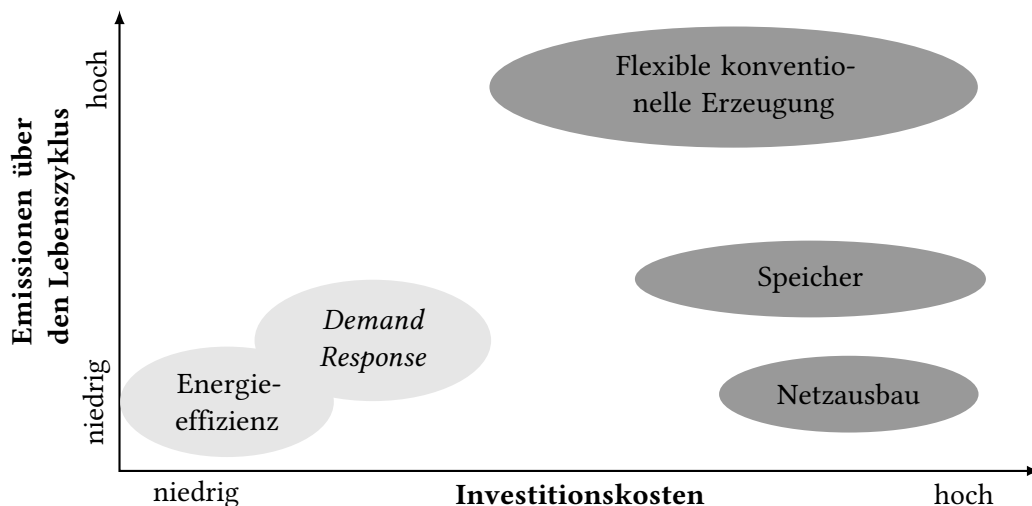


Abbildung 1.3: Ökonomische und ökologische Einordnung verschiedener Maßnahmen zur Synchronisation von Stromerzeugung und Stromverbrauch (modifiziert nach [15])

1.2 Darstellung des Entwicklungsstandes

In diesem Abschnitt wird der aktuelle Entwicklungsstand auf dem Gebiet des *Demand Side Management* bzw. *Demand Response* dargelegt. Dazu wird zunächst in den Abschnitten 1.2.1 und 1.2.2 der energietechnische Rahmen vorgestellt, in dem Maßnahmen zur verbraucherseitigen Lastbeeinflussung eingebettet sind. In Abschnitt 1.2.3 wird auf die Definition und auf den Einsatz von *Demand Side Management* eingegangen. Abschnitt 1.2.4 widmet sich dem *Demand Response* und stellt ausgewählte DR-Feldstudien vor, bei denen variable Strompreise zur Beeinflussung der Haushaltsstromkunden verwendet wurden. Das automatische *Demand Response*

wird in Abschnitt 1.2.5 beschrieben. Ein wichtiger Schritt bei der datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens stellt die Datenvorverarbeitung dar. Der *State-of-the-Art* bei der Datenvorverarbeitung von *Smart-Meter*-Datensätzen wird in Abschnitt 1.2.6 präsentiert. Gängige Verfahren zur Prognose der Verbraucherlast werden in Abschnitt 1.2.7 vorgestellt. Die Modellierung des preisbeeinflussten Verbrauchsverhaltens stellt den Kern der vorliegenden Doktorarbeit dar. Aus diesem Grund werden verschiedene Modellierungsansätze in Abschnitt 1.2.8 ausführlich diskutiert.

1.2.1 Versorgungssicherheit Strom

Die elektrische Energieversorgung in Deutschland ist im internationalen Vergleich sehr zuverlässig. Die durchschnittliche Unterbrechungsdauer der angeschlossenen Letztverbraucher (*System Average Interruption Duration Index*, SAIDI) betrug im Jahr 2015 lediglich 12,70 Minuten (10,45 min Unterbrechung in der Mittelspannung, 2,25 min Unterbrechung in der Niederspannung [3]) und liegt damit deutlich unter dem Durchschnitt der vergangenen neun Jahre (2006-2014: 16,23 min) [3, 16]. Die langfristige Versorgungssicherheit bzgl. der Elektrizitätsversorgung in Deutschland ist gekennzeichnet durch

- die Sicherstellung der Verfügbarkeit benötigter Primärenergieträger,
- die Sicherstellung ausreichender Erzeugungskapazitäten und
- die Sicherstellung des Elektrizitätstransports durch ausreichende Transportkapazitäten [17].

Werden dagegen kürzere Zeiträume betrachtet, ist vielmehr die Gewährleistung der Systemstabilität für die Versorgungssicherheit entscheidend. Die Gewährleistung der Systemstabilität gehört zu den Kernaufgaben des Übertragungsnetzbetreibers und umfasst die Frequenzhaltung (Primärregelung, Sekundärregelung, Minutenreserve), die Vermeidung von Betriebsmittelüberlastungen und die Spannungshaltung [16]. Zur Gewährleistung der Sicherheit der öffentlichen Stromversorgung werden Stromnetze nach dem (n-1)-Prinzip geplant und gebaut. Dieses Prinzip besagt, dass bei Störung oder Ausfall einer Komponente (z.B. Transformator oder Hochspannungsfreileitung) die Netzsicherheit bei maximalen Übertragungs- und Versorgungsaufgaben gewährleistet bleibt [18].

Maßnahmen zur Beeinflussung der Nachfrage, sprich DSM bzw. DR, eignen sich sowohl für die Erhöhung der kurzfristigen, als auch der langfristigen Versorgungssicherheit in Deutschland. Wie bereit in Abschnitt 1.1 geschildert, wird durch DSM bzw. DR die erforderliche Flexibilität zur Synchronisation von Stromerzeugung und Stromverbrauch bereitgestellt, was die kurzfristige Versorgungssicherheit begünstigt [19]. Die langfristige Versorgungssicherheit wird insofern verbessert, als der Bedarf an Primärenergieträgern und bereitzustellende Erzeugungs- und Transportkapazitäten durch Nachfrageflexibilisierung minimiert wird [17].

1.2.2 Intelligente Stromnetze - *Smart Grids*

Der Stromfluss in traditionellen Stromnetzen ist *Top-Down* gerichtet, was bedeutet, dass er von hohen zu niedrigen Spannungen verläuft. Wenige große, zentrale Erzeugungsanlagen versorgen Stromkunden auf allen Spannungsebenen. Die Integration der erneuerbaren Energien erfordert die Neugestaltung der Stromnetze auf allen Ebenen (Übertragungsnetz- und Verteilnetzebene). Durch die Vielzahl an dezentralen Einspeiseanlagen, die vermehrt in unteren Spannungsebenen in das Stromnetz einspeisen, ist eine strenge Ausrichtung des Stromflusses nicht mehr gegeben. Neben den Veränderungen auf der Erzeugerseite wird auch die Verbraucherscharakteristik, beispielsweise durch Elektrofahrzeuge oder Wärmepumpen, einem Wandel unterliegen [20]. Die bidirektionalen Lastflüsse und neuen Dynamiken im Verteilnetz erfordern daher eine Anpassung der Netzinfrastruktur. *Smart Grids* sind notwendig, da die derzeitige Elektrizitätsinfrastruktur auf diese neuen Herausforderungen nicht ausgelegt ist. Laut [20] sollen dafür intelligente Ortsnetzstationen in das bestehende Netzsystem eingebaut werden. Zudem ist eine zeitlich hoch aufgelöste Messung, Regelung und Automatisierung des Stromflusses vonnöten.

Definition von *Smart Grid*: „Der Begriff *Smart Grid* (intelligentes Energieversorgungssystem) umfasst die Vernetzung und Steuerung von intelligenten Erzeugern, Speichern, Verbrauchern und Netzbetriebsmitteln in Energieübertragungs- und Verteilnetzen mit Hilfe von Informations- und Kommunikationstechnik (IKT). Ziel ist, auf Basis eines transparenten, energie- und kosteneffizienten sowie sicheren und zuverlässigen Systembetriebs, die nachhaltige und umweltverträgliche Sicherstellung der Energieversorgung.“ [21]

Eine detaillierte Beschreibung des Systems *Smart Grid*, dessen Komponenten und möglichen Marktteilnehmern liefern [22–26]. Des Weiteren ist in [27] die Einbindung von *Demand Response* im *Smart Grid* inklusive der technologischen Rahmenbedingungen (z.B. *Smart Technologies*, Steuergeräte, IKT) umfangreich beschrieben. Auf der Ebene von Verteilnetzen stellen *Smart-Meter*-Technologien eine entscheidende Komponente für die Steuerung von Verbrauchern in einem *Smart Grid* dar. Diese neuen (intelligenten) Energiezähler ermöglichen zum einen die Reflektion des eigenen Verbrauchsverhaltens auf der Ebene des Haushaltsstromkunden durch Visualisierung des vergangenen und aktuellen Stromverbrauchs [28, 29]. Zum anderen unterstützen *Smart Meter* die Einführung und Abrechnung zeitvariabler Stromtarife, indem sie den Zeitpunkt und die Amplitude der Haushaltslast aufzeichnen [25, 30].

In zahlreichen europäischen *Smart-Grid*-Projekten wurden variable Stromtarife als Mittel zur Steuerung der Stromnachfrage eingesetzt [31]. Eine umfangreiche Auflistung von vergangenen und aktuellen *Smart-Grid*-Projekten in der EU ist in Tabelle B.1 zu sehen.

1.2.3 *Demand Side Management*

Bei unerwarteten Ereignissen oder Abweichungen des Leistungsgleichgewichts im Stromnetz, die zur Destabilisierung eines Versorgungsnetzes führen, müssen so schnell wie möglich und nach einem vorgeschriebenen Einsatzplan Eingriffe von Seiten des Netzbetreibers vorgenommen werden [32]. Grundsätzlich kann die Kompensation der Abweichung zwischen Leistungsbedarf

und erwartetem Leistungsangebot durch Leistungsanpassungen bei schnell regelbaren Kraftwerken (z.B.: Einsatz von Gasturbinen-Kraftwerken bei Spitzenlast oder bei Ausfall eines Kraftwerks als Minutenreserve [33]), durch den Einsatz von Pumpspeicherkraftwerken (Bereitstellung von Regelleistung in Übertragungsnetzen mit geringer Hochlaufzeit von nur ca. 90 Sekunden [33]) oder durch nachfrageseitige Laständerungen [34] erreicht werden. Letztgenanntes wird mit dem Begriff *Demand Side Management* beschrieben.

Definition von Demand Side Management: „Alle Maßnahmen, die die Last auf der Verbraucherseite beeinflussen, werden als Lastmanagement oder *Demand Side Management* bezeichnet.“ [35]

In der Literatur wird vorgeschlagen, DSM-Maßnahmen in anreizorientierte Programme (*Incentive-Based Programs*, IBP) und preisorientierte Programme (*Price-Based Programs*, PBP) zu unterscheiden [36–40]. Programme wie die direkte Laststeuerung (*Direct Load Control*, DLC) oder unterbrechbare Lasten (*Interruptible/Curtailable load*, I/C) werden dem klassischen IBP zugeordnet. Daneben existieren zahlreiche Maßnahmen, die den *Market-Based* Programmen (*Emergency Demand Response*, *Capacity Market*, *Demand Bidding* oder *Ancillary Service Markets*) zugewiesen werden. Eine Zusammenstellung aller anreizorientierter Programme inklusive ausgewählter Literaturbeiträge ist in Tabelle 1.1 zu finden. Einen guten Überblick über Architekturen und Komponenten des DSM und über DSM-Optimierungsmodelle liefert [41]. Als praktisches Bei-

Tabelle 1.1: Zusammenstellung verschiedener anreizorientierter Programme (modifiziert nach [38, 42])

Programmbezeichnung engl., deutsch	Abk.	Beschreibung	Bemerkung	Lit.
<i>Direct Load Control</i> , Direkte Laststeuerung	DLC	zyklische Steuerung und Schaltung von Geräten	Klimageräte, Warmwasserboiler; auch im Haushaltssektor eingesetzt	[38, 43–47]
<i>Interruptible load</i> , Unterbrechbare Lasten <i>Curtailable load</i> , Lastbegrenzung	I/C	EVU nimmt Laständerung beim Stromkunden auf vorher vereinbartes Level vor, Zahlung niedriger Tarife	nur bei großen Industriebetrieben (ca. 200 kW bis 3 MW)	[48]
<i>Emergency DR</i> , Notfallprogramme	EDR	Ausgleichszahlungen für Lastreduktion im Eventfall	Stromkunde (Großabnehmer) entscheidet von Fall zu Fall über Laständerung	[49–51]
<i>Capacity Market</i>	CAP	Stromkunden nehmen vereinbarte Laständerung im Eventfall vor	Strafzahlung bei ausbleibender Laständerung	[48]
<i>Demand Bidding</i> , NachfrageBieterverfahren	DB	Großverbraucher bieten per Auktion Laständerungen an	Preis- oder Mengengebote	[42, 52]
<i>Ancillary Service Markets</i>	A/S	Substitut für Systemdienstleistungen	ähnlich der Sekundärregelung im europäischen Stromnetz	[53–55]

spiel für ein anreizorientiertes Programm kann ein Kühlhaus genannt werden, welches durch ein Steuersignal des Energieversorgungsunternehmens (EVU) veranlasst wird, seinen Betriebszustand zu ändern [35]. Der zulässige Temperaturbereich des Kühlgutes muss selbstverständlich gewährleistet sein.

Zur deutlichen Abgrenzung von IBP gegenüber PBP wird die Differenzierung von DSM in direktes und indirektes DSM eingeführt. Anreizorientierte Programme (IBP) sind im Folgenden mit

direktem DSM gleichzusetzen, da in diesen Fällen die Laständerung durch eine direkte Kommunikation bzw. Interaktion zwischen Stromkunde und DSM-Anbieter stattfindet [38]. Demzufolge werden PBP und indirektes DSM synonym verwendet. Maßnahmen des direkten DSM sind i.d.R. von kurzer Dauer und verschlechtern die Prozessqualität [39].

Als indirektes DSM wird die Einflussnahme auf die Verbraucherlast über ein zeitlich änderndes Preissignal verstanden. Indirektes DSM ist mit *Demand Response* gleichzusetzen. *Demand Response* umfasst dabei jede vorsätzliche Änderung des Verbrauchsmusters, die dazu bestimmt ist, den Zeitpunkt, das Niveau der momentanen Nachfrage oder den gesamten Verbrauch zu ändern [35, 36]. Eine ausführliche Beschreibung von Maßnahmen des indirekten DSM (PBP) erfolgt in Abschnitt 1.2.4. Die Differenzierung in direktes und indirektes DSM sollte der Vollständigkeit halber um die Kategorie „Energieeffizienz“ ergänzt werden. Energieeffizienzmaßnahmen bewirken eine dauerhafte Veränderung bzw. Beeinflussung der Kundenlast und bedingen zudem eine erhöhte Prozessqualität, was mit einer dauerhaften Reduktion des Gesamtenergiebedarfs einhergeht [39]. Abb. 1.4 zeigt abschließend eine Zusammenfassung verschiedener DSM-Programme und deren Kategorisierung in direktes, indirektes DSM und Energieeffizienz.

Zielstellung	Ökonomisch-/ Marktgeführt		Umweltgeführt		Netzgeführt	
Beeinflussung	Leistung			Leistung und Energie		
Akteure	Ein Akteur		Mehrere Akteure		Mehrere Akteure - gegenläufige Beeinfl.	
Strategie	Peak Clipping	Valley Filling	Load Shifting	Strategic Conservation	Strategic Load Growth	Flexible Load Shape
Kategorisierung	Direktes DSM, Incentive-Based Programs (IBP)			Indirektes DSM, Price-Based Programs (PBP), Demand Response		Energieeffizienz
	Classical IBP		Market-Based IBP			
Programme	Direct Load Control	Interruptible rates	Emergency DR programs	Capacity Market programs	Demand Bidding programs	Real-Time Pricing (RTP) Critical Peak Pricing (CPP) Time-of-Use (TOU)
Signale	Steuersignal bekannt		Steuersignal nicht bekannt		Preissignal bekannt	Preissignal unbekannt

Abbildung 1.4: Unterteilung und Kategorisierung von *Demand Side Management* (modifiziert nach [39, 43, 56, 57])

1.2.4 Demand Response

Definition von Demand Response: „Demand Response ist eine kurzfristige und planbare Veränderung der Verbraucherlast als Reaktion auf Preissignale im Markt oder auf eine Aktivierung im Rahmen einer vertraglichen Leistungsreserve. Diese Marktpreise oder Leistungsabrufe werden durch ungeplante, unregelmäßige oder extreme energiewirtschaftliche Ereignisse ausgelöst.“ [58]

Preisorientierte DSM-Programme nutzen einen variablen Stromtarif, um den Stromverbrauch von Haushaltsstromkunden zu flexibilisieren. Je nach Dynamik und Beschaffenheit des Stromtarifs ergeben sich zahlreiche Tarifmodellstrukturen. Die drei gängigsten preisorientierten Pro-

gramme sind auch in Abb. 1.4 zu erkennen. Zudem zeigen die Abb. 1.5a, 1.5b und 1.5c den beispielhaften tageszeitlichen Verlauf des Preissignals dieser drei Programme.

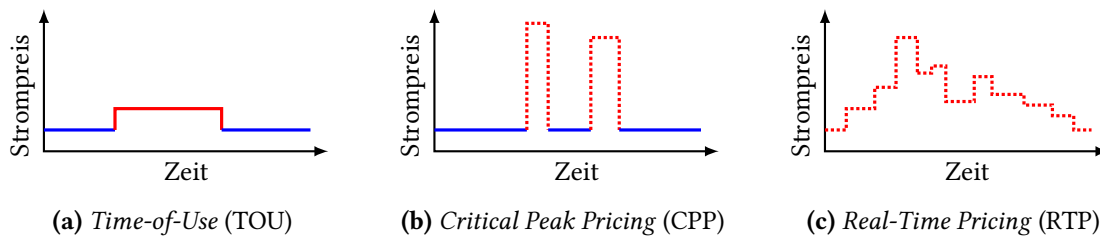


Abbildung 1.5: Typische Verläufe des Preissignals bei verschiedenen *Demand-Response*-Programmen; die gestrichelten (roten) Linien deuten an, dass der Zeitraum und die Höhe des Strompreises in horizontaler und vertikaler Richtung verschiebbar sind (modifiziert nach [59])

Bereits in den 1970 und 1980 Jahren fanden zahlreiche Feldstudien statt, bei denen der Einfluss zeitvariabler Stromtarife auf das Lastverhalten von Haushaltsstromkunden gemessen wurde. Eine Zusammenfassung dieser Studien liefert [60]. Hervorgerufen durch die kalifornische Energiekrise 2000/2001 wurde eine zweite Welle von groß angelegten experimentellen Feldstudien mit Verbrauchsbeeinflussung durch variable Stromtarife losgetreten [61–63]. Diese unterscheiden sich von früheren Feldstudien insofern, als Tarife mit anderen, z.T. komplexeren, Tarifstrukturen eingesetzt wurden.

In zahlreichen weiteren DR-Pilotstudien konnte die Wirksamkeit von variablen Stromtarifen zur Veränderung des Verbrauchsverhaltens gezeigt werden (eTelligence [64], Intelliekon [65], Modellstadt Mannheim [66]). Durch Laständerungen können Spitzenlasten in Haushalten reduziert und stark fluktuierende Lastkurven geglättet werden. Um die Lastverlagerungspotenziale im Industriebereich zu nutzen, werden dort entsprechende Klimatisierungs- und Kälteanwendungen eingesetzt [67]. Koch [46] liefert hierfür umfassende Untersuchungen zur Modellierung großer Gruppen thermostat geregelter Lasten, welche für Heiz- und Kühlanwendungen zum Einsatz gebracht werden. Der Fokus der vorliegenden Doktorarbeit richtet sich jedoch auf Anwendungen im Haushaltsbereich.

Eine genaue quantitative Angabe des DR-Potenzials ist wegen der unterschiedlich gestalteten Rahmenbedingungen (Tarifstruktur, Geräteausstattung) schwierig. Eine Übersicht über gängige preisorientierte Programme liefert Tabelle 1.2. Des Weiteren fasst die Tabelle die charakteristischen Merkmale der Programme zusammen. Grundlage hierfür bilden die Ergebnisse aus [68], wo 15 neuere Feldstudien ausführlich diskutiert und bewertet wurden. Weitere internationale Projekte mit dem Einsatz variabler Stromtarife sind in [69–71] aufgeführt.

Die Daten aus Tabelle 1.2 belegen, dass sich der Einfluss zeitvariabler und dynamischer Stromtarife durch den Einsatz von Technologien zur automatischen Laständerung (engl. *Automated Demand Response* (ADR), in Tabelle 1.2 mit *w/T* gekennzeichnet; siehe Abschnitt 1.2.5) deutlich steigern lässt. Zudem wird deutlich, dass CPP-Tarifmodelle tendenziell besser zur Spitzenlastreduktion geeignet sind als TOU-Tarifmodelle. Die Menge an auswertbaren Feldstudien mit RTP-Tarifen ist sehr gering und lässt daher keine allgemeingültigen Aussagen zum Potenzial von RTP-Tarifen zu. Im Folgenden sollen für jedes Tarifmodell ein oder zwei Feldstudien vorgestellt und diskutiert werden.

Time-of-Use Programme: Ein Stromtarif, welcher aus wenigen Preisstufen besteht, eine langfristige Gültigkeit besitzt und sich an der zeitlichen Struktur der Nachfrage orientiert, wird als *Time-of-Use* (TOU) Tarif bezeichnet. TOU-Tarifmodelle wurden bereits in zahlreichen Feldstudien als Flexibilisierungsoption eingesetzt (vgl. Tabelle 1.2).

Es hat sich bei TOU-Tarifen gezeigt, dass das Tarifdesign einen wesentlichen Einfluss auf die Reaktion von Haushaltsstromkunden hat. Sowohl die absolute Höhe der Tarifstufen, als auch die Spreizung zwischen Hoch- und Niedertarif müssen bei der Beurteilung von TOU-Tarifen berücksichtigt werden. In Abb. 1.6 ist dazu der Zusammenhang zwischen Lastreduktion und dem Quotienten aus *on-peak/off-peak* Strompreis dargestellt. Eine größere Spreizung zwischen *on-* und *off-peak* Strompreis führt dazu, dass Haushaltsstromkunden eher bereit sind, die Last zu Spitzenlastzeiten zu reduzieren bzw. zu verschieben. Dieser Effekt wird durch ADR-Technologien (siehe Abschnitt 1.2.5) weiter verstärkt. Weitere Untersuchungen hierzu bietet [72].

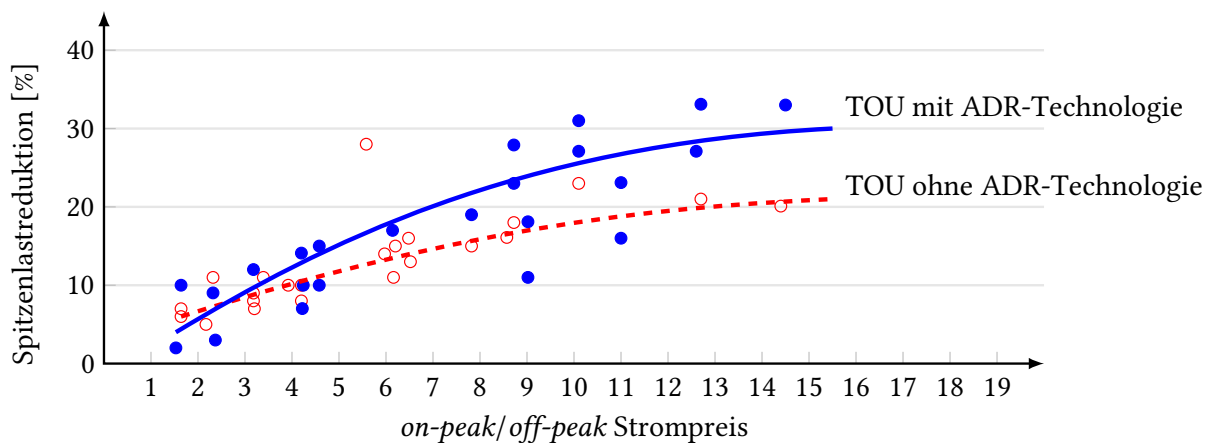


Abbildung 1.6: Zusammenhang zwischen Spitzenlastreduktion und Strompreisspreizung in experimentellen Demand-Response-Studien mit TOU-Tarifen und mit/ohne ADR-Technologien; die Linien stellen eine Approximation der Messpunkte aus den Feldstudien dar (modifiziert nach [70])

Im Teilprojekt *Olympic Peninsula Project* [73] des *GridWise Olympic Peninsula Project* [74] sollte der Einfluss verschiedener variabler Stromtarife auf das Verbrauchsverhalten über ein Zeitraum von einem Jahr untersucht werden. 112 Haushalte aus dem Bundesstaat Washington, USA, nahmen an der Feldstudie teil, wobei die Haushalte gleich verteilt einer der vier Tarifgruppen TOU-, RTP-, FIXED- oder einer Kontrollgruppe angehörten. Alle Haushalte wurden mit intelligenten Haushaltsgeräten ausgestattet. Der TOU-Tarif bestand aus drei Tarifstufen: einer *off-peak*, *on-peak* und einer *Critical Peak Pricing* Tarifstufe. Das Verhältnis des *on-peak* zum *off-peak* Strompreis betrug 2.7 bzw. 2.95. Haushalte der TOU-Tarifgruppe wiesen eine signifikante Lastreduktion in Hochtarifzeiten von ungefähr 20% verglichen mit der FIXED-Gruppe auf [75]. Die *Critical Peak Pricing* Tarifstufe, welche lediglich an einem Tag während der einjährigen Projektlaufzeit ausgerufen wurde, führte sogar zu einer Lastreduktion von 40% [76]. Die Spitzenlastreduktion der Haushalte der *Real-Time Pricing* Tarifgruppe belief sich auf 15-17% im Vergleich zur FIXED-Gruppe [68, 75].

Im Forschungsprojekt RESIDENS⁷ [77, 78] wurde ein dreistufiger TOU-Tarif als Tarifmodell eingesetzt, um nachfrageseitige Laständerungen zu bewirken. Im Untersuchungszeitraum (Sep. 2010 - Okt. 2011) wurde das Verbrauchsverhalten von 64 Mietern und 88 Hausbewohnern in Thüringen aufgezeichnet. Sowohl die Messdaten aus dem *Olympic Peninsula Project* als auch aus dem Forschungsprojekt RESIDENS werden im Rahmen der Arbeit untersucht.

Critical Peak Pricing Programme: Ein *Critical Peak Pricing* Stromtarif besteht aus wenigen Preisstufen und einem sehr hohen Spitzenlastpreis. Ein Energieversorger nutzt diese Tarifvariante, um auf seltene und kritische Netzbelastungen (Events) zu reagieren. Der Haushaltsstromkunde erhält die Information über das zeitliche Auftreten der CPP-Events meist nur einen Tag im Voraus. Die Auftretenswahrscheinlichkeit der Preissprünge liegt bei ca. 1% der Jahreslaststunden [79].

Ein Projekt, bei dem zwei verschiedene CPP-Tarife zum Einsatz kamen, war Kaliforniens *Statewide Pricing Pilot* (SPP) zwischen Juli 2003 und Dezember 2004. Die 2500 Projektteilnehmer setzten sich aus Haushaltsstrom-, Gewerbe- und Industriekunden zusammen. Die Haushalte wurden so ausgewählt, dass Kaliforniens unterschiedliche Klimabedingungen repräsentiert waren. Neben einem klassischen TOU-Tarif, der an 350 Tagen im Jahr galt, wurden an 15 kritischen Tagen zwei unterschiedliche CPP-Tarife, mit einem Verhältnis von *on-peak* zu *off-peak* Strompreis von ungefähr fünf, eingesetzt. Der *Critical Peak Pricing-Fixed* (CPP-F) Tarif wurde einen Tag im Voraus angekündigt und der Zeitraum der kritischen Spitzenlast blieb konstant (14 bis 19 Uhr). Es wurde zudem keine ADR-Technologie verwendet. Beim *Critical Peak Pricing-Variable* (CPP-V) Tarif war die Startzeit und die Länge des Zeitraums der kritischen Spitzenlast variabel (eine bis fünf Stunden im Zeitraum von 14 bis 19 Uhr) und die Haushaltsstromkunden wurden mit ADR-Technologie ausgestattet. Der in dieser Studie eingesetzte TOU-Tarif führt zu einer deutlich geringeren Spitzenlastreduktion als der CPP-F Tarif (5.9% zu 14.4% in den Sommermonaten Juli bis September) [68]. Ein Vergleich von CPP-F mit CPP-V zeigt, dass ADR-Technologien in Haushalten eine Spitzenlastreduktion begünstigen. Die Spitzenlastreduktion lag für unterschiedliche Haushalte der CPP-V Tarifgruppe im Bereich von 16 bis 27%.

Real-Time Pricing Programme: Im Vergleich zu TOU- und CPP-Programmen stellt das *Real-Time Pricing* die flexibelste Form der Tarifgestaltung dar [79]. Die Höhe des Strompreises wird in diesem Fall dynamisch an die Marktgegebenheiten gemäß Angebot und Nachfrage angepasst. Die Information über die Höhe des RTP Strompreises liegt dem Stromkunden i.d.R. in einem Bereich von einer bis 24 Stunden im Voraus vor. Somit wird das Preisrisiko bei einem in Echtzeit schwankenden Strompreis vollständig an den Stromkunden weitergegeben. Die dynamische Abbildung des Marktgeschehens in einem RTP-Tarif bietet das größte Potenzial zur Stromkostenreduktion, wird allerdings ohne ADR-Technologien nicht umzusetzen sein [70].

⁷Abkürzung für „EffizienteRe Energienutzung durch Systemtechnische Integration Des privaten ENdabnehmerS“

Einziges Projekt aus [68], bei dem RTP Strompreise eingesetzt wurden, ist das *Olympic Peninsula Project* (OPP) [73]. Wie oben bereits beschrieben, konnten Haushalte der RTP-Tarifgruppe ihre Spitzenlast im Vergleich zu einer Kontrollgruppe um 15-17% reduzieren.

Tabelle 1.2: Zusammenstellung verschiedener preisorientierter Programme und deren Einfluss auf die Spitzenlastreduktion, wie in [68] gezeigt (modifiziert nach [36, 38, 79–81]): S - Anzahl an Studien.

Bezeichnung <i>englisch</i> , deutsch	Abk.	Beschreibung, Bemerkung	S	Mittl. Spitzen- lastreduktion (Min, Max) (%)
<i>Time-of-Use</i> , Zeitvariabler Tarif mit statischen Preisstufen	TOU	Zeitvariabler Tarif mit wenigen, meist 2 bis 3 Preisstufen; i.d.R. abhängig von Tageszeit oder Wochentag mit saisonalen Schwankungen (Winter-/ Sommertarif); langfristige Gültigkeit über Monate oder ein Jahr; Bsp. siehe Abb. 1.5a	5	4 (2, 6)
<i>Time-of-Use w/T</i>	TOU w/T	siehe TOU; automatisierte Lastverlagerung	4	26 (21, 32)
<i>Critical Peak Pricing</i> , Zeitvariabler Tarif mit Event-Preisstufen	CPP	Zeitvariabler Tarif mit zusätzlichen Preissprüngen bei Events, die mit bestimmter Frist bekannt gegeben wird; Eventdauer zwischen wenigen Stunden bis ganze Tage; Bsp. siehe Abb. 1.5b	8	17 (12, 25)
<i>Critical Peak Pricing w/T</i>	CPP w/T	siehe CPP; automatisierte Lastverlagerung	8	36 (16, 51)
<i>Peak-Time Rebate</i>	PTR	Zeitvariabler Tarif mit expliziten Spitzenlastperioden; Gutschrift bei Unterschreitung des haushaltsindividuellen mittleren Leistungswertes vergangener (fünf) Spitzenlasttage	3	13 (9, 18)
<i>Real-Time Pricing</i> , Dynamischer Tarif	RTP	Tarifmodell mit dynamischen Preisstufen und Preisbekanntgabe am Vortag (<i>day ahead</i>) oder stündlich (<i>hour ahead</i>); Bsp. siehe Abb. 1.5c	1	15
<i>Real-Time Pricing w/T</i>	RTP w/T	siehe RTP; automatisierte Lastverlagerung	1	17

1.2.5 Automatisches *Demand Response*

Durch eine Zunahme von steuer- und programmierbaren Haushaltsgeräten, die sich zur automatischen Lastverlagerung eignen, erhöht sich das Lastverlagerungspotenzial in Haushalten. Tabelle B.2 fasst das technische Lastverlagerungspotenzial verschiedener Anwendungen zusammen. Für eine umfassende Abschätzung des technischen Flexibilitätspotenzials von Gebäuden in Deutschland sei auf [82] verwiesen.

Hinsichtlich des Automatisierungsgrades muss zwischen manuellem, halb-automatischem und automatischem *Demand Response* unterschieden werden [79]. Beim manuellen *Demand Response* reagiert der Haushaltsstromkunde auf eine Preisänderung, indem Elektrogeräte, Thermostate, Schalter oder Steuerungseinheiten per Hand bedient werden und somit auf die neue Preissituation reagiert wird. Die Information über eine zukünftige Strompreisänderung erreicht den Stromkunden per E-Mail, Telefon oder *in-home* Display. Langzeituntersuchungen mit Beteiligung solcher Feedback-Technologien liefern [28, 83]. Bei halb-automatischen DR-Systemen werden wenige Verhaltensmuster vorprogrammiert, ein Großteil der Haushalts- und Klimageräte kann aber weiterhin manuell bedient werden [79]. Beispielsweise könnte ein programmierbarer Raumthermostat auf kritische Strompreise (CPP-Event) insofern reagieren, als Klimageräte bis zu einem voreingestellten Temperaturniveau heruntergeregelt werden. Im Gegensatz dazu

muss bei automatischen DR-Systemen kein manueller Eingriff von Seiten des Stromkunden als Reaktion auf ein Preissignal vorgenommen werden. Eine Deaktivierung der Automatik sollte gewährleistet werden.

1.2.6 Datenvorverarbeitung bei *Smart-Meter*-Datensätzen

Wie in Abschnitt 1.2.2 angedeutet, ermöglichen *Smart Meter* die Aufzeichnung der Haushaltslast und des zeitvariablen Strompreises. Die aufgezeichneten Rohdaten der Haushaltslast und des Strompreises lassen sich dabei als *Smart-Meter*-Datensatz zusammenfassen. Die Rohdaten vieler *Smart-Meter*-Datensätze weisen typische Fehler wie fehlende Messwerte oder Ausreißer auf, die auf den Ausfall des Datenerfassungssystems vor Ort (*Smart Meter*) zurückzuführen sind. Bei der Fernübertragung⁸ der Messdaten vom *Smart Meter Equipment* an den Messdienstleister (vgl. [26]) können ebenfalls fehlerhafte Messwerte generiert werden.

Umgang mit fehlenden Messwerten: Die Anzahl und Verteilung von fehlenden Messwerten und die Größe von zusammenhängenden Regionen mit fehlenden Messwerten entscheidet über den Einsatz der Methoden der Datenvorverarbeitung. Grundsätzlich sind für den Umgang mit fehlenden Messwerten verschiedene Strategien, wie Verfahren der Fallreduktion, Sample-Selection-Modelle, Gewichtungungsverfahren und Imputationsverfahren, ratsam [85–87].

Zu den **Verfahren der Fallreduktion** werden Datenvorverarbeitungsmethoden gezählt, bei denen fehlerhafte Datentupel, Teile von Messzeitreihen oder gesamte Messzeitreihen gelöscht werden, wenn der entsprechende Datensatz fehlende Werte aufweist [86]. **Sample-Selection-Modelle** und **Gewichtungsverfahren** spielen in der vorliegenden Arbeit keine Rolle, werden aber der Vollständigkeit halber aufgeführt. Eine Kurzbeschreibung beider Verfahren ist in [86] zu finden.

Alle Verfahren, die das Ziel verfolgen, fehlende Werte durch geeignete Ersatzwerte zu ersetzen, werden den **Imputationsverfahren** zugeordnet. In der Literatur existieren zahlreiche Imputationsverfahren, deren Beschreibungen und Charakteristiken in Tabelle 1.3 aufgeführt sind. Unter der Vielzahl an Imputationsverfahren zählen die Interpolationsverfahren zu den am häufigsten angewendeten Methoden zum Umgang mit fehlenden Messwerten in Energiezeitreihen (vgl. [76, 97]). Grundsätzlich meint eine Interpolation die Berechnung eines Verhaltens innerhalb eines Bereichs gesicherter Datentupel. Im Gegensatz dazu ermitteln Extrapolationsverfahren das Verhalten über den durch Datentupel gesicherten Bereich hinaus.

Ausreißerbehandlung: Ausreißer stellen neben fehlenden Messwerten eine wichtige Fehlerart in den Rohdaten vieler *Smart-Meter*-Datensätze dar. Im Regelfall ist ein Ausreißer ein Extremwert, der sich von den benachbarten Werten der Zeitreihe stark unterscheidet. Da ausreißerbehaftete Messzeitreihen eine nachfolgende datengetriebene Analyse und Modellierung

⁸Technische Realisierungen der Datenfernübertragung, wie beispielsweise *Powerline Communication* (PLC), werden in [22, 84] diskutiert.

Tabelle 1.3: Übersicht über Verfahren zum Umgang mit fehlenden Werten (Imputationsverfahren, engl. *to impute*: unterstellen) in Energiezeitreihen; Beim DCT-PLS-Verfahren werden die Daten mithilfe einer diskreten Kosinustransformation (DCT) in eine Form überführt, die aus zahlreichen Sinus- und Kosinusfunktionen besteht, und anschließend wird eine Spline-Interpolation durchgeführt (PLS; Abkürzung für *Penalized Least Square*) [88].

Imputationsverfahren	Beschreibung	Literatur
Expertenschätzung	Füllen von Datenlücken durch Expertenwissen; (-) sehr aufwändiges Verfahren, da die Schätzung meist in verbaler Form erfolgt und somit bei einer großen Datenmenge ineffektiv ist	[86]
Mittelwertabschätzung	Fehlende Werte werden mit dem Mittelwert aller zur Verfügung stehenden Messwerte einer Zeitreihe ersetzt; (-) sehr ungenau bei stark variierenden Messwerten	[86]
<i>Nearest Historical Average</i>	Ersetzt fehlende Messwerte einer Zeitreihe mit dem Mittelwert historischer Messwerte, die zur gleichen Tageszeit an benachbarten Tagen gemessen werden	[89]
LOCF, NOCB, NEOI	Verfahren, bei denen der letzte noch vorhandene Wert (<i>Last-Observation-Carried-Forward</i> , LOCF) fortgeschrieben oder der nächste vorhandene Wert (<i>Next-Observation-Carried-Backward</i> , NOCB) „zurückgeschrieben“ wird; Verfahren, bei dem der fehlende Wert durch den nächstgelegenen Nicht-NaN-Wert ersetzt wird (<i>NEarest-Observation-Imputation</i> , NEOI)	[86]
Interpolationsverfahren	Fehlende Werte werden durch lineare Interpolation, Polynominterpolation, Spline-Interpolation (z.B. <i>Piecewise Cubic Hermite Interpolating Polynomial</i> (pchip), DCT-PLS), Stineman-Interpolation bestimmt	[88, 90–92]
Univariate Zeitreihenmodelle	Fehlende Werte werden durch ein univariates Zeitreihenmodell mit definierter Modellstruktur (z.B. AR-Modell (siehe Seite 185) oder ARIMA-Modell) geschätzt. Das Zeitreihenmodell berücksichtigt dabei keine Zusammenhänge zwischen mehreren Messgrößen, sondern nur die Abhängigkeit von einer Messgröße.	[85, 93, 94]
Multivariate Zeitreihenmodelle	Fehlende Werte werden durch ein multivariates Zeitreihenmodell mit definierter Modellstruktur (z.B. ARIMAX-Modell) geschätzt. Hierbei werden auch Abhängigkeiten mehrerer Messgrößen modelliert.	[86, 89, 95, 96]

erheblich beeinflussen, müssen geeignete statistische Verfahren zur Ausreißerdetektion angewandt werden [96, 98, 99]. Eine Vielzahl der Ausreißertests bedient sich statistischer Verteilungsannahmen, anhand derer entschieden wird, ob ein Messwert zu der angenommenen Verteilung gehört oder nicht [85]. Die Anwendung weiterer Ausreißertests, wie z.B. der Q-Test nach Dixon und die Tests nach Grubbs oder Nalimov, wird in [100–103] beschrieben.

1.2.7 Verfahren zur Prognose der Verbraucherlast

Der Nachfrageseite kommt im heutigen und zukünftigen Energiesystem eine wichtige Rolle zu. Die Last einzelner Haushalte bzw. die Last einer Vielzahl von Haushalten gilt es mit geeigneten Verfahren möglichst genau zu prognostizieren, um u.a. das Vorhalten teurer Erzeugungskapazitäten (z.B. thermische Spitzenlastkraftwerke) so gering wie möglich zu halten [18]. Die datengetriebene Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens ist der Kern der vorliegenden Arbeit. Daher werden in Abschnitt 1.2.8 die in der Literatur bekannten Ansätze für die *Demand-Response*-Modellierung vorgestellt. Zunächst soll jedoch in Abschnitt 1.2.7 auf gängige Verfahren zur Prognose der unbeeinflussten Verbraucherlast⁹ eingegangen werden.

⁹Anstatt der Bezeichnung „unbeeinflusste Verbraucherlast“ müsste korrekterweise „nicht durch Preissignale beeinflusste Verbraucherlast“ geschrieben werden. Der Einfachheit halber wird die erstgenannte und kürzere Bezeichnung verwendet.

In der englischsprachigen Literatur sind zwei verwandte Begriffe für die Prognose der unbeeinflussten Verbraucherlast zu finden: *load forecasting/estimation* (Lastprognose) und *Customer Baseline Load estimation* (CBL-Prognose). Da sowohl für die Lastprognose [104–109] als auch für die CBL-Prognose [110–114] zahlreiche Übersichtsarbeiten existieren, sollen beide Verfahren nur kurz vorgestellt und auf die Unterschiede eingegangen werden.

Lastprognose (*Load forecasting/estimation*): Grundsätzlich können Lastprognosen hinsichtlich des Prognosehorizonts in Kurz-, Mittel- und Langfristprognose unterschieden werden, wobei die Einteilung in „Kurzfrist“, „Mittelfrist“ und „Langfrist“ in der Literatur nicht einheitlich ist. In der Regel wird von einer Kurzfristprognose gesprochen, wenn es sich um einen Prognosehorizont von bis zu einer Woche handelt. Bei einem Prognosehorizont im Bereich von einer Woche bis zu einem Jahr kommen Mittelfristprognosen zum Einsatz. Bei Zeiträumen über einem Jahr wird von einer Langfristprognose gesprochen. Eine geeignetere Unterscheidung verschiedener Lastprognoseansätze bietet die Unterscheidung hinsichtlich des verwendeten Verfahrens. Daher werden im Folgenden ausgewählte Verfahren zur Lastprognose vorgestellt.

Das Lastprofil einer Vielzahl von Haushalten kann aggregiert als ein Standardlastprofil (SLP) beschrieben werden [115]. Zur Abschätzung des Stromverbrauchs und des Lastverlaufs werden heute i.d.R. Standardlastprofile verwendet, die vom Bundesverband der Energie- und Wasserwirtschaft (BDEW) bereitgestellt werden. Diese konventionellen (manuellen) Lastprognosemodelle entstehen durch Aggregation der historischen Lastkurven an gleichnamigen Wochentagen, bei denen der Wettereinfluss herausgerechnet und die Erfahrung des Lastverteilerpersonals mit einbezogen wird [18]. Untersuchungen zeigen jedoch, dass einzelne Haushalte große Unterschiede in ihrem Stromverbrauch und Verbrauchsmuster aufweisen [116]. Sowohl die Wertebereiche als auch die Lage von Mindest- bzw. Spitzenlasten können erheblich voneinander abweichen. Verschiedene Einflussfaktoren, wie Anzahl und Alter der Haushaltsbewohner, Wettereinflüsse oder die Geräteausstattung (vgl. Tabelle B.2), beeinflussen das Verbrauchsverhalten und damit das Lastprofil eines Haushaltes [117]. Sogar in vergleichbaren Haushalten unter vergleichbaren Wetterbedingungen sind deutliche Variationen im Verbrauchsmuster und der bezogenen Energie auszumachen [118]. Die zentrale Rolle bei der Ausprägung des Lastprofils eines Haushaltes scheint das individuelle Verbrauchsverhalten der Haushaltsbewohner zu sein [119]. Diese Individualität ist auch bei der Reaktion von Haushaltsstromkunden auf zeitvariable und dynamische Stromtarife auszumachen (vgl. Tabelle 1.2).

In einer Vielzahl von Veröffentlichungen wird die Verbraucherlast mithilfe von *Data-Mining*-Verfahren modelliert, um Prognosemodelle zu erhalten, die die zukünftige Verbraucherlast prognostizieren. Hierbei kann zwischen dynamischen Regressionsmodellen, wie beispielsweise ARIMA- (*AutoRegressive Integrated Moving Average*) [120, 121] oder ARMAX-Modellen (*AutoRegressiv Moving Average model with eXogeneous input*) [122], künstlichen neuronalen Netzen (KNN; engl. *Artificial Neural Network*, ANN) [123–130], *Support Vector Machines* (SVM) [131–133] und *Support Vector Regression* (SVR) [134] unterschieden werden. Eine weitere Modellklasse, welche eng mit den oben genannten ARIMA- bzw. ARMAX-Modellen verwandt ist und häufig für Lastprognosen verwendet wird, ist die Klasse der MA- (*Moving Average*) und ES-Modelle

(*Exponential Smoothing*) [99, 106]. Bei einem reinen MA-Modell fließen nur aktuelle und vergangene Rauschterme in das Prognosemodell ein, wobei mithilfe der MA-Koeffizienten die Gewichtung der Rauschterme vorgenommen wird. Bei den ES-Modellen werden im Vergleich zu den MA-Modellen weiter zurückliegende Vergangenheitswerte schwächer gewichtet. Lastprognosen auf Basis von MA- oder ES-Modellen wurden in [135–139] durchgeführt. In [106] sind zudem zahlreiche Hybridmodelle aus zwei oder mehreren *Data-Mining*-Verfahren beschrieben.

Die unbeeinflusste Verbraucherlast kann zudem mit ökonomischen Modellen geschätzt werden, die den Zusammenhang zwischen Energiebedarf und makro-ökonomischen Variablen, wie z.B. das Einkommen oder das Bruttoinlandsprodukt, schätzen und hauptsächlich für Mittel- und Langfristprognosen eingesetzt werden. Mehr Details zu ökonomischen Modellen liefern [108, 109]. Sogenannte *End-Use*-Modelle schätzen die zukünftige Verbraucherlast durch die Nutzung umfangreicher Informationen über die Verbraucher und die technische Gebäudeausstattung [109, 115].

Ferner kann die unbeeinflusste Verbraucherlast als Signal begriffen werden, welches aus verschiedenen Signalkomponenten aufgebaut ist und mithilfe des additiven Komponentenmodells beschrieben werden kann [56, 140–142]. Das additive Komponentenmodell ist folgendermaßen definiert:

$$P_U[k] = \beta_T \cdot P_{U,T}[k] + \beta_P \cdot P_{U,P}[k] + \beta_M \cdot P_{U,M}[k] + \beta_S \cdot P_{U,S}[k] + \beta_E \cdot P_{U,E}[k], \quad (1.1)$$

wobei P_U die unbeeinflusste Verbraucherlast, $P_{U,T}$ die Trendkomponente, $P_{U,P}$ die periodische Komponente, $P_{U,M}$ die Musterkomponente, $P_{U,S}$ die stochastische Komponente und $P_{U,E}$ die exogene Komponente der unbeeinflussten Verbraucherlast ist. Die Koeffizienten β_T , β_P , β_M , β_S und β_E stellen die Gewichtungsfaktoren der Signalkomponenten dar und müssen in Summe Eins ergeben. Die deterministischen Signalkomponenten (Trendkomponente, periodische Komponente, Musterkomponente) bilden den zeitlichen Einfluss auf die Verbraucherlast ab. Die periodische Signalkomponente setzt sich aus einem Tages-, Wochen- und Jahreszyklus zusammen und wird durch Überlagerung mehrerer Kosinusfunktionen realisiert. Laut [141] treten in realen Signalverläufen typische wiederkehrende Musterverläufe auf, die vom vorangegangenen Signalverlauf oder anderen Einflussgrößen, wie beispielsweise Datum oder Uhrzeit, abhängen. Dieser Signalanteil wird durch die Musterkomponente $P_{U,M}$ abgebildet, die wiederum aus Standardlastprofilen besteht. Der stochastische Anteil der Verbraucherlast wird durch die stochastische Komponente modelliert. Einflüsse der Temperatur, Windgeschwindigkeit oder Solarstrahlung werden im Komponentenmodell durch die exogene Komponente abgebildet. In Abschnitt 2.3.2 wird das additive Komponentenmodell zur Generierung von synthetischen unbeeinflussten Verbraucherlasten eingesetzt.

Abschließend muss an dieser Stelle angemerkt werden, dass sich viele der hier genannten Verfahren (z.B. dynamische Regressionsmodelle, KNNs) auch zur Prognose der beeinflussten Verbraucherlast und somit zur DR-Modellierung (siehe Abschnitt 1.2.8) eignen.

CBL-Prognose (*Customer Baseline Load estimation*): Die CBL-Prognose kann als Unterkategorie der Lastprognose betrachtet werden. Die *Customer Baseline Load* ist definiert als die

Last, die der Stromkunde ohne das DR-Event (Preissignal) gehabt hätte [111]. Eine möglichst genaue Schätzung der CBL ist deshalb wichtig, da die Auswirkung eines DR-Programms bewertet werden und zu Abrechnungszwecken bekannt sein muss. Es muss an dieser Stelle betont werden, dass CBL-Prognoseverfahren nur dann eingesetzt werden können, wenn es sich bei dem DR-Programm um ein event-basiertes DR-Programm, wie z.B. CPP oder PTR (vgl. Tabelle 1.2), handelt. D.h., in den Messdaten der Haushalte müssen sowohl Tage enthalten sein, bei denen die Last durch ein Event-Preissignal beeinflusst ist (sogenannte DR-Eventtage), als auch Tage, bei denen die Last nicht durch ein Event-Preissignal beeinflusst ist (sogenannte Nicht-DR-Eventtage). Zur Schätzung der CBL haben sich in den letzten Jahren zahlreiche Verfahren etabliert, die im Folgenden kurz erläutert werden.

Die *Customer Baseline Load* für einen bestimmten DR-Eventtag wird mithilfe des *HighXofY*-Verfahrens wie folgt berechnet. In einem ersten Schritt wird eine Menge aus Y zurückliegenden Nicht-DR-Eventtagen des gleichen Wochentages gebildet. Aus dieser Menge werden nun die Tage mit den X höchsten Tagesstromverbräuchen ausgewählt. Die mittlere Tageskurve aller X Tage bildet die *Customer Baseline Load*. In analoger Weise wird die *LowXofY*-CBL aus den Tagen mit den X niedrigsten Tagesstromverbräuchen geschätzt. Die *MidXofY*-CBL wird durch Weglassen der Tage mit den jeweils Z niedrigsten bzw. höchsten Tagesstromverbräuchen berechnet. Auch hier wird die CBL durch Mittelung der verbliebenen $X = Y - 2Z$ Tage gebildet. Beispiele für die Berechnung der CBL mithilfe der genannten Verfahren sind in [113, 143–146] zu finden.

Ein weiteres CBL-Prognoseverfahren basiert auf dem Verfahren der exponentiellen Glättung [144]. Für die Schätzung der *Customer Baseline Load* an einem bestimmten DR-Eventtag werden historische Lastkurven der Nicht-DR-Eventtage des gleichen Wochentags gemittelt, wobei weiter zurückliegende Tageslastkurven an Gewichtung verlieren. Mithilfe eines Glättungsfaktors kann festgelegt werden, ob weiter zurückliegende (Glättungsfaktor $\rightarrow 0$) oder aktuellere Tageslastkurven (Glättungsfaktor $\rightarrow 1$) einen größeren Einfluss bei der Berechnung des CBL haben. Allerdings eignet sich dieses Verfahren erst dann zur CBL-Prognose, wenn eine Mindestanzahl an Tageslastkurven zur Verfügung steht. Die genaue Berechnungsvorschrift und Beispiele sind in [113, 144] gegeben.

Die *Customer Baseline Load* kann auch mithilfe eines Regressionsmodells berechnet werden. Dabei wird anhand der Messwerte der Y Nicht-DR-Eventtage des gleichen Wochentags der Zusammenhang zwischen verschiedenen unabhängigen Variablen und der CBL als abhängigen Variable durch ein multiples lineares Regressionsmodell geschätzt. Als mögliche unabhängige Variablen kommen historische Lastdaten, Wetterdaten (z.B. Außentemperatur, Luftfeuchtigkeit) oder Sonnenauf- und -untergangszeiten in Frage [113].

Des Weiteren wird in [113] vorgeschlagen, die *HighXofY*-, *LowXofY*- und *MidXofY*-CBL durch Einbeziehung von Messdaten am DR-Eventtag zu verfeinern. Eine mögliche Abweichung zwischen der geschätzten *Customer Baseline Load* und der tatsächlichen Last kann dann ca. 2-4 Stunden vor dem Event durch additive und multiplikative Korrekturterme korrigiert werden.

1.2.8 Demand-Response-Modellierung

Das ökonomische Potenzial von DSM- und DR-Programmen macht eine umfangreiche und detaillierte Systemuntersuchung notwendig [147]. Eine zentrale Rolle nimmt dabei die Modellierung des preisbeeinflussten Verbrauchsverhaltens einzelner Haushalte bzw. Haushaltsgruppen als Reaktion auf verschiedene Preissignale ein. In der Literatur finden sich zahlreiche Ansätze zur Modellierung des beeinflussten Verbrauchsprozesses als Reaktion auf Preissignale. Allerdings ähneln sich viele Modellansätze, sodass eine Klassifizierung ratsam ist. Die gängigsten Modellklassen, wie

- ökonomische Modelle,
- dynamische Regressionsmodelle,
- Optimierungsmodelle,
- physikalisch motivierte Modelle,
- stochastische Modelle oder
- *Data-Mining*-Modelle,

sollen im Folgenden eingeführt werden, wobei zu jeder Modellklasse ein bis zwei ausgewählte Modellbeispiele vorgestellt werden. Abschnitt 1.2.8 schließt mit einer tabellarischen Zusammenfassung der Modellklassen.

Ökonomische Demand-Response-Modelle: Die Eigenpreiselastizität und Kreuzpreiselastizität¹⁰ dienen in zahlreichen Untersuchungen als einheitsloses Maß für die Beschreibung des preisbeeinflussten Verbrauchsverhaltens. Ganz allgemein gibt die Preiselastizität das Verhältnis zwischen der prozentualen Änderung der nachgefragten Menge und der prozentualen Änderung des Preises an [148] (vgl. Abb. B.1). Diese allgemeingültige, mikroökonomische Theorie lässt sich auch für die Darstellung des preisbeeinflussten Verbrauchsverhaltens von Haushaltsstromkunden übertragen.

Das Verbrauchsverhalten als Reaktion auf variable Strompreise kann im Sinne der ökonomischen *Demand-Response*-Modelle wie folgt zusammengefasst werden [59, 149–151]:

- Eigenpreiselastizität -- Nicht verschiebbare Lasten (z.B. elektrische Beleuchtung) werden infolge eines Preissignals entweder an- oder ausgeschaltet. Die Laständerung $\Delta P[k]$ findet somit zum selben Zeitpunkt k wie die Preisänderung $\Delta p[k]$ statt. Die Eigenpreiselastizität $\varepsilon[k, k]$ berechnet sich gemäß Gleichung (1.2):

$$\varepsilon[k, k] = \frac{p[k]}{P[k]} \cdot \frac{\Delta P[k]}{\Delta p[k]}, \quad (1.2)$$

wobei $\frac{p[k]}{P[k]}$ das Verhältnis von Strompreis $p[k]$ und Last $P[k]$ zum Zeitpunkt k darstellt.

¹⁰In vielen Veröffentlichungen auch Substitutionselastizität (*elasticity of substitution*) genannt.

- Kreuzpreiselastizität -- Verschiebbare Lasten (z.B. Kühl-/ Heizanwendungen) können infolge eines Preissignals tageszeitlich verschoben werden. Eine Preisänderung $\Delta p[j]$ zum Zeitpunkt j führt demnach zu einer Laständerung $\Delta P[k]$ zu einem anderen Zeitpunkt k (*load shifting*). Die Kreuzpreiselastizität $\varepsilon[j, k]$ ist wie folgt definiert:

$$\varepsilon[j, k] = \frac{p[j]}{P[k]} \cdot \frac{\Delta P[k]}{\Delta p[j]}. \quad (1.3)$$

Mit $\varepsilon[j, k]$ kann somit angegeben werden, wie ein variabler Stromtarif tageszeitliche Lastverlagerungen verursacht (siehe Gleichung (B.1)). Die Laständerung als Funktion der Preisänderung kann mit einem linearen oder isoelastischen Ansatz beschrieben werden [150].

Die Modellierung des preisbeeinflussten Verbrauchsverhaltens mit ökonomischen *Demand-Response*-Modellen, die Maße wie die Kreuzpreiselastizität und *a-priori*-Annahmen über den funktionalen Zusammenhang zwischen Preis und Last beinhalten, wurde z.B. in [61, 68, 149, 150, 152–156] durchgeführt.

Ökonomische Modelle können, wie in [157] gezeigt, mit Anreizgewichtungsfaktoren ergänzt werden. Diese Faktoren hängen von der Art des DSM-Programms (IBP oder PBP), der Größe des zu beeinflussenden Verbrauchers, dem Sektor, der Höhe des Einkommens des Stromkunden und von weiteren sozialen und kulturellen Aspekten ab. Dies setzt allerdings eine umfassende Untersuchung der beteiligten Stromkunden voraus.

Dynamische Regressionsmodelle: Dynamische Regressionsmodelle kommen bei der Modellierung des preisbeeinflussten Verbrauchsverhaltens von Haushaltsstromkunden zum Einsatz. Modellansätze mit Vergangenheitswerten der Ausgangsgröße werden dabei als autoregressive Modelle bezeichnet [85]. So wurde in [158] ein *Finite Impulse Response* (FIR) Modell zur Schätzung der Laständerung als Reaktion auf Preisänderungen und auf Änderungen einer externen Größe (Wetter) verwendet. Zur Schätzung der Parameter des FIR-Prozesses wurde ein rekursives Verfahren der kleinsten Fehlerquadrate (*Recursive Least Squares*, RLS) angewandt.

Wie in [159] dargestellt, eignen sich in Ergänzung zu dem in [158] gezeigten Verfahren, NFIR- (*Nonlinear Finite Impulse Response*), ARX- (*AutoRegressive model with eXogeneous input*) und ARMAX- (*AutoRegressiv Moving Average model with eXogeneous input*) Modelle zur Schätzung des Zusammenhangs zwischen Preissignal und Verbraucherlast. In [159] ist ein ARX(n_a, n_b, n_k)-Modell zur Schätzung des Zusammenhangs zwischen Preissignal $p[k]$ als Eingangsgröße und preisbeeinflusstem Anteil der Verbraucherlast $P_R[k]$ als Ausgangsgröße wie folgt definiert:

$$\hat{P}_R[k] + \sum_{i=1}^{n_a} a_i P_R[k-i] = \sum_{j=1}^{n_b} b_j p[k-n_k-j+1] + e[k], \quad (1.4)$$

wobei n_a, n_b die Modellordnungen des ARX-Modells, n_k die Totzeit des ARX-Modells und $e[k]$ ein gaußverteilt weißes Rauschen bezeichnen. Die Parameter (a_i, b_j) dieser Modelle werden in der Regel für komplette Tarifgruppen und längere Zeiträume geschätzt.

Optimierungsmodelle: Der in [160] vorgestellte Modellierungsansatz verwendet einen einfachen LP-Algorithmus (Lineare Programmierung), der in einem *Energy Management System* (EMS) implementiert wird und den Stromverbrauch eines Haushaltes unter Berücksichtigung eines Echtzeit-Strompreises optimiert. Es wird vorausgesetzt, dass zwischen Stromkunde und Stromversorger eine bidirektionale Kommunikation besteht. Der Stromkunde empfängt stündlich eine Preisinformation, die jeweils für die nächste Stunde gilt. Ein ARMA- (*AutoRegressiv Moving Average model*) Modell schätzt aus der Preishistorie den Verlauf des zukünftigen Preises. Des Weiteren gibt der Stromkunde die Randbedingungen für die Optimierung vor, wie den Gesamtstromverbrauch des Tages und minimale/ maximale Laständerungen. Das Ergebnis der Optimierungsrechnung, sprich optimaler stündlicher Verbrauch und optimaler Lastverlauf, wird wiederum an den Stromversorger geschickt.

Weitere Modellansätze, die den Optimierungsmethoden zuzuschreiben sind, sind beispielsweise die stochastische Optimierung [161], die konvexe Programmierung [162], die heuristische Optimierung [163], Chance-Constrained-Optimierung [158], spieltheoretische Optimierungsverfahren [164–166] oder evolutionäre Algorithmen [167, 168]. Eine Übersicht über DSM-Optimierungsmodelle liefern [41, 169].

Modellierung mit Markov-Ketten (stochastische Modelle): Ein weiteres Verfahren nutzt Methoden des maschinellen Lernens (Reinforcement-Lernen, engl. *reinforcement learning*) zur Beschreibung des preisbeeinflussten Verbrauchsverhaltens. [170] schlägt vor, An- und Abschaltwünsche bei elektrischen Haushaltsgeräten und den zukünftigen Strompreis als Markov-Ketten mit unbekanntem Übergangswahrscheinlichkeiten zu modellieren. Im *Consumer Automated Energy Management System* (CAES) laufen beide Informationen zusammen und der CAES-Algorithmus berechnet daraus einen optimalen, sprich kostengünstigen, Fahrplan für die gewählten Haushaltsgeräte.

In [171] wurde ein Simulationsmodell vorgestellt, welches die Auswirkungen von Strompreissignalen und CO₂-Signalen in einem DR-Programm simuliert. Das Simulationsmodell besteht aus einem Markov-Ketten-Verbrauchermodell und einem Scheduling-Programm, das optimale Fahrpläne für intelligente Haushaltsgeräte auf Basis der Preis- und CO₂-Signale ausgibt.

Der Belade- und Entladefahrplan von Elektrofahrzeugen aus dem bzw. in das Stromnetz (engl. *vehicle-to-grid*, V2G) unter Berücksichtigung von Preisunsicherheiten wurde in [172] mit ähnlichen Methoden modelliert.

Physikalisch motivierte Modelle: Physikalisch motivierte Modelle bilden als *Bottom-Up*-Modelle¹¹ das Verhalten einzelner Haushaltsgeräte im Betrieb ab. In [173] handelt es sich dabei um Geräte zur elektrischen Warmwasserbereitung, Luftheizungsanlagen, Wäschetrockner und Elektrofahrzeuge. Der Fokus liegt nicht in der Modellierung der Reaktion des Haushaltsstromkunden, sondern in der Modellierung des physikalischen Verhaltens verschiedener Geräte. Das

¹¹In manchen Veröffentlichungen werden diese Modelle auch *End-Use*-Modelle genannt (vgl. Abschnitt 1.2.7) [109, 115].

DR-Signal geht als *on-/off*-Faktor in die Berechnung ein. Des Weiteren werden die einzelnen Modelle mithilfe stochastischer Methoden zu Gesamtmodellen aggregiert. Ähnliche Modelle sind in [174, 175] beschrieben.

Clusterverfahren zur Bestimmung von typischen Lastprofilen: Ein *Data-Mining*-Verfahren, das sich grundsätzlich zur Analyse von Haushaltslastdaten eignet, ist das Clustering. Das Ziel des Clustering ist das Auffinden von Gruppen ähnlicher Lastzeitreihen in großen Datensätzen (z.B. *Smart-Meter*-Datensatz), die als Typtage¹² interpretiert werden können. Die Grundlagen des Clustering werden umfangreich in [176, 177] behandelt. Eine sehr gute Übersicht zum Thema Zeitreihen-Clustering liefert [178].

In der Literatur existieren zudem viele Veröffentlichungen, die sich mit dem Clustering von Haushaltslastdaten befassen, um daraus ein typisches Verbrauchsverhalten abzuleiten [179–181]. Das Zeitreihen-Clustering von Lastdaten kalifornischer Haushalte wurde in [182] durchgeführt. Das Clustering von *Smart-Meter*-Daten mit ähnlichen Ansätzen ist in [183–192] gezeigt. Zudem wurde in [187] nach der Clusteranalyse untersucht, ob ein Zusammenhang zwischen den gefundenen Typtagen und sozioökonomischen Faktoren besteht. Eine Übersicht über verschiedene Clusterverfahren (z.B. hierarchisches Clustering, C-Means-Clustering, Fuzzy-C-Means-Clustering, Support-Vector-Clustering oder Self-Organizing Maps) zur Erzeugung typischer Lastprofile liefern [193–195].

In allen hier genannten Veröffentlichungen handelte es sich bei den untersuchten Haushalten um Haushalte, die keine Beeinflussung durch variable Stromtarife erfahren haben. D.h., es wurden keine typischen preisbeeinflussten Lastprofile generiert, sondern typische Lastprofile.

Eine Übersicht zum Forschungsstand auf dem Gebiet der Modellierung des preisbeeinflussten Verbrauchsverhaltens bei Haushaltsstromkunden zeigt Tabelle 1.4.

1.3 Ungelöste Probleme und offene Fragen

Aus den vorangegangenen Abschnitten ergeben sich für die Modellierung des preisbeeinflussten Verbrauchsverhaltens folgende offene Probleme:

1. *Systematisierung der Datenstruktur von Smart-Meter-Datensätzen:*

Die meisten der vorgestellten Modellierungskonzepte beruhen auf Modellstrukturannahmen, die anhand von realen Messdaten getroffen bzw. mit ihnen validiert werden. Allerdings existiert keine Anleitung zur systematischen Beschreibung dieser Messdaten, d.h. zur Systematisierung der Datenstruktur von *Smart-Meter*-Datensätzen.

¹²Anstatt des Begriffes „Typtage“ kann auch die Bezeichnung „typische Lastprofile“ verwendet werden. Stammen die Lastdaten von Haushalten, welche eine Beeinflussung durch variable Strompreise erfahren haben, muss von „preisbeeinflussten Typtagen“ bzw. von „typischen preisbeeinflussten Lastprofilen“ gesprochen werden.

Tabelle 1.4: Übersicht zum Forschungsstand auf dem Gebiet der Modellierung des preisbeeinflussten Verbrauchsverhaltens bei Haushaltsstromkunden

Modellklasse	Kurzbeschreibung, Charakteristiken	Anforderungen, Einschränkungen	Literatur
ökonomische Modelle	Prinzip der Preiselastizität	eher „statische“ als „dynamische“ Modelle, spiegelt nicht das Zeitverhalten eines Haushaltsstromkunden wider	[68, 149, 150, 152–154, 157, 196–201]
dynamische Regressionsmodelle	FIR-, NFIR-, ARX-, ARMAX-Modelle	komplexe Modellstrukturen mit einer hohen Modellordnung möglich	[158, 159, 202]
Optimierungsmodelle	Lineare Programmierung, konvexe Programmierung, stochastische Optimierung, heuristische Optimierung, <i>Chance-Constrained</i> -Optimierung, spieltheoretische Optimierung	spiegeln nicht das reale Verbrauchsverhalten als Reaktion auf Preissignale wider; nur in Haushalten mit Energiemanagementsystemen und automatischem DR anwendbar	[158, 160–164, 166, 203–207]
stochastische Modelle	Markov-Ketten mit unbekanntem Übergangswahrscheinlichkeiten	spiegeln nicht das reale Verbrauchsverhalten als Reaktion auf Preissignale wider	[170–172, 208]
physikalisch motivierte Modelle	<i>Bottom-up</i> -Modelle; Modelle der Haushaltslasten auf Geräteebe; Aggregation der Teilmodelle, um Lastprofile zu erhalten	viele Teilmodelle auf Geräteebe; umfangreicher Aufwand bei der Parameterschätzung; aufwändige Datenerfassung	[173–175]
Clustermodelle	<i>Data-Mining</i> -Verfahren; hierarchisches Clustering, <i>C-Means</i> -Clustering, <i>Fuzzy-C-Means</i> -Clustering, <i>Support-Vector</i> -Clustering, <i>Self-Organizing Maps</i> ; Clustering von <i>Smart-Meter</i> -Datensätzen	Clusterkurven als Ergebnis des Clusterings haben einen beschreibenden Charakter und sind als Prognosemodelle ungeeignet	[76, 179–181, 183–190, 192–195, 209]

2. Ganzheitliches Konzept zur Strukturierung und Analyse von *Smart-Meter*-Datensätzen:

Wie bereits formuliert, dienen reale Messdaten der Parametrierung von DR-Modellstrukturannahmen. Da *Smart-Meter*-Datensätze fehlerhafte Rohdaten enthalten, müssen vor dem Modellentwurf zahlreiche Vorverarbeitungsschritte durchlaufen werden. Es existiert bislang keine Anleitung zum Umgang mit fehlerbehafteten Messdaten, die das Verbrauchsverhalten von Haushaltsstromkunden als Reaktion auf Preissignale abbilden. Des Weiteren lassen alle genannten DR-Modellierungskonzepte eine umfassende Analyse der Auswirkung verschiedener Preissignale auf das Verbrauchsverhalten vermissen.

Der Fokus des Clusterings von *Smart-Meter*-Daten lag bisher in der Generierung von typischen Lastprofilen, die das unbeeinflusste Verbrauchsverhalten von Haushalten abbilden. Es fehlt bisher ein Ansatz, um aus preisbeeinflussten Lastdaten typische preisbeeinflusste Lastprofile zu erzeugen.

3. Haushaltsindividuelle, datengetriebene Modellierung des beeinflussten Verbrauchsprozesses:

Bei *bottom-up* Modellen sind eine Fülle an Informationen für den Modellentwurf notwendig, die nicht immer zur Verfügung stehen und darum nicht zufriedenstellende Prognoseergebnisse liefern. Die Modellierung des preisbeeinflussten Verbrauchsverhaltens von Haushaltsstromkunden sollte ohne Kenntnis der Geräteausstattung oder anderer datenschutzsensibler Informationen möglich sein. Unterschiedliche Stromtarifdesigns werden in den Literaturmodellen bisher nur unzureichend berücksichtigt.

Wie bereits auf Seite 20 angesprochen, wird das preisbeeinflusste Verbrauchsverhalten in zahlreichen Veröffentlichungen durch Optimierungsmodelle beschrieben. In der Regel führt das *Energy Management System* des Haushaltes eine Optimierung durch, die eine

optimale Lastkurve (optimaler Lastfahrplan) als Ergebnis hat. Allerdings kann dies nur in Haushalten erfolgen, in denen steuer- und programmierbare Haushaltsgeräte vorhanden sind, die die Haushaltslast gemäß der berechneten optimalen Lastkurve „abfahren“. Das reale Verbrauchsverhalten von Haushalten als Reaktion auf Preissignale (manuelles *Demand Response*) kann mithilfe von Optimierungsmodellen nicht abgebildet werden.

4. *Systemtechnische Einordnung und Simulation des beeinflussten Verbrauchsprozesses in einem rückgekoppelten System:*
DR-Maßnahmen finden i.d.R. in Verteilnetzen der Niederspannungsebene statt und unterliegen verschiedenen Interessen und Einflüssen. Es fehlt bislang die systemtechnische Einordnung des preisbeeinflussten Verbrauchsprozesses in ein Smart Grid, die sowohl Lastprognosen, unterschiedliche Stromtarifdesigns, als auch Rückkopplungen innerhalb des Systems „Smart Grid“ berücksichtigt.

1.4 Ziele und Aufgaben

Das Ziel dieser Arbeit ist die Entwicklung eines neuen, gesamtheitlichen Konzeptes zur datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsprozesses im Umfeld eines Smart Grids. Das neue Konzept erfordert folgende Entwicklungsschritte:

1. *Generalisierte Beschreibung und Strukturierung von Smart-Meter-Datensätzen:*
Die Analyse des beeinflussten Verbrauchsprozesses auf der Basis realer Messdaten erfordert die Aufbereitung und Strukturierung der zugrunde liegenden Rohdaten. Die Rohdaten durchlaufen zahlreiche Vorverarbeitungs-, Aggregierungs-, Normierungs- und Mittelungsschritte, die derart kombiniert werden, dass am Ende eine systematisierte Beschreibung des *Smart-Meter*-Datensatzes möglich ist („Checkliste“).
2. *Datengetriebene Analyse des preisbeeinflussten Verbrauchsprozesses:*
Die Analyse eines strukturierten *Smart-Meter*-Datensatzes stellt einen wichtigen Baustein für das gesamtheitliche Konzept dar. Die Anwendbarkeit der Analysemethoden richtet sich nach der Beschaffenheit des *Smart-Meter*-Datensatzes.
3. *Entwicklung eines neuen Konzeptes zur Modellierung des Verbrauchsverhaltens von Haushaltsstromkunden als Reaktion auf verschiedene Preissignale:*
Es sollen Modelle erarbeitet werden, welche eine genaue Prognose für Einzelhaushalte oder für aggregierte Haushaltsgruppen unter Berücksichtigung von verbrauchsbeeinflussenden Preissignalen ermöglichen. Die Ergebnisse des Analyseprozesses sollen in die parametrierbaren Modelle integriert werden und die interne Modellstruktur soll verständlich sein (*Gray-Box*-Modell). Des Weiteren soll durch die Modellstruktur und eine relativ geringe Anzahl an zu schätzenden Modellparametern die Interpretierbarkeit der DR-Modelle gewährleistet werden.
4. *Validierung der Modelle und Modellbenchmark:*
Zum Nachweis der Funktionalität des neuen Modellkonzeptes sollen die entwickelten DR-Modelle mittels realer *Smart-Meter*-Daten validiert werden. Des Weiteren ist der Nachweis

zu erbringen, dass die Prognosegüte der neuen DR-Modelle gleichwertig gegenüber oder besser als bei etablierten, schlechter interpretierbaren DR-Modellen ist.

Kapitel 2 stellt das neue, gesamtheitliche Konzept zur datengetriebenen Analyse und Modellierung des beeinflussten Verbrauchsprozesses im Umfeld eines *Smart Grids* vor. Die generalisierte Beschreibung von *Smart-Meter*-Datensätzen wird in Abschnitt 2.2 vorgenommen und dient als Basis für die datengetriebene Analyse der Auswirkungen von variablen Stromtarifen auf das Verbrauchsverhalten von Haushaltsstromkunden (Abschnitt 2.5). Das neue Konzept zur Modellierung des preisbeeinflussten Verbrauchsverhaltens wird in Abschnitt 2.6 vorgestellt. Die konkrete Umsetzung des neuen Konzepts erfolgt mithilfe der grafischen Benutzeroberfläche der MATLAB Toolbox SciXMiner. Die Entwicklungsumgebung wird ausführlich in Kapitel 3 vorgestellt. Kapitel 4 widmet sich der Anwendung der in Kapitel 2 eingeführten neuen Verfahren an zwei *Smart-Meter*-Datensätzen (*Olympic Peninsula Project*, RESIDENS). Die Arbeit schließt mit einer Zusammenfassung und einem Ausblick in Kapitel 5.

2 Neues Konzept zur datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens

Kapitel 2 befasst sich mit einem neuen Konzept zur Untersuchung und Beschreibung des Verbrauchsverhaltens von Haushaltsstromkunden als Reaktion auf Preissignale. Um die Rolle des Haushaltsstromkunden im Gesamtsystem *Smart Grid* zu bewerten, wird in Abschnitt 2.1 eine systemtechnische Beschreibung des preisbeeinflussten Verbrauchsprozesses und dessen Einordnung im System *Smart Grid* vorgenommen. Hierbei soll sowohl auf die Planung des Preissignals, als auch auf die Verbrauchsbeeinflussung im rückgekoppelten System eingegangen werden. Um eine erfolgreiche Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens zu gewährleisten, ergeben sich im Voraus spezifische Anforderungen an die Eigenschaften des zugrunde liegenden *Smart-Meter*-Datensatzes (SMD), die in Abschnitt 2.2 erläutert und in zwei generalisierten Checklisten zusammengefasst werden. Mithilfe der beiden Checklisten soll der zu untersuchende *Smart-Meter*-Datensatz dahingehend bewertet werden, ob und inwieweit die Eigenschaften des SMD eine konkrete Aufgabenstellung ermöglichen bzw. begünstigen. Zur Demonstration der Wirksamkeit des neuen Konzeptes und zur Validierung der einzelnen Methoden werden in Abschnitt 2.3.2 auf Basis der Checklisten zwölf synthetische *Smart-Meter*-Benchmarkdatensätze definiert. Im neuen Konzept zur datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens werden die einzelnen Methoden derart miteinander kombiniert, dass von einer Methodenkette oder -pipeline¹ gesprochen werden kann. In Abb. 2.1 ist in einem Prozessschaubild die Abfolge der einzelnen Methoden im neuen Konzept zur datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens dargestellt.

Die Methoden der Pipeline lassen sich hinsichtlich der Zielsetzung in

- analyseorientierte Methoden (Abschnitt 2.5) und
- prognoseorientierte Methoden (Abschnitt 2.6)

zusammenfassen. Die analyseorientierten Methoden der Pipeline dienen dazu, qualitative und quantitative Aussagen über den untersuchten *Smart-Meter*-Datensatz zu machen. So soll mithilfe der analyseorientierten Methodenpipeline beispielsweise die Frage geklärt werden, inwieweit haushaltsindividuelle Verbrauchsunterschiede als Reaktion auf eine DR-Maßnahme auftreten. Die prognoseorientierten Methoden der Pipeline haben die Erstellung datenbasierter

¹Der Begriff Methodenpipeline meint in diesem Zusammenhang die Aneinanderreihung bestimmter Verarbeitungsschritte und Methoden.

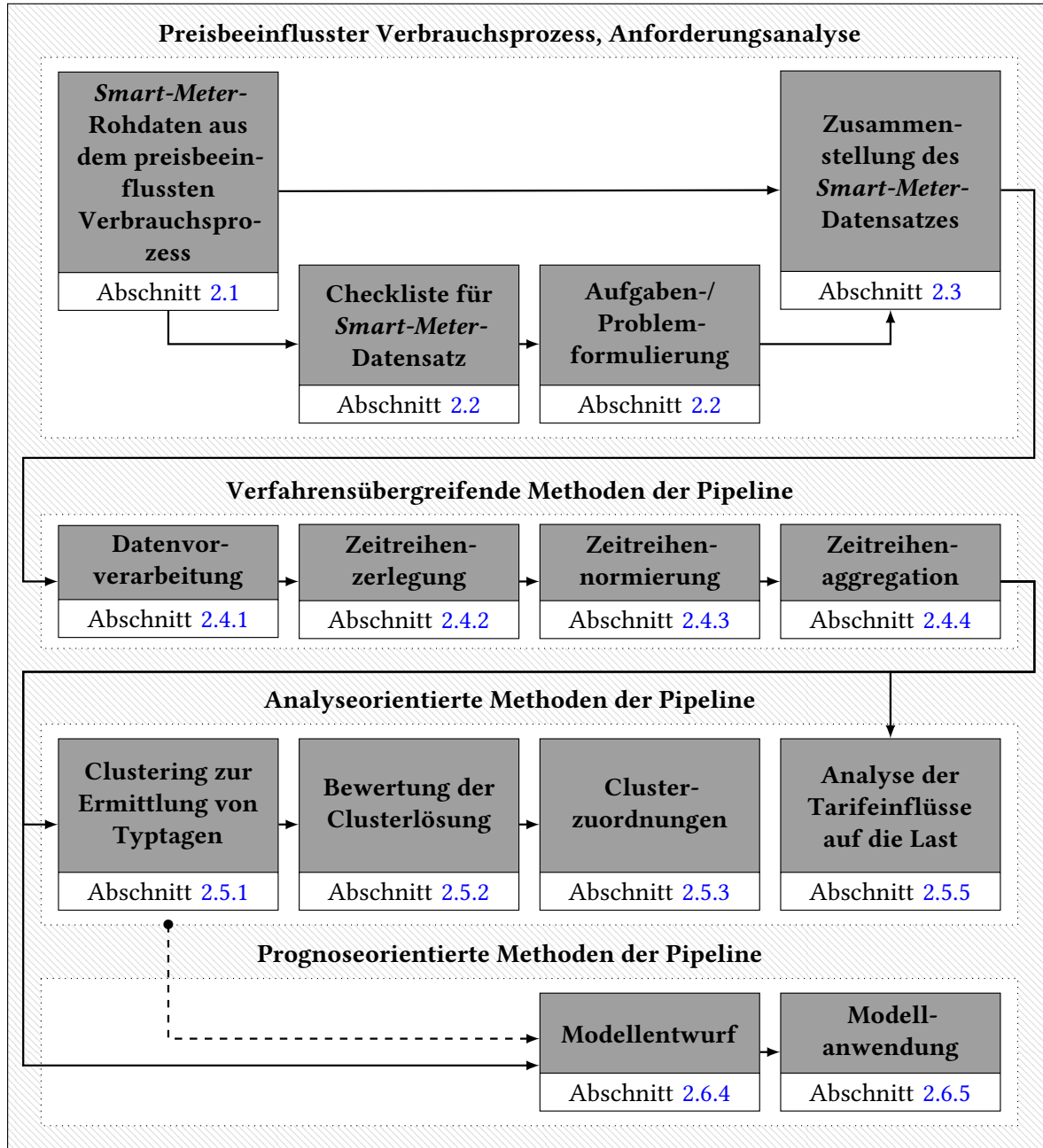


Abbildung 2.1: Struktur der Methodenpipeline des neuen Konzeptes zur datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens

Prognosemodelle zur Zielsetzung, welche eine Prognose für Einzelhaushalte oder für aggregierte Haushaltsgruppen unter Berücksichtigung von verbrauchsbeeinflussenden Preissignalen ermöglichen sollen.

Unabhängig von der gewählten Zielsetzung beinhaltet die Gesamtpipeline Methoden, die sowohl bei der analyseorientierten als auch bei der prognoseorientierten Pipeline zum Einsatz kommen und in Abschnitt 2.4 vorgestellt werden. Die verfahrensübergreifenden Methoden umfassen dabei wichtige Verarbeitungsschritte, die klassischerweise der Datenvorverarbeitung zugeschrieben werden [85]. Weitere Methoden führen zur systematischen Strukturierung des *Smart-Meter*-Datensatzes, was die nachfolgende Analyse und datengetriebene Modellierung des preisbeeinflussten Verbrauchsverhaltens begünstigt.

Der gestrichelte Verknüpfungspfeil zwischen der analyse- und prognoseorientierter Pipeline in Abb. 2.1 soll deutlich machen, dass die Analyseergebnisse in die Erstellung der Prognosemodelle einfließen. Dazu wird in Abschnitt 2.6.2 auf das *a-priori*-Wissen eingegangen, welches zur Erstellung der Prognosemodelle verwendet wird.

2.1 Preisbeeinflusster Verbrauchsprozess mit Systemrückkopplung

Die Beeinflussung von Stromkunden durch Preissignale (allg.: Flexibilisierung der Nachfrage) führt dazu, dass ein zuvor bekanntes Verbrauchsverhalten einzelner Stromkunden verändert wird und sich so eine veränderte Elektrizitätsnachfrage insgesamt einstellt. Die Beeinflussung kann zum Beispiel dadurch motiviert sein, dass kurzfristige Schwankungen auf der Erzeugerseite durch eine veränderte Nachfrage ausgeglichen werden müssen, was somit zur Gewährleistung der Systemstabilität (vgl. Abschnitt 1.2.1) beiträgt. Allerdings ist eine isolierte Betrachtung der Beeinflussung von Stromkunden durch Preissignale unzureichend. Der Stromkunde reagiert zwar auf ein bestimmtes Preissignal, jedoch führt diese Reaktion im Stromnetz zu einer veränderten Nachfragesituation, die der Netzbetreiber bei der Synchronisation zwischen Stromerzeugung und Stromverbrauch berücksichtigen muss, was sich wiederum auf die Planung des neuen Preissignals auswirkt. Zusammengefasst lässt sich sagen, dass es sich bei der Beeinflussung von Stromkunden durch Preissignale um ein rückgekoppeltes System handelt.

Für das neue Konzept zur datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens ist eine systemtechnische Einordnung des *Demand-Response*-Prozesses ratsam [56]. Abb. 2.2 zeigt in einem Prozessschaubild wie die **Planung des Preissignals** (Block A), der **beeinflusste Verbrauchsprozess** (Block B) und das **Prognosemodell des Gesamtverbrauchs** (Block C) interagieren. Im Folgenden werden die Inhalte der einzelnen Blöcke und deren Wechselwirkungen kurz dargestellt.

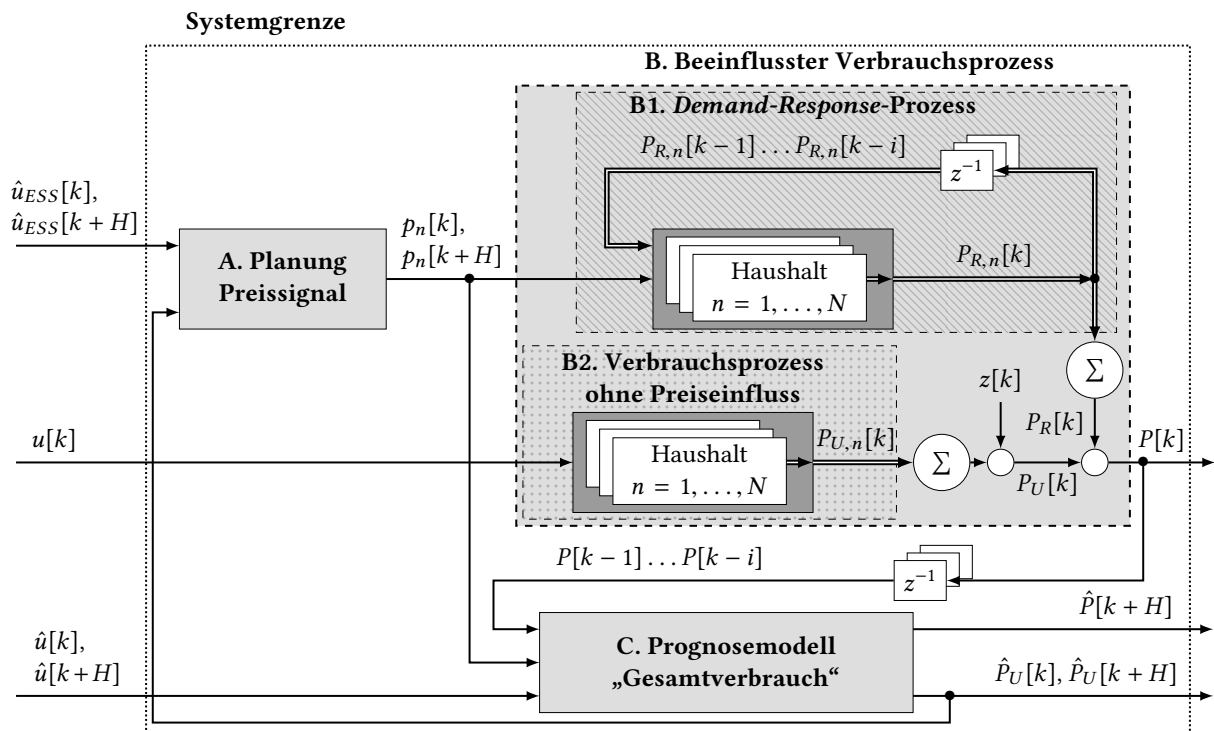


Abbildung 2.2: Systemtechnische Einordnung des *Demand-Response*-Prozesses: Beeinflusster Verbrauchsprozess mit Prognosemodell und Planung (modifiziert nach [56, 210]): H - Vorhersagehorizont, n - Haushalts-/Verbraucherindex, $p[k]$ - Preissignal, \hat{x} - Schätzung für x , $u[k]$ - Eingangsgrößen, $u_{ESS}[k]$ - exogene Einflussgrößen des elektrischen Energiesystems, $P[k]$ - beeinflusste Verbraucherlast, $P_R[k]$ - beeinflusster Anteil der Verbraucherlast, $P_U[k]$ - unbeeinflusster Anteil der Verbraucherlast, $z[k]$ - Störgröße, z^{-1} Verschiebeoperator.

A. Planung des Preissignals

Das Energieversorgungsunternehmen (z.B. der Regelzonenbetreiber), welches die Planung des Preissignals $p_n[k]$ vornimmt, berücksichtigt sowohl die Prognose der exogenen Einflussgrößen des elektrischen Energiesystems $\hat{u}_{ESS}[k]$, als auch die Prognose des unbeeinflussten Anteils der Verbraucherlast $\hat{P}_U[k]$. Der funktionale Zusammenhang der drei Größen kann wie folgt dargestellt werden:

$$p_n[k] = f(\hat{u}_{ESS}[k], \hat{P}_U[k]). \quad (2.1)$$

Die exogenen Einflussgrößen des elektrischen Energiesystems sind vielfältiger Natur und lassen sich in

- meteorologische Einflussgrößen (Wetterprognosen),
- energiewirtschaftliche Rahmenbedingungen (Börse, OTC²-Markt, Risiken, Beschaffungs-/Anlagenkosten, steuerliche Rahmenbedingungen) und

²Abkürzung für *over the counter*: finanzielle Transaktionen zwischen Teilnehmern am Finanzmarkt, die nicht über die Börse abgewickelt werden

- energietechnische Rahmenbedingungen (Grund- und Regellast, Verfügbarkeiten, Anlagenrestriktionen, Speichereinsatz, Fahrpläne im Bilanzraum³)

einteilen [211]. Nur durch die Berücksichtigung aller Prognosen der Einflussgrößen kann ein optimaler Betrieb des elektrischen Energieversorgungsnetzes im Sinne einer sicheren und kostenoptimalen Energieversorgung gewährleistet werden. Für die Planung des Preissignals ist zudem die Kenntnis des ungestörten Netzzustandes und somit die Kenntnis des unbeeinflussten Anteils der Verbraucherlast $P_U[k]$ im Bilanzraum von entscheidender Bedeutung. Die Planung des Preissignals erfolgt unter Zuhilfenahme einer Lastflussberechnung⁴, die Auskunft über den optimalen Betriebszustand des elektrischen Energieversorgungsnetzes gibt. Einen gesamtheitlichen Ansatz zur Planung eines optimalen Preissignals mithilfe einer Lastflussberechnung und Kostenoptimierung liefert [213].

B. Beeinflusster Verbrauchsprozess

Der beeinflusste Verbrauchsprozess setzt sich aus zwei Teilprozessen zusammen: dem unbeeinflussten Verbrauchsprozess und dem *Demand-Response*-Prozess.

Der **unbeeinflusste Verbrauchsprozess**⁵ (Block B2 in Abb. 2.2) beschreibt den Anteil der Haushaltslast $P_{U,n}[k]$, welcher nicht durch Preissignale, wohl aber durch haushaltsindividuelle, soziodemografische, meteorologische, tageszeitabhängige, saisonale und singuläre Faktoren beeinflusst ist. In der Literatur existieren zahlreiche Veröffentlichungen, die sich mit der Evaluation dieser Einflussfaktoren beschäftigen [17, 58, 117, 214]. Die Einflussfaktoren auf $P_{U,n}[k]$ lassen sich als $u[k]$ zusammenfassen:

$$u[k] = f(u_{sd}[k], u_n[k], u_m[k], u_t[k], u_s[k]), \quad (2.2)$$

wobei $u_{sd}[k]$ soziodemografische Einflüsse, wie den sozialen Status, Einkommen, Anzahl und Alter der Haushaltsbewohner, beschreibt. $u_n[k]$ stellen Einflüsse auf die Haushaltslast dar, die auf die Größe, Lage, Alter und technische Ausstattung des Gebäudes zurückzuführen sind (vgl. Abschnitt 1.2.5 und Tabelle B.2). Meteorologische Faktoren werden mit $u_m[k]$, tageszeitabhängige und saisonale Faktoren mit $u_t[k]$ abgebildet. Singuläre Faktoren wie Sonderereignisse, die die Haushaltslast in besonderem Maße beeinflussen, und mehr oder weniger zufällige Ereignisse (z.B. spontane Aktivität außer Haus), sind in Gleichung (2.2) mit $u_s[k]$ definiert. Wird als Systemgrenze ein räumlich begrenztes Versorgungsgebiet angenommen, werden die unbeeinflussten Anteile der Haushaltslasten $P_{U,n}[k]$ aller N Verbraucher aufsummiert und als unbeeinflusster Anteil der Verbraucherlast $P_U[k]$ für ein Versorgungsgebiet angegeben. Der mittlere (oder

³Ein Bilanzraum (regelungstechnische Sichtweise) ist hier nicht mit Bilanzkreis (energiewirtschaftliche Sichtweise) gleichzusetzen. Ein Bilanzraum ist als einen abstrakten Raum zu verstehen, der durch eine Systemgrenze von seiner Umgebung abgegrenzt ist.

⁴Bei der Lastflussberechnung werden Werte der Knotenspannungen im stationären Zustand und davon ausgehend die Ströme und Leistungen im Netz ermittelt [212].

⁵Anstatt der Bezeichnung „unbeeinflusster Anteil der Verbraucherlast“ müsste korrekterweise „nicht durch Preissignale beeinflusster Anteil der Verbraucherlast“ geschrieben werden. Der Einfachheit halber wird in der vorliegenden Arbeit die erstgenannte und kürzere Bezeichnung verwendet.

aggregierte) unbeeinflusste Anteil der Verbraucherlast für ein Versorgungsgebiet ist demnach als

$$\bar{P}_U[k] = \frac{1}{N} \sum_{n=1}^N P_{U,n}[k] \quad (2.3)$$

definiert.

Der **Demand-Response-Prozess** (Block B1 in Abb. 2.2) ist die haushaltsindividuelle Verbrauchsreaktion auf ein Preissignal (vgl. Strategien in Abb. 1.4) und beschreibt den Anteil der Haushaltslast $P_{R,n}[k]$, der ausschließlich auf den Einfluss des Preissignals $p_n[k]$ zurückzuführen ist, was wie folgt definiert ist:

$$P_{R,n}[k] = f(p_n[k]). \quad (2.4)$$

Bei der Betrachtung von Gleichung (2.4) fällt auf, dass keine zusätzlichen Abhängigkeiten (z.B. zwischen $P_{R,n}[k]$ und $u[k]$) berücksichtigt werden. Es handelt sich hierbei um eine vereinfachende Annahme, um die Modellierung des preisbeeinflussten Verbrauchsverhaltens zu erleichtern.

Aus Gleichung (2.4) wird

$$\bar{P}_R^{T_i}[k] = \frac{1}{N_{T_i}} \sum_{n=1}^{N_{T_i}} P_{R,n}[k] = f(p^{T_i}[k]), \quad (2.5)$$

wenn der aggregierte beeinflusste Anteil der Verbraucherlast aller N_{T_i} Haushalte einer Tarifgruppe T_i betrachtet wird und

$$\bar{P}_R[k] = \frac{1}{N} \sum_{n=1}^N P_{R,n}[k] = f(p[k]), \quad (2.6)$$

wenn der aggregierte beeinflusste Anteil der Verbraucherlast aller N Haushalte betrachtet wird.

Die Rückführung der historischen Werte der Verbrauchsreaktion $P_{R,n}[k-1] \dots P_{R,n}[k-i]$ in Abb. 2.2 soll verdeutlichen, dass eine aktuelle Verbrauchsreaktion $P_{R,n}[k]$ nicht losgelöst von bereits erfolgten Laständerungen betrachtet werden kann. Analog zum unbeeinflussten Verbrauchsprozess werden die beeinflussten Anteile der Verbraucherlast $P_{R,n}[k]$ aller N Haushalte aufsummiert und als beeinflusster Anteil der Verbraucherlast $P_R[k]$ angegeben. Der mittlere (oder aggregierte) beeinflusste Anteil der Verbraucherlast für ein Versorgungsgebiet ist demnach durch Gleichung (2.6) definiert.

Die Summation des unbeeinflussten und beeinflussten Anteils der Verbraucherlast ergibt die beeinflusste Verbraucherlast:

$$P_n[k] = P_{U,n}[k] + P_{R,n}[k] \quad \text{für Einzelhaushalte } n \quad (2.7)$$

Die aggregierte beeinflusste Verbraucherlast ergibt sich direkt aus Gleichung (2.7) durch Addition der Summen der entsprechenden Verbrauchsanteile und Division mit der Anzahl an

betrachteten Haushalten, wie in den Gleichungen (2.5) und (2.6) beispielhaft gezeigt. Die aggregierte beeinflusste Verbraucherlast einer Tarifgruppe bzw. aller Haushalte ist folgendermaßen definiert:

$$\begin{aligned}\bar{P}^{T_i}[k] &= \bar{P}_U^{T_i}[k] + \bar{P}_R^{T_i}[k] && \text{für alle } N_{T_i} \text{ Haushalte einer Tarifgruppe } T_i \\ \bar{P}[k] &= \bar{P}_U[k] + \bar{P}_R[k] && \text{für alle } N \text{ Haushalte}\end{aligned}\tag{2.8}$$

C. Prognosemodell des Gesamtverbrauchs

Für die Planung des Preissignals $p_n[k]$ durch das EVU ist die Kenntnis des ungestörten Netzzustandes und somit die Kenntnis des unbeeinflussten Anteils der Verbraucherlast $P_U[k]$ im Versorgungsgebiet von entscheidender Bedeutung. Da allerdings nur die beeinflusste Verbraucherlast $P[k]$ gemessen werden kann, muss $P_U[k]$ durch das Prognosemodell aus den historischen Werten $P[k-1] \dots P[k-i]$, dem Preissignal $p[k]$ und der Schätzung der haushaltsindividuellen, soziodemografischen, meteorologischen, tageszeitabhängigen, saisonalen und singulären Faktoren $\hat{u}[k]$ gemäß

$$\hat{P}_U[k] = f(P[k-1] \dots P[k-i], p[k], \hat{u}[k])\tag{2.9}$$

prognostiziert werden. Erste Ansätze für das Prognosemodell sind in [56, 210, 215] vorgestellt. Da das Ergebnis des Prognosemodells als Eingangsgröße für die Planung des neuen Preissignals dient, wird von einem rückgekoppelten System gesprochen.

Die Planung des Preissignals (Block A in Abb. 2.2) und das Prognosemodell des Gesamtverbrauchs (Block C in Abb. 2.2) stellen wichtige Bausteine bei der systemtechnischen Modellierung des preisbeeinflussten Verbrauchsverhaltens dar, sollen aber nicht Kern der vorliegenden Arbeit sein. Stattdessen soll in dieser Arbeit ein neues Konzept zur datengetriebenen Analyse und Modellierung des DR-Prozesses in einem Smart Grid vorgestellt werden (Block B in Abb. 2.2). Um das neue Konzept hinsichtlich seiner Anwendbarkeit und Einsetzbarkeit bewerten zu können, erfolgt in Abschnitt 2.2 eine umfassende Anforderungsanalyse für *Smart-Meter*-Datensätze.

2.2 Anforderungsanalyse

Das in der vorliegenden Arbeit vorgestellte Konzept basiert auf der Systemanalyse, Wissensextraktion und Modellbildung anhand von Messdaten, die beispielsweise in Feldstudien erhoben wurden. Die Beschaffenheit und Qualität der Messdaten trägt maßgeblich dazu bei, welche wissenschaftlichen Aufgabenstellungen mit dem vorgestellten Konzept untersucht werden können. Folgende Aufgabenstellungen (A1 bis A4) ergeben sich für die datengetriebene Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens (vgl. [76]):

- **Aufgabenstellung A1: Identifikation von saisonalen und wochentagsbedingten Unterschieden für eine *Demand-Response*-Maßnahme**
Es sollen saisonale und wochentagsbedingte Schwankungen für eine *Demand-Response*-Maßnahme, z.B. für Fälle mit ungewöhnlichen Preissignalen, identifiziert werden. Das neue Konzept soll beispielsweise Unterschiede in der Verbrauchsreaktion auf ein Preissignal zwischen Winter- und Sommermonaten aufdecken.
- **Aufgabenstellung A2: Identifikation von tarifbedingten Unterschieden für eine *Demand-Response*-Maßnahme**
Ein Echtzeitstrompreis (vgl. Abb. 1.5c) bewirkt eine andere Verbrauchsreaktion als ein Zweistufentarif (vgl. Abb. 1.5a), da die Preisstruktur des Zweistufentarifs i.d.R. im Voraus bekannt ist und der Haushaltsstromkunde im Sinne einer Tagesoptimierung auf das Preissignal reagiert. Das neue Konzept soll den unterschiedlichen Stromtarifen Rechnung tragen und Unterschiede in der Verbrauchsreaktion, hervorgerufen durch die verschiedenen Stromtarifstrukturen, erklären.
- **Aufgabenstellung A3: Identifikation von (Non-) Responder-Haushalten**
Erfolgt die Verbrauchsreaktion auf ein Preissignal durch manuelles Ein- und Ausschalten bestimmter Haushaltsgeräte, spielt die Anwesenheit des Menschen und die individuelle Bereitschaft zur Laständerung eine entscheidende Rolle. Mithilfe des neuen Konzeptes sollen haushaltsindividuelle Unterschiede in der Verbrauchsreaktion analysiert und sogenannte „Responder-Haushalte“ (Abk.: R-Haushalt; Haushalt reagiert stark auf Preissignal), „Non-Responder-Haushalte“ (Abk.: NR-Haushalt; Haushalt reagiert nicht bzw. schwach auf Preissignal) und „Semi-Responder-Haushalte“ (Abk.: SR-Haushalt; Haushalt reagiert mäßig auf Preissignal) identifiziert werden.
- **Aufgabenstellung A4: Prognose des preisbeeinflussten Verbrauchsverhaltens**
Die Ergebnisse aus A1 bis A3 haben analytischen Charakter und eignen sich daher nur begrenzt für den Einsatz als Prognosemodell zur Vorhersage des preisbeeinflussten Verbrauchsverhaltens von Haushaltsstromkunden. Aus diesem Grund soll, in Ergänzung zu den Aufgabenstellungen A1 bis A3, das preisbeeinflusste Verbrauchsverhaltens mit einem systemtechnischen *Demand-Response*-Modell beschrieben werden. Das zu entwickelnde Prognosemodell (vgl. Gleichung (2.4)) soll die haushaltsindividuelle Verbrauchsreaktion auf ein bestimmtes Preissignal beschreiben. Das Prognosemodell soll in der Lage sein, die analytischen Ergebnisse der Aufgabenstellung A1 bis A3 bei der Modellentwicklung zu berücksichtigen. Des Weiteren soll das Prognosemodell in komplexen Simulationsszenarien als Simulationsmodell eingesetzt werden.

Für jede der vier Aufgabenstellungen gilt es nun, Eigenschaften eines *Smart-Meter*-Datensatzes zu definieren und deren Auswirkung auf die Erfüllung der Aufgabenstellung zu diskutieren. Der Übersichtlichkeit halber werden die relevanten Eigenschaften⁶ in einer generalisierten Checkliste für *Smart-Meter*-Datensätze (siehe Tabelle 2.1 und 2.2) zusammengefasst. Die Bewertung und Gewichtung der Eigenschaften ist von qualitativer Natur und soll beispielsweise verdeutlichen, welche Eigenschaften für die Erfüllung der Aufgabenstellungen unbedingt notwendig sind

⁶Die Liste der Eigenschaften des *Smart-Meter*-Datensatzes erhebt keinen Anspruch auf Vollständigkeit und kann fortwährend ergänzt werden.

und damit einen signifikanten Einfluss (Kennzeichnung (+) in Tabelle 2.1 und 2.2) haben. Eigenschaften, die eher einen geringen oder keinen signifikanten Einfluss auf die Erfüllung der Aufgabenstellungen haben, werden mit (-) gekennzeichnet. In den beiden Tabellen ist zudem eine kurze Begründung für die Bewertungen mit angegeben.

Es muss an dieser Stelle deutlich gemacht werden, dass es hier zur Überlagerung zweier Kategorien kommt: die Qualität der Verbrauchsreaktion auf der einen Seite und die Qualität der Messung auf der anderen Seite. Eine stärkere Verbrauchsreaktion auf ein Preissignal, die durch die Existenz elektrischer Geräte zur Strom- bzw. Wärmebereitstellung oder durch automatisches *Demand Response* hervorgerufen wird, wird nur dann als solche in den *Smart-Meter*-Messdaten zu sehen sein, wenn die Messung in ausreichender Qualität hinsichtlich der Granularität der Messdaten und des Messzeitraumes erfolgt.

In den Tabellen 2.1 und 2.2 ist zu berücksichtigen, dass sich viele Eigenschaften gegenseitig beeinflussen, wie z.B. die Eigenschaften „E1: Messung der Last einzelner Haushaltsgeräte“, „E2: Messung der Haushaltslast“ und „E5: Granularität der Messdaten“. So können die Lastkurven einzelner Haushaltsgeräte (E1) durch Aggregation zu der Gesamtlast eines Haushaltes (E2) zusammengefasst werden. Außerdem existieren in der Literatur zahlreiche Ansätze, die bei alleiniger Kenntnis der Gesamthaushaltslast das Nutzerverhalten für einzelne Haushaltsgeräte extrahieren [216–223]. Diese Verfahren werden unter dem Begriff *non-intrusive appliance load monitoring* (NIALM) zusammengefasst, wobei die Granularität der Messdaten (E3) einen entscheidenden Einfluss auf die Qualität des NIALM hat [224].

Tabelle 2.1: Checkliste für *Smart-Meter*-Datensätze für die datengetriebene Analyse des preisbeeinflussten Verbrauchsverhaltens (Aufgabenstellungen A1 und A2). Weitere Eigenschaften, die das Verbrauchsverhalten von Haushaltsstromkunden beeinflussen, sind in Tabelle B.3 aufgeführt; Kennzeichnung: (+) = signifikanter Einfluss, (-) = geringer Einfluss. Abkürzung: HVAC - *Heating, Ventilation and Air Conditioning*.

Nr.: Eigenschaft des <i>Smart-Meter</i> -Datensatzes	A1: Identifikation von saisonalen, wochentags- und situationsbedingten Unterschieden für eine DR-Maßnahme	A2: Identifikation von tarifbedingten Unterschieden für eine DR-Maßnahme
E1: Messung der Last einzelner Haushaltsgeräte	(+) Der Einfluss der Jahreszeit (Außentemperatur) auf elektr. Geräte zur Wärme-/ Strombereitstellung ist besser ermittelbar, wenn die Last dieser Haushaltsgeräte aufgezeichnet wird.	(-) Unterschiede im Verbrauchsverhalten, die rein auf das Preissignal zurückzuführen sind, lassen sich auch ohne die Messung der Last einzelner Haushaltsgeräte identifizieren.
E2: Messung der Haushaltslast	(+) Die Haushaltslast ist, zusammen mit dem Strompreis, die entscheidende Messgröße zur DR-Bewertung. Die Haushaltslast kann aber auch aus den Lasten aller Haushaltsgeräte durch Aggregation berechnet werden.	
E3: Messung der Einstellung der Anlagentechnik	(+) Sind Thermostateinstellungen bekannt und als Messgröße im SMD enthalten, kann der Einfluss der Außentemperatur besser berechnet und modelliert werden.	(-) Unterschiede im Verbrauchsverhalten, die rein auf das Preissignal zurückzuführen sind, lassen sich auch ohne die Kenntnis der Einstellung der Anlagentechnik (z.B. Thermostateinstellung) identifizieren.
E4: Messzeitraum	(+) Jahreszeitliche Einflüsse auf eine DR-Maßnahme sind nur bewertbar, wenn der Messzeitraum Sommer-, Winter- und Übergangsmonate beinhaltet.	(-) Unterschiede im Verbrauchsverhalten, die rein auf das Preissignal zurückzuführen sind, lassen sich auch in kurzen Messzeiträumen bewerten
E5: Granularität der Messdaten	Der Einfluss der Granularität der Messdaten auf die Erfüllung der Aufgabenstellungen lässt sich nur schwer abschätzen. Untersuchungen über das Verbrauchsverhalten unter Berücksichtigung verschiedener Granularitäten liefert [224].	
E6: Existenz von ADR	(+) ADR erhöht das Lastverlagerungspotenzial und reduziert Unsicherheiten in der Verbrauchsreaktion, da diese unabhängig vom menschlichen Verhalten ist. Insofern hat die Existenz von ADR einen großen Einfluss auf die Erfüllung der Aufgabenstellungen A1 und A2.	

Fortsetzung der Tabelle

Nr.: Eigenschaft des Smart-Meter-Datensatzes	A1: Identifikation von saisonalen, wochentags- und situationsbedingten Unterschieden für eine DR-Maßnahme	A2: Identifikation von tarifbedingten Unterschieden für eine DR-Maßnahme
E7: Existenz elektr. Geräte zur Strom-/Wärmebereitstellung (HVAC)	(+) Der Einsatz von Klimageräten, Luftheizungsanlagen und Geräten zur elektr. Warmwasserbereitung ist stark abhängig von der Außentemperatur. Somit hat die Existenz und die Messung der Leistungsaufnahme dieser Geräte einen starken Einfluss auf ein saisonal geprägtes Verbrauchsverhalten.	(+) Die Existenz von Klimageräten, Luftheizungsanlagen und Geräten zur elektr. Warmwasserbereitung erhöht das Lastverlagerungspotenzial. Somit steigt die Lastverschiebung tendenziell und tarifbedingte Unterschiede sind besser erkennbar.
E8: Existenz von Feedback-Technologien	(+) Feedback-Technologien (z.B. <i>in-home displays</i>) erhöhen das Potenzial zur Energieeinsparung und zur Lastverlagerung und begünstigen somit die Erfüllung der Aufgabenstellungen A1 und A2.	
E9: Existenz von Wetterdaten	(+) Der Einfluss der Außentemperatur auf elektr. Geräte zur Wärme-/ Strombereitstellung ist besser ermittelbar, wenn neben der Last auch Wetterdaten im SMD enthalten sind.	(-) Unterschiede im Verbrauchsverhalten, die rein auf das Preissignal zurückzuführen sind, lassen sich ohne die Existenz von Wetterdaten im SMD bewerten, sofern es eine Kontrollgruppe mit vergleichbaren Wetterbedingungen gibt.
E10: Existenz einer Kontrollgruppe	(-) Auch ohne eine Kontrollgruppe können Unterschiede identifiziert werden, die auf saisonale, wochentags- und situationsbedingte Einflüsse zurückzuführen sind.	(+) Um die Verbrauchsreaktion auf unterschiedliche Strompreissignale bewerten zu können, ist das Vorhandensein einer Kontrollgruppe zwingend notwendig.
E11: Zusammenstellung der Kontrollgruppe	(-) Auch ohne eine Kontrollgruppe können Unterschiede identifiziert werden, die auf saisonale, wochentags- und situationsbedingte Einflüsse zurückzuführen sind.	(+) im Idealfall: die gleichen Haushalte befinden sich zeitlich gestaffelt in der Tarifgruppe und in der Kontrollgruppe; alternativ: zufällige Auswahl einer großen Anzahl an Haushalten; kritisch: Selection Bias, bei dem sich die Haushalte den für Sie passenden Tarif aussuchen könnten
E12: Anzahl der Tarifgruppen	(+) Es ist mindestens eine Tarifgruppe notwendig.	(+) Es ist mindestens eine Tarifgruppe plus eine Kontrollgruppe notwendig.
E13: Anzahl der gemessenen Haushalte	(+) Soll das aggregierte Verbrauchsverhalten bzw. Unterschiede im aggregierten Verbrauchsverhalten identifiziert werden, spielt die Anzahl der Haushalte eine wichtige Rolle. Je mehr Haushalte bei der Aggregation zur Verfügung stehen, desto geringer ist der Einfluss der stochastischen Lastkomponente eines Einzelhaushaltes. (-) Wenn saisonale, wochentags- und situationsbedingte Unterschiede für eine DR-Maßnahme auf Einzelhaushaltsebene bewertet werden sollen, spielt die Anzahl der gemessenen Haushalte keine Rolle.	

Tabelle 2.2: Checkliste für *Smart-Meter*-Datensätze für die datengetriebene Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens (Aufgabenstellungen A3 und A4). Weitere Eigenschaften, die das Verbrauchsverhalten von Haushaltsstromkunden beeinflussen, sind in Tabelle B.3 aufgeführt; Kennzeichnung: (+) = signifikanter Einfluss, (-) = geringer Einfluss. Abkürzung: HVAC - *Heating, Ventilation and Air Conditioning*.

Nr.: Eigenschaft des Smart-Meter-Datensatzes	A3: Identifikation von (Non-) Responder-Haushalten	A4: Prognose des preisbeeinflussten Verbrauchsverhaltens
E1: Messung der Last einzelner Haushaltsgeräte	(-) Unterschiede im Verbrauchsverhalten, die rein auf das Preissignal zurückzuführen sind, lassen sich auch ohne die Messung der Last einzelner Haushaltsgeräte identifizieren.	Der Einfluss von E1 auf die Erfüllung von A4 ist abhängig von der Klasse des Prognosemodells: (+) bei physikalisch motivierten Modellen; (-) bei restlichen Modellklassen (vgl. Tabelle 1.4)
E2: Messung der Haushaltslast	(+) Die Haushaltslast ist, zusammen mit dem Strompreis, die entscheidende Messgröße zur DR-Bewertung. Die Haushaltslast kann aber auch aus den Lasten aller Haushaltsgeräte durch Aggregation berechnet werden.	
E3: Messung der Einstellung der Anlagentechnik	(-) Haushaltsindividuelle Verbrauchsunterschiede lassen sich auch ohne die Kenntnis der Thermostateinstellung identifizieren.	Der Einfluss von E3 auf die Erfüllung von A4 ist abhängig von der Klasse des Prognosemodells: (+) bei physikalisch motivierten Modellen; (-) bei restlichen Modellklassen (vgl. Tabelle 1.4)

Fortsetzung der Tabelle

Nr.: Eigenschaft des Smart-Meter-Datensatzes	A3: Identifikation von (Non-) Responder-Haushalten	A4: Prognose des preisbeeinflussten Verbrauchsverhaltens
E4: Messzeitraum	(-) Haushaltsindividuelle Verbrauchsunterschiede lassen sich auch in kurzen Messzeiträumen bewerten. Dies gilt nicht, wenn ein Haushalt z.B. im Winter Responder und im Sommer Non-Responder ist.	(+) Jahreszeitliche Einflüsse auf eine DRMaßnahme sind nur bewertbar, wenn der Messzeitraum Sommer-, Winter- und Übergangsmo-nate beinhaltet.
E5: Granularität der Messdaten	Der Einfluss der Granularität der Messdaten auf die Erfüllung der Aufgabenstellung lässt sich nur schwer abschätzen. Untersuchungen über das Verbrauchsverhalten unter Berücksichtigung verschiedener Granularitäten liefert [224].	
E6: Existenz von ADR	(+) ADR erhöht das Lastverlagerungspotenzial und reduziert Unsicherheiten in der Verbrauchsreaktion, da diese unabhängig vom menschlichen Verhalten ist. Insofern hat die Existenz von ADR einen großen Einfluss auf die Erfüllung der Aufgabenstellung A3.	Der Einfluss von E6 auf die Erfüllung von A4 ist abhängig von der Klasse des Prognosemodells: (+) bei physikalisch motivierten Modellen; (-) bei restlichen Modellklassen (vgl. Tabelle 1.4)
E7: Existenz elektr. Geräte zur Strom-/Wärmebereitstellung (HVAC)	(+) Der Einsatz von Klimageräten, Luftheizungsanlagen und Geräten zur elektr. Warmwasserbereitung erhöht den Gesamtstromverbrauch und führt zu ausgeprägteren Lastprofilen. Die Existenz dieser Geräte hat somit einen großen Einfluss auf die Erfüllung von A3 und A4.	
E8: Existenz von Feedback-Technologien	(+) Feedback-Technologien (z.B. <i>in-home displays</i>) erhöhen das Potenzial zur Energieeinsparung und zur Lastverlagerung und begünstigen somit die Erfüllung der Aufgabenstellungen A3 und A4.	
E9: Existenz von Wetterdaten	(-) Die Identifikation von (Non-) Responder-Haushalten ist auch ohne die Existenz von Wetterdaten im SMD möglich.	(+) Insbesondere bei physikalisch motivierten Modellen können Wetterdaten als weitere Eingangsgrößen berücksichtigt werden.
E10: Existenz einer Kontrollgruppe	(-) Eine Kontrollgruppe hilft bei der Identifizierung von (Non-) Responder-Haushalten. Unterschiedlich starke Verbrauchsreaktionen können allerdings auch ohne eine Kontrollgruppe ausfindig gemacht werden.	(+) Ohne eine Kontrollgruppe kann keine Modellvalidierung vorgenommen werden. Daher ist die Existenz einer Kontrollgruppe für die Erfüllung von A4 zwingend notwendig.
E11: Zusammenstellung der Kontrollgruppe	(-) Eine Kontrollgruppe hilft bei der Identifizierung von (Non-) Responder-Haushalten. Unterschiedlich starke Verbrauchsreaktionen können allerdings auch ohne eine Kontrollgruppe ausfindig gemacht werden.	(+) im Idealfall: die gleichen Haushalte befinden sich zeitlich gestaffelt in der Tarifgruppe und in der Kontrollgruppe; alternativ: zufällige Auswahl einer großen Anzahl an Haushalten; kritisch: Selection Bias, bei dem sich die Haushalte den für Sie passenden Tarif aussuchen konnten
E12: Anzahl der Tarifgruppen	(+) Es ist mindestens eine Tarifgruppe plus eine Kontrollgruppe notwendig.	(+) Es ist mindestens eine Tarifgruppe plus eine Kontrollgruppe notwendig.
E13: Anzahl der gemessenen Haushalte	(+) Sollen Unterschiede im aggregierten Verbrauchsverhalten identifiziert werden, spielt die Anzahl der Haushalte eine große Rolle. Je mehr Haushalte bei der Aggregation zur Verfügung stehen, desto geringer ist der Einfluss der stochastischen Lastkomponente eines Einzelhaushaltes. (-) Wenn bewertet werden soll, ob ein Haushalt stark oder wenig auf ein Preissignal reagiert, spielt die Anzahl der gemessenen Haushalte keine Rolle.	(+) Die Aggregation von Haushalten führt zur Reduktion von Unsicherheiten, was die Modellierung begünstigt. Die Anzahl der gemessenen Haushalte hat somit einen großen Einfluss auf die Erfüllung von A4.

2.3 Datenstruktur und Benchmarkdatensätze

Im vorliegenden Abschnitt wird die grundsätzliche Datenstruktur der untersuchten *Smart-Meter*-Datensätze erläutert. Für die Bewertung der Wirksamkeit des Gesamtkonzeptes und für die Validierung der einzelnen Verfahren aus Abb. 2.1 gilt es, Benchmarkdatensätze zu erzeugen. Dazu stellt Abschnitt 2.3.2 das Verfahren zur Generierung von Benchmarkdatensätzen vor. Insgesamt wurden zwölf Benchmarkdatensätze erzeugt, deren Charakteristik und Parameterwerte in Abschnitt 2.3.3 behandelt werden.

2.3.1 Datenstruktur

Die Rohdaten eines *Smart-Meter*-Datensatzes bestehen aus $n_D = 1, \dots, N_D$ Datentupeln (englisch: *data point*, *data sample*). Für jedes Datentupel existieren idealerweise

$$\begin{aligned} j = 1, \dots, s_z & \text{ Zeitreihen} & x_{j,n_D}[k], \\ l = 1, \dots, s & \text{ Einzelmerkmale} & x_{l,n_D} \text{ und} \\ i = 1, \dots, s_y & \text{ Ausgangsgrößen} & y_{i,n_D}, \end{aligned}$$

wobei K die Länge der Zeitreihen mit $k = 1, \dots, K$ ist. Abb. 2.3 veranschaulicht die Datenstruktur in einem *Smart-Meter*-Datensatz.

Grundsätzlich beschreibt ein Merkmal die Eigenschaft eines Datenobjektes. Im Fall des preisbeeinflussten Verbrauchsprozesses mit Systemrückkopplung (vgl. Abschnitt 2.1) ergeben sich zahlreiche Merkmale, wie beispielsweise der Strompreis, die Haushaltslast, die verbrauchte tägliche Strommenge eines Haushaltes, die mittlere Tagestemperatur der Umgebung oder die Zugehörigkeit eines Haushaltes zu einer Tarifgruppe. Merkmale lassen sich in qualitative und quantitative Merkmale einteilen. Qualitative Merkmale sind von beschreibender Natur und deren Ausprägungen sind durch linguistische Ausdrücke charakterisiert. Um qualitative Merkmale in die oben gezeigte Datenstruktur einzuordnen, werden die Merkmalsausprägungen quantifiziert und als Klassen $b = 1, \dots, m_y$ der $i = 1, \dots, s_y$ Ausgangsgrößen y_{i,n_D} beschrieben. $N_{D,b}$ ist dabei die Anzahl der Datentupel, die zu einer bestimmten Klasse b gehören, wobei

$$N_D = \sum_{b=1}^{m_y} N_{D,b} \quad (2.10)$$

gilt. Als Beispiel für ein qualitatives Merkmal in einem *Smart-Meter*-Datensatz ist die Ausprägung des Wochentages zu nennen. Die Ausgangsgröße „Wochentag“ ließe sich hierbei in $m_y = 7$ Klassen („Montag“, „...“, „Sonntag“) einteilen.

Quantitative Merkmale sind Merkmale, bei denen die Merkmalsausprägungen zahlenwertig sind. Sie werden zudem häufig in wertediskrete und wertekontinuierliche Merkmale eingeteilt. Die Merkmalsausprägungen eines wertediskreten Merkmals bilden eine diskrete Zahlenmenge, was am Beispiel eines zweistufigen TOU-Tarifs mit den Tarifstufen „Niedertarif“ (z.B. $p = 4 \text{ ¢/kWh}$)

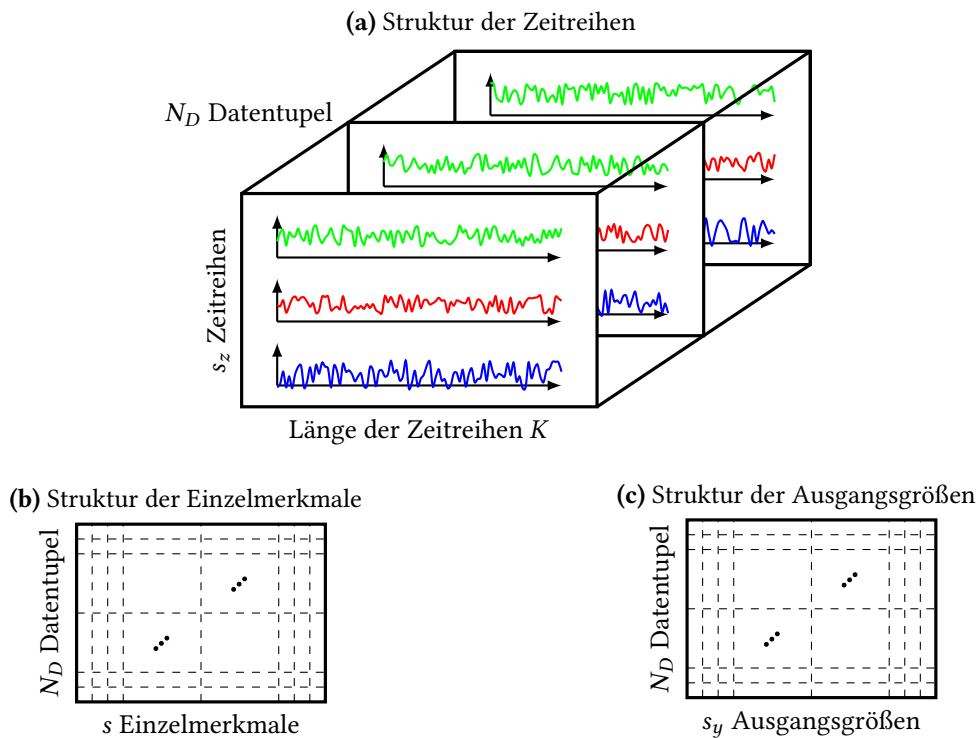


Abbildung 2.3: Schematische Darstellung der Datenstruktur in einem *Smart-Meter*-Datensatz (modifiziert nach [85]).

und „Hochtarif“ (z.B. $p = 12 \text{ ¢/kWh}$) deutlich wird. Bei wertekontinuierlichen Merkmalen können die Merkmalsausprägungen jeden denkbaren Wert annehmen. Die reellwertige Haushaltslast $P_n[k]$ in kW ist ein Beispiel für ein wertekontinuierliches Merkmal. In der Regel werden quantitative Merkmale im *Smart-Meter*-Datensatz als Zeitreihen oder Einzelmerkmale abgebildet.

2.3.2 Generierung von Benchmarkdatensätzen - Benchmarkgenerator

Für das neue Konzept zur datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens werden Benchmarkdatensätze definiert. Die synthetischen Datensätze dienen zum einen der Bewertung der Wirksamkeit des Gesamtkonzeptes, zum anderen der Verdeutlichung und Validierung der einzelnen Methoden aus Abb. 2.1. In Abb. 2.5 ist das prinzipielle Vorgehen beim Erstellen eines Benchmarkdatensatzes in einem Flussdiagramm dargestellt. Bei der Generierung eines Benchmarkdatensatzes müssen drei wesentliche Elemente genauer betrachtet werden: der **Lastsignalgenerator**, der **Demand-Response-Simulator** und der **Messfehlergenerator**, die im Folgenden erläutert werden.

2.3.2.1 Lastsignalgenerator

Nach Vorgabe der Anzahl an betrachteten Haushalten N wird durch den Lastsignalgenerator ein haushaltsindividuelles Lastprofil ($P_{U,n}[k]$) mit deterministischen und stochastischen Signalanteilen erzeugt. Der Aufbau des synthetischen Lastsignals orientiert sich am additiven Komponentenmodell, das in Gleichung (1.1) beschrieben ist. Die Signalkomponenten lassen sich einzeln konfigurieren und gewichten, sodass das Gesamtsignal als Summe der Einzelsignalkomponenten 100% ergibt. In Beispiel 1 sind die Signalkomponenten für ein Beispiellastsignal dargestellt.

Neben der Konfiguration und Gewichtung der Einzelsignalkomponenten können die Signallänge K , die Abtastperiodendauer T_S und der Gesamtenergieinhalt⁷ E_n der Lastzeitreihe $P_{U,n}[k]$ mit

$$E_n = \sum_{k=1}^K P_{U,n}[k] \quad (2.11)$$

vorgegeben werden. Besteht der Benchmarkdatensatz aus einer Vielzahl an Haushalten, schwanken die Gesamtstromverbräuche der einzelnen Haushalte um einen bestimmten Mittelwert μ_E . Diese Schwankungen werden durch die normalverteilte Zufallsvariable

$$X_E^n \sim \mathcal{N}(\mu_E, \sigma_E^2) \quad (2.12)$$

beschrieben, wobei σ_E^2 die mittlere quadratische Abweichung der einzelnen Gesamtstromverbräuche um den Erwartungswert $\mu_E = E_n$ angibt.

Der Lastsignalgenerator wurde nicht im Rahmen der vorliegenden Arbeit, sondern am Fraunhofer-IOSB AST⁸ von Stefan Klaiber entwickelt [226]. Aufgrund des verständlichen Aufbaus und der vielfältigen Konfigurationsmöglichkeiten wurde dieser Lastsignalgenerator zur Erzeugung synthetischer, haushaltsindividueller Lastprofile ausgewählt.

Ein ähnlicher (verhaltensbasierter) Lastprofilgenerator zur Generierung von synthetischen Haushaltlastprofilen wurde in [227–230] vorgestellt. Ein alternativer Ansatz zur Generierung der synthetischen, unbeeinflussten Verbraucherlast $P_{U,n}[k]$ stellt die Verwendung bereits existierender Datensätze dar. So beinhaltet der in [231] vorgestellte Datensatz hoch aufgelöste Lastprofile ($T_S = 1$ s), die typische Lastfluktuationen aufweisen.

⁷Im Jahr 2015 betrug der durchschnittliche Haushaltsstromverbrauch 2842 kWh (ohne Elektro-Speicherheizung) bzw. 3105 kWh (mit Elektro-Speicherheizung). Quelle: BMWi; Statistisches Bundesamt; BDEW [225]

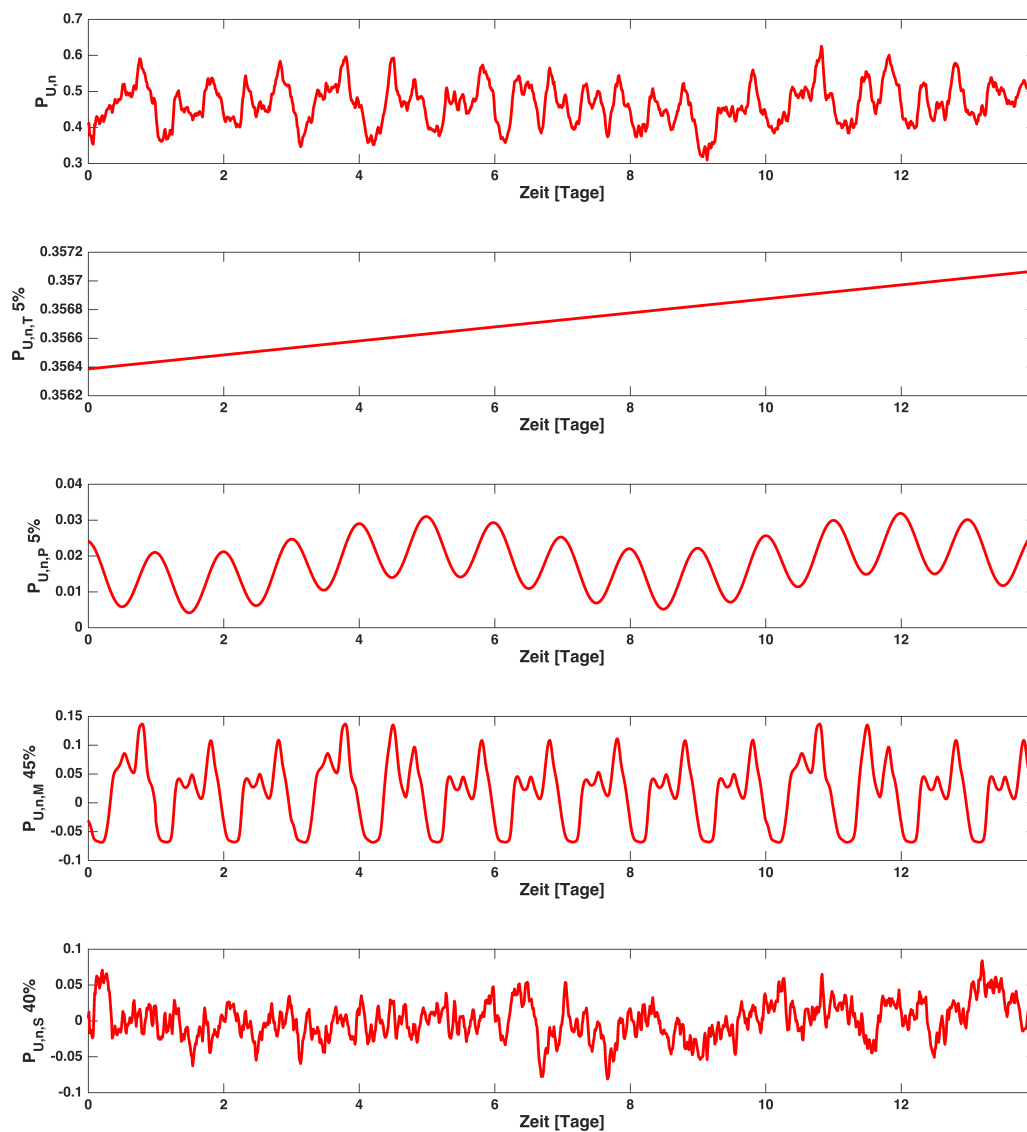
⁸Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung (IOSB), Institutsteil Angewandte Systemtechnik (AST)

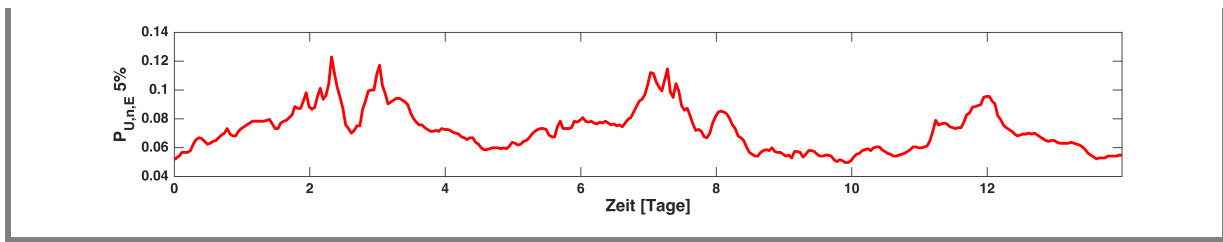
Beispiel 1: Signalkomponenten eines Beispiellastsignals

Die Zusammensetzung einer synthetischen, unbeeinflussten Verbraucherlast $P_{U,n}$ aus den unterschiedlichen Signalkomponenten ist in diesem Beispiel gezeigt. Die synthetische, unbeeinflusste Verbraucherlast $P_{U,n}$ berechnet sich wie folgt:

$$P_{U,n}[k] = 0.05 \cdot P_{U,T}[k] + 0.05 \cdot P_{U,P}[k] + 0.45 \cdot P_{U,M}[k] + 0.40 \cdot P_{U,S}[k] + 0.05 \cdot P_{U,E}[k].$$

Die Anteile der Signalkomponenten sind zudem an den Ordinatenachsen angegeben. Aus den unten stehenden Abbildungen wird ersichtlich, dass das tageszeitliche Verbrauchsverhalten mit einer Lastspitze am Abend und einer oder zwei Lastspitzen am Morgen im Lastprofil der Musterkomponente abgebildet ist. Das unterschiedliche Verbrauchsverhalten an Werk- und Wochenendtagen ist sowohl in der periodischen Komponente als auch in der Musterkomponente enthalten.





2.3.2.2 Demand-Response-Simulator

Der *Demand-Response-Simulator* simuliert die haushaltsindividuelle Verbrauchsreaktion auf ein definiertes Strompreisprofil. Grundsätzlich können im DR-Simulator verschiedene Modelle für die preisbeeinflusste Verbrauchsreaktion eines Haushaltes (vgl. Abschnitt 1.2.8) und verschiedene synthetische Strompreisprofile (Abb. 1.5a, 1.5b und 1.5c) verwendet werden. Die synthetischen Strompreisprofile, die im *Demand-Response-Simulator* ausgewählt werden können, sind in Abb. 2.4 dargestellt. Die statischen Strompreistarife aus Abb. 2.4(a) und Abb. 2.4(b)

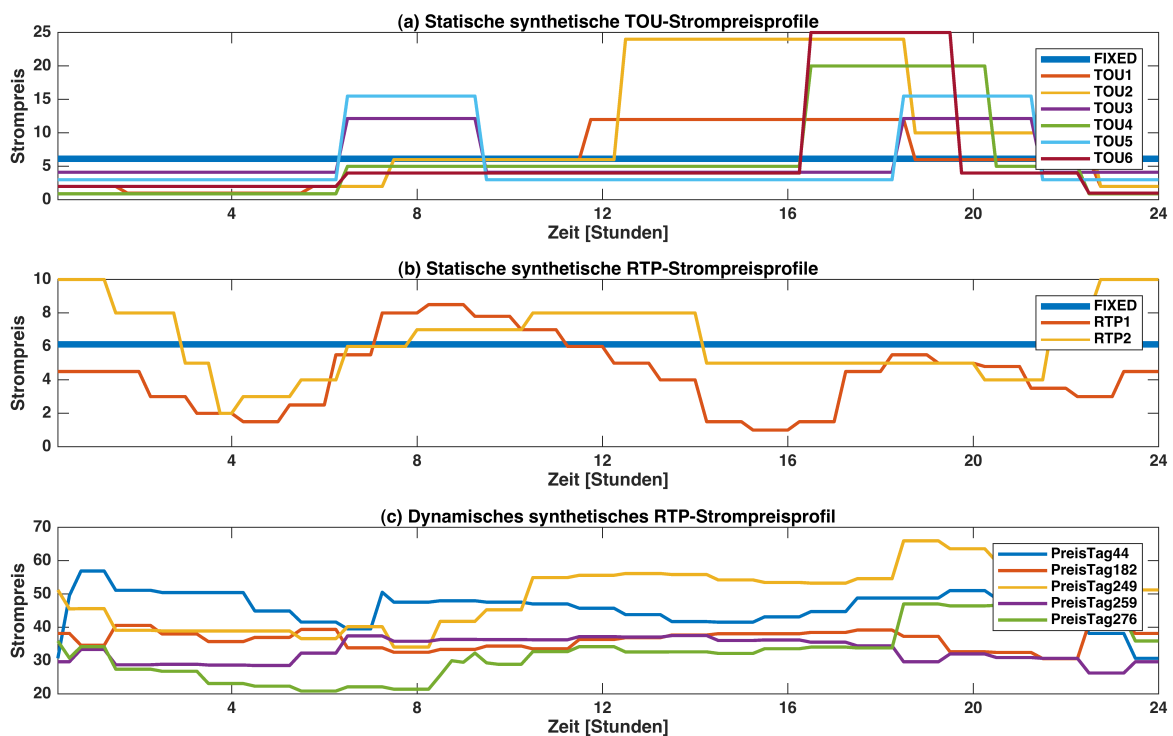


Abbildung 2.4: Synthetische Strompreisprofile des *Demand-Response-Simulators* zur Erzeugung von Benchmarkdatensätzen: (a) Tagesverläufe der statischen TOU-Strompreistarife, (b) Tagesverläufe der statischen RTP-Strompreistarife und (c) Verlauf des dynamischen RTP-Strompreistarifs an fünf ausgewählten Tagen.

zeichnen sich dadurch aus, dass die dargestellten Verläufe an jedem Tag des Jahres gleich sind. Der dynamische RTP-Tarif in Abb. 2.4(c) hat dagegen für jeden Tag des Jahres einen unterschiedlichen Verlauf. Die Strompreisstufen in Abb. 2.4(a) wurden so gewählt, dass der mittlere Strompreis ($\bar{p} = 6,125 \text{ ¢/kWh}$) in fast allen Fällen gleich ist. Einzig der „TOU2“-Tarif weist einen erhöhten mittleren Strompreis von $\bar{p} = 9,9 \text{ ¢/kWh}$ auf.

Wie bereits in Abb. 1.6 gezeigt, führt eine größere Spreizung zwischen *on*- und *off-peak* Strompreis dazu, dass Haushaltsstromkunden eher bereit sind, die Last zu Spitzenlastzeiten zu reduzieren bzw. zu verschieben. Des Weiteren zeigten die Untersuchungen des *MeRegio*-Projektes, dass auch die Tarifstufendauer eine wichtige Rolle bei der Lastverschiebung in Haushalten spielt [17, 151]. Sollen nun Benchmarkdatensätze mit unterschiedlichen TOU-Stromtarifen generiert werden, können sowohl Tarife mit einer geringen Preisspreizung und einer kurzen Tarifstufendauer (z.B. „TOU3“-Tarif), als auch Tarife mit einer größeren Preisspreizung und langen Tarifstufendauer (z.B. „TOU2“-Tarif) ausgewählt werden.

In der vorliegenden Arbeit wurde ein ökonomisches *Demand-Response*-Modell verwendet, welches sich an Gleichung (3) aus [151] orientiert. Dabei ist die preisbeeinflusste Haushaltslast $P_n[k]$ des n -ten Haushaltes durch

$$P_n[k] = P_{U,n}[k] + \underbrace{P_{U,n}[k] \cdot \varepsilon[k, k] \cdot \Delta p[k] \cdot X_\varepsilon^n}_{P_{R,n}[k]} \quad (2.13)$$

mit

$$\Delta p[k] = \frac{p[k] - \bar{p}}{\bar{p}} \quad (2.14)$$

und

$$X_\varepsilon^n = \begin{cases} \mathcal{U}(0, 0.1) & , \text{ für Non-Responder-Haushalte} \\ \mathcal{U}(0.1, 0.5) & , \text{ für Semi-Responder-Haushalte} \\ \mathcal{U}(0.5, 1) & , \text{ für Responder-Haushalte} \end{cases} \quad (2.15)$$

beschrieben. Die Werte für die Preiselastizität ε in Gleichung (2.13) hängen von der Jahreszeit, dem Wochentag, der Tageszeit und der Tarifstufendauer ab, und wurden anhand realer Messdaten aus dem *MeRegio*-Projekt gewonnen (vgl. [17, 151]) und sind in der *Lookup*-Tabelle B.4 dargestellt.

Die Zufallsvariable X_ε^n gibt an, wie stark eine Verbrauchsreaktion auf ein Preissignal ausfällt. Für Haushalte, die nicht oder nur sehr gering auf ein Preissignal reagieren (Non-Responder-Haushalte), liegt der Wert für X_ε^n im Bereich von 0 bis 0.1 und wird mit einer gleichverteilten Zufallsvariablen \mathcal{U} berechnet. Haushalte, die stark auf ein Preissignal reagieren, werden als Responder-Haushalte bezeichnet (vgl. Aufgabenstellung A3 auf Seite 32). Für die Responder-Haushalte gilt $X_\varepsilon^n = \mathcal{U}(0.5, 1)$. Neben den Responder- und Non-Responder-Haushalten können im DR-Simulator auch Semi-Responder-Haushalte simuliert werden. Die Reaktion dieser Haushalte auf ein Preissignal fällt weniger stark aus als bei Responder-Haushalten. Der Wert für X_ε^n liegt bei Semi-Responder-Haushalten im Bereich von 0.1 bis 0.5. Mithilfe von Gleichung (2.15) soll sichergestellt werden, dass verschiedene Haushalte unterschiedlich stark auf ein Preissignal reagieren.

2.3.2.3 Messfehlergenerator

Für die Datenanalyse und Prognose im Energieumfeld (z.B. Analyse der Haushaltslast [76], Prognose der Leistung einer Photovoltaik-Anlage [87, 97]) spielt die Qualität von Messzeitreihen

eine entscheidende Rolle. Fehlende oder undefinierte Werte und Ausreißer in Messzeitreihen gilt es geeignet zu erkennen und zu beseitigen, wobei hierfür zahlreiche Methoden und Handlungsempfehlungen existieren [85, 87, 95, 97], die in Abschnitt 2.4.1 näher vorgestellt werden.

Da reale *Smart-Meter*-Datensätze selten frei von Messfehlern (Ausreißer oder fehlende Werte) sind, werden bei der Generierung von Benchmarkdatensätzen derartige Fehler künstlich erzeugt. Im Messfehlergenerator können für jede Lastzeitreihe die drei Fehlerarten „Bereiche mit fehlenden Werten“, „einzelne fehlende Werte“ und „Ausreißer“ generiert werden. Dazu muss im Messfehlergenerator

- die Anzahl an Bereichen mit fehlenden Werten b_{nan}^n ,
- der prozentuale Anteil $\mu_{nan,g}$ an zusammenhängenden fehlenden Werten $K_{nan,g}$ bezogen auf die Zeitreihenlänge K mithilfe von

$$\mu_{nan,g} = \frac{K_{nan,g}}{K}, \quad (2.16)$$

- der prozentuale Anteil $\mu_{nan,s}$ an einzelnen fehlenden Werten $K_{nan,s}$ bezogen auf die Zeitreihenlänge K abzüglich der Bereiche mit fehlenden Werten $K_{nan,g}$ mithilfe von

$$\mu_{nan,s} = \frac{K_{nan,s}}{K - K_{nan,g}} \quad (2.17)$$

- und der prozentuale Anteil μ_a an Ausreißern K_a bezogen auf die Zeitreihenlänge K abzüglich der fehlenden Werte ($K_{nan,g} + K_{nan,s}$) mithilfe von

$$\mu_a = \frac{K_a}{K - (K_{nan,s} + K_{nan,g})} \quad (2.18)$$

festgelegt werden.

Nachdem nun die Anzahl an fehlenden Werten und Ausreißern definiert wurde, muss noch auf die Verteilung der Messfehler entlang der Zeitreihe eingegangen werden. Die Positionierung der Messfehler beinhaltet folgende Schritte:

- **Schritt 1:** Zunächst muss die Anzahl an Bereichen mit fehlenden Werten b_{nan}^n , der prozentuale Anteil $\mu_{nan,g}$ an Bereichen mit fehlenden Werten, der prozentuale Anteil $\mu_{nan,s}$ an einzelnen fehlenden Werten und der prozentuale Anteil μ_a an Ausreißern definiert werden.
- **Schritt 2:** Es wird die prozentuale Aufteilung aller $K_{nan,g}$ fehlenden Werte auf die b_{nan}^n Bereiche mithilfe einer gleichverteilten Zufallsvariablen bestimmt. Dadurch ist die Länge aller b_{nan}^n Bereiche mit fehlenden Werten festgelegt.
- **Schritt 3:** Es werden alle b_{nan}^n Startindizes für die Bereiche mit fehlenden Werten mithilfe der gleichverteilten ganzzahligen Zufallsvariablen $X_{nan,g,pos}^n \sim \mathcal{U}(1, K)$ berechnet.

- **Schritt 4:** Die Schritte 2 und 3 waren erfolgreich, wenn sich die Bereiche mit fehlenden Werten nicht überlappen und der letzte Bereich die Indexgrenze K nicht überschreitet. Falls dies nicht erfüllt ist, werden die Schritte 2 und 3 wiederholt. Die fehlenden Messwerte werden anschließend mit NaN (*Not a Number*) gekennzeichnet.
- **Schritt 5:** Die einzelnen fehlenden Werte werden gleichmäßig über die Länge der Zeitreihe verteilt, wobei die Bereiche mit fehlenden Werten zuvor ausgeschlossen werden müssen. Die einzelnen fehlenden Messwerte werden anschließend mit NaN (*Not a Number*) gekennzeichnet.
- **Schritt 6:** Die Ausreißer werden gleichmäßig über die Länge der Zeitreihe verteilt, wobei die Bereiche mit fehlenden Werten und die einzelnen fehlenden Werte zuvor ausgeschlossen werden müssen.

Damit nicht alle Lastzeitreihen eines Benchmarkdatensatzes eine exakt gleiche Anzahl an einzelnen fehlenden Werten aufweisen, kann mit der normalverteilten Zufallsvariablen $X_{nan,s}^n$ und der Varianz $\sigma_{nan,s}^2$ gemäß

$$X_{nan,s}^n \sim \mathcal{N}(\mu_{nan,s}, \sigma_{nan,s}^2) \quad (2.19)$$

eine Streuung der Anzahl an einzelnen fehlenden Werten um den Erwartungswert $\mu_{nan,s}$ gewährleistet werden. In analoger Weise kann mit der normalverteilten Zufallsvariablen $X_{nan,g}^n$ und der Varianz $\sigma_{nan,g}^2$ gemäß

$$X_{nan,g}^n \sim \mathcal{N}(\mu_{nan,g}, \sigma_{nan,g}^2) \quad (2.20)$$

eine Streuung des prozentualen Anteils an zusammenhängenden Bereichen mit fehlenden Werten um den Erwartungswert $\mu_{nan,g}$ sichergestellt werden.

Auch im Falle der Ausreißer ist eine Streuung des prozentualen Anteils an Ausreißern um den Erwartungswert μ_a durch die normalverteilten Zufallsvariablen X_a^n und der Varianz σ_a^2 gemäß

$$X_a^n \sim \mathcal{N}(\mu_a, \sigma_a^2) \quad (2.21)$$

vorgesehen. Neben der Anzahl (K_a) und Position der Ausreißer (k_a) ist der numerische Wert der Ausreißer zu definieren. Mithilfe der gleichverteilten Zufallsvariablen

$$X_{ha}^n \sim \mathcal{U}(a, b) \quad (2.22)$$

und der Intervallgrenzen a und b , die das kleinste bzw. größte mögliche Ergebnis der Zufallsvariablen darstellen, kann durch Multiplikation des ursprünglichen Wertes $P_n[k_a]$ mit der Zufallsvariablen X_{ha}^n der Wert des Ausreißers wie folgt berechnet werden:

$$P_n^*[k_a] = X_{ha}^n \cdot P_n[k_a]. \quad (2.23)$$

2.3.3 Benchmarkdatensätze

Nachdem im vorherigen Abschnitt das Verfahren zur Generierung von Benchmarkdatensätzen vorgestellt wurde, geht Abschnitt 2.3.3 auf die Benchmarkdatensätze ein, die zur Bewertung der Wirksamkeit des Gesamtkonzeptes und zur Validierung der einzelnen Methoden aus Abb. 2.1 erzeugt wurden. Der in Abschnitt 2.3.2 vorgestellte Benchmarkgenerator erlaubt es, Benchmarkdatensätze mit bestimmten Eigenschaften und Charakteristiken zu erzeugen. Die im Rahmen der Arbeit generierten Benchmarkdatensätze und die hierfür notwendigen Parametereinstellungen wurden in Abschnitt 2.3.3 präsentiert. Die Methodik der Benchmarkgenerierung und anschließender Analyse des Datensatzes ermöglicht die Untersuchung weiterer Aspekte des preisbeeinflussten Verbrauchsverhaltens.

2.3.3.1 Benchmarkdatensatz zur systematischen Untersuchung verschiedener Imputationsverfahren

Im Rahmen der Arbeit wurde ein Benchmarkdatensatz generiert, der die Basis für ein Testverfahren bildet, welches systematisch den Umgang mit fehlenden Werten (siehe Imputationsverfahren in Abschnitt 1.2.6) untersuchen soll. Die Parametereinstellungen des Benchmarkgenerators zur Erzeugung dieses Benchmarkdatensatzes sind in Tabelle B.5 dargestellt.

2.3.3.2 Benchmarkdatensatz zur Ermittlung von optimalen Hampel-Filterparametern

Ferner wurde ein zweiter Benchmarkdatensatz generiert, der die Datenbasis für ein Testverfahren bildet, welches systematisch die Wahl von optimalen Hampel-Filterparametern untersucht. Der Hampel-Filter wurde in einer Vielzahl von Veröffentlichungen als Verfahren zur Ausreißerbehandlung erfolgreich eingesetzt [76, 102, 232] und wird ausführlich in Abschnitt 2.4.1.2 beschrieben. Die Parametereinstellungen des Benchmarkgenerators zur Erzeugung des Benchmarkdatensatzes sind in Tabelle B.6 dargestellt. Die Ergebnisse der beiden Testverfahren auf Basis der hier präsentierten Benchmarkdatensätze werden in Abschnitt 2.4.1 vorgestellt.

2.3.3.3 Benchmarkdatensätze zur Validierung der analyseorientierten Methoden der Pipeline

In diesem Abschnitt werden die Benchmarkdatensätze vorgestellt, die der Validierung der analyseorientierten Methoden der Pipeline dienen. Tabelle B.7 fasst die Parametereinstellungen des Benchmarkgenerators zur Erzeugung dieser Benchmarkdatensätze zusammen. Die Bezeichnung der Datensätze gibt Rückschlüsse auf die Anzahl der Haushalte N , die Länge der Zeitreihen K , den Gewichtungsfaktor der Musterkomponente β_M und den Gewichtungsfaktor der stochastischen Komponente β_S . So enthält zum Beispiel der Datensatz „Settlement210H365D_60M30S“ Zeitreihen mit einer Länge von 365 Tagen (35040 Abtastzeitpunkte $\cdot 0,25$ h = 8760 h = 365 d) für eine Anzahl von 210 Haushalten, wobei der Gewichtungsfaktor der Musterkomponente $\beta_M = 0.60$ und der Gewichtungsfaktor der stochastischen Komponente $\beta_S = 0.30$ beträgt. Zur

Vergleichbarkeit der Analyseergebnisse wurden die Zeitreihenlänge, die Abtastperiodendauer, die Anzahl an Bereichen mit fehlenden Werten und die untere und obere Grenze des Ausreißermultiplikators für alle zehn Benchmarkdatensätze konstant gehalten. Die Zufallsvariablen X_E^n , $X_{nan,g}^n$, $X_{nan,s}^n$, X_a^n und X_{ha}^n sorgen dafür, dass kein Lastprofil eines Haushaltes identisch mit dem eines anderen Haushaltes ist. Zum anderen wird durch das Konstanthalten der Parameterwerte σ_E^2 , $\sigma_{nan,g}^2$, $\sigma_{nan,s}^2$, σ_a^2 , a und b sichergestellt, dass sich die Lastprofile nicht beliebig stark voneinander unterscheiden können. Die Beeinflussung des Verbrauchsverhaltens der Haushalte erfolgt für alle Benchmarkdatensätze durch synthetische Strompreise (vgl. Abb. 2.4). Es handelt es sich hierbei um drei Tarifgruppen, wobei eine davon die Kontrollgruppe mit einem konstanten Strompreis (FIXED) ist. Die beiden anderen Tarifgruppen „TOU2“ (große *on-peak/off-peak* Strompreisspreizung und lange Tarifstufendauer) und „TOU3“ (geringe *on-peak/off-peak* Strompreisspreizung und kurze Tarifstufendauer) weisen statische Mehrstufentarife auf. Der Verlauf und die Höhe des Strompreises sind in Abb. 2.4 zu sehen.

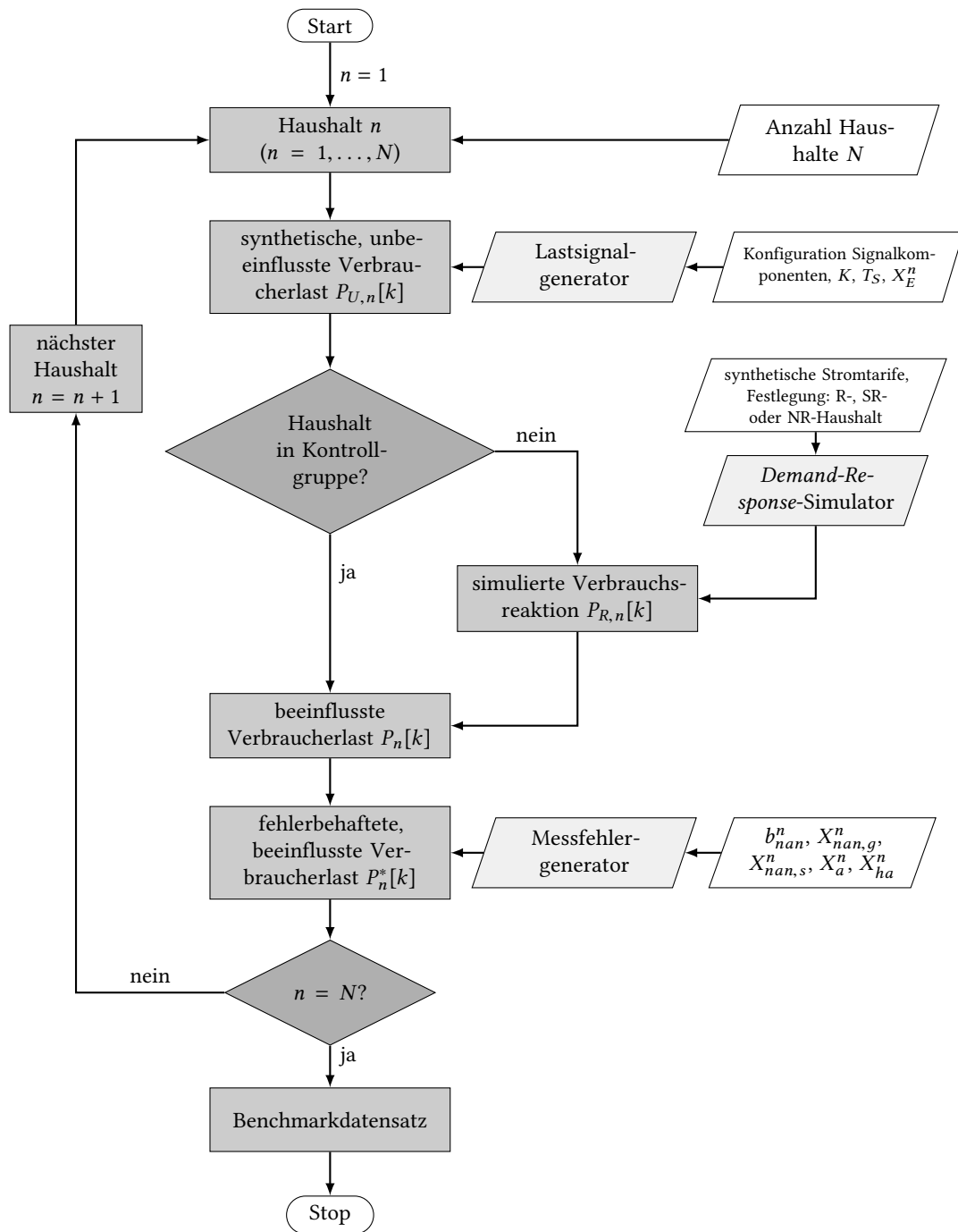


Abbildung 2.5: Flussdiagramm für die Generierung eines Benchmarkdatensatzes: R - Responder-Haushalt, SR - Semi-Responder-Haushalt, NR - Non-Responder-Haushalt, b_{nan}^n - Anzahl an Bereichen mit fehlenden Werten, K - Länge der Zeitreihe, n - Haushalts-/ Verbraucherindex, N - Anzahl der Haushalte, $P_{U,n}[k]$ - unbeeinflusster Anteil der Verbraucherlast, $P_{R,n}[k]$ - beeinflusster Anteil der Verbraucherlast, $P_n[k]$ - beeinflusste Verbraucherlast, $P_n^*[k]$ - fehlerbehaftete, beeinflusste Verbraucherlast, T_S - Abtastperiodendauer, $\{X_E^n, X_{nan,g}^n, X_{nan,s}^n, X_a^n\}$, - normalverteilte Zufallsvariable des Gesamtenergieinhalts der Zeitreihe, der Verbrauchsreaktion, des Anteils an Bereichen mit fehlenden Werten, des Anteils an fehlenden Werten und des Anteils an Ausreißern, X_{ha}^n - gleichverteilte Zufallsvariable der Höhe der Ausreißer.

2.4 Verfahrensübergreifende Methoden der Pipeline

Die datengetriebene Analyse und Prognose des preisbeeinflussten Verbrauchsverhaltens setzt eine Vorverarbeitung und Bereinigung der Messdaten des *Smart-Meter*-Datensatzes voraus (vgl. Abschnitt 1.2.6). Erst durch diese Vorverarbeitungsschritte können nachfolgende Methoden der analyse- und prognoseorientierten Pipeline angewandt werden. Da die Vorverarbeitungsschritte unabhängig von den nachfolgenden Methoden durchzuführen sind, werden sie unter dem Begriff „Verfahrensübergreifende Methoden der Pipeline“ zusammengefasst.

Im Einzelnen umfassen die verfahrensübergreifenden Methoden der Pipeline die Datenvorverarbeitung (Abschnitt 2.4.1), die Zeitreihenzerlegung (Abschnitt 2.4.2), die Zeitreihennormierung (Abschnitt 2.4.3) und die Zeitreihenaggregation (Abschnitt 2.4.4). Abschnitt 2.4 endet mit einer Zusammenfassung über die verfahrensübergreifenden Methoden der Pipeline.

2.4.1 Datenvorverarbeitung

Wie bereits in Abschnitt 1.2.6 erläutert, existieren zahlreiche etablierte Verfahren zum Umgang mit fehlenden Werten und mit Ausreißern. Allerdings hängt die Entscheidung, welches Verfahren zum Einsatz kommt, von vielen Faktoren (z.B. Abtastperiodendauer, Anzahl der Datentupel, Länge und Charakteristik der Messzeitreihen) ab und muss für jeden Datensatz und jede Energiezeitreihe gesondert bewertet werden. In der Literatur fehlt bislang eine Anleitung für die Datenvorverarbeitung bei preisbeeinflussten Haushaltslastkurven.

Energiezeitreihen, wie beispielsweise Zeitreihen der Haushaltslast [209, 233] oder Zeitreihen der Leistung aus Erneuerbare-Energien-Anlagen [97, 234], haben i.d.R. eine Abtastperiodendauer im Bereich von 5-30 Minuten. Bei einzelnen fehlenden Messwerten innerhalb dieser Zeitreihen sind Interpolationsverfahren geeignete Verfahren zur Berechnung der Ersatzwerte der *NaN*-Werte, da dadurch wenig bis keine Information verloren geht und diese Verfahren leicht anzuwenden sind. Die Entscheidung, bis zu welcher Anzahl an fehlenden Werten Interpolationsverfahren sinnvolle Ersatzwerte liefern, hängt auch von der Charakteristik der Messzeitreihen ab. Wie in Tabelle B.2 dargestellt, haben elektrische Anwendungen, die sich zur Lastverlagerung in Haushalten eignen, erhebliche Spitzenlasten, die zu starken Schwankungen im Lastprofil führen. Diese Lastcharakteristik führt dazu, dass bei größeren Bereichen mit fehlenden Messwerten äußerst sorgsam bei der Wahl des Imputationsverfahrens und bei der Ausreißerbehandlung vorgegangen werden muss.

Im Rahmen der vorliegenden Arbeit wurden zwei Testverfahren durchgeführt. Das erste Testverfahren soll Aufschluss darüber geben, welche Imputationsverfahren bei fehlerbehafteten Lastzeitreihen idealerweise zum Einsatz kommen sollten und wird im Folgenden vorgestellt. Das zweite Testverfahren befasst sich mit der Ermittlung von optimalen Hampel-Filterparametern und wird in Abschnitt 2.4.1.2 beschrieben. Die Basis der Testverfahren stellen zwei Benchmarkdatensätze dar, die mithilfe des Benchmarkgenerators (siehe Abb. 2.5) und den in den Tabellen B.5 und B.6 gezeigten Parametereinstellungen erzeugt wurden.

2.4.1.1 Testverfahren zur Bewertung verschiedener Imputationsverfahren bei fehlerbehafteten Lastzeitreihen

Wie eben beschrieben, dient das erste Testverfahren dem Vergleich verschiedener Imputationsverfahren bei fehlerbehafteten Lastzeitreihen. Als Imputationsverfahren wurden folgende Methoden aus Tabelle 1.3 getestet: Mittelwertabschätzung, *Nearest Historical Average*, LOCF, NOCB, NEOI, lineare Interpolation, pchip-Verfahren, DCT-PLS-Verfahren und zwei AR-Modelle⁹. Für die Evaluierung wurde die Wurzel der mittleren quadratischen Abweichung und der Korrelationskoeffizient zwischen der originalen (d.h., ohne fehlende Werte) und der bereinigten Jahreslastzeitreihe verwendet. Die Formeln zur Berechnung der Wurzel der mittleren quadratischen Abweichung und des Korrelationskoeffizienten stehen in Tabelle 2.11.

Die Ergebnisse des ersten Testverfahrens sind in Tabelle B.8 und in Abb. 2.6, B.3 und B.4 dargestellt. Das Imputationsverfahren „Mittelwertabschätzung“ lieferte in den meisten Fällen sehr schlechte Ergebnisse, was dazu führt, dass sich der Wertebereich der Ergebnisse des Imputationsverfahrens „Mittelwertabschätzung“ stark von den Wertebereichen der Ergebnisse der anderen Imputationsverfahren unterscheidet. Aus Gründen der Lesbarkeit wurden deshalb alle Werte des Imputationsverfahrens „Mittelwertabschätzung“ aus Tabelle B.8 und den Abb. 2.6, B.3 und B.4 entfernt. Aus den Abb. 2.6, B.3 und B.4 lassen sich folgende Aussagen ableiten:

- Bei steigenden Anteilen an fehlenden Werten ($\mu_{nan,s}$) und zusammenhängenden fehlenden Werten ($\mu_{nan,g}$) nimmt die Abweichung zwischen den originalen (d.h., ohne fehlende Werte) und bereinigten Jahreslastzeitreihen im Mittel zu.
- $b_{nan}^n = 0$: Befindet sich in den Jahreslastzeitreihen kein Bereich mit fehlenden Werten, schätzt das *Nearest-Historical-Average*-Verfahren im Vergleich zu allen anderen Verfahren am schlechtesten die Ersatzwerte für die NaN-Werte. Damit die Unterschiede zwischen den anderen acht Imputationsverfahren („LOCF“, „NOCB“, „NEOI“, „lin“, „pchip“, „DCTPLS“, „AR96“ und „AR672“) besser erkannt werden können, wurde das Imputationsverfahren „*Nearest Historical Average*“ aus den Abb. 2.6a, 2.6c und 2.6e entfernt. Die fünf Imputationsverfahren „lin“, „pchip“, „DCTPLS“, „AR96“ und „AR672“ eignen sich gleichermaßen zur Schätzung der Ersatzwerte für die NaN-Werte. Die drei Verfahren „LOCF“, „NOCB“ und „NEOI“ schneiden dagegen schlechter ab.
- $1 \leq b_{nan}^n \leq 5$: Befinden sich in den Jahreslastzeitreihen Bereiche mit fehlenden Werten, liefern die Imputationsverfahren „*Nearest Historical Average*“, „AR96“ und „AR672“ die besten Ersatzwerte für die NaN-Werte.
- Die Signalcharakteristik der Lastzeitreihe (HH001 mit $\beta_M = 0.90$, $\beta_S = 0.00$; HH002 mit $\beta_M = 0.63$, $\beta_S = 0.27$; HH003 mit $\beta_M = 0.36$, $\beta_S = 0.54$) hat lediglich einen geringen Einfluss auf die Qualität der Imputation. Sowohl bei Jahreslastzeitreihen ohne stochastischen Anteil (HH001) als auch bei Jahreslastzeitreihen mit einem hohen stochastischen Anteil (HH003) liefern die Imputationsverfahren sinnvolle Ersatzwerte für die NaN-Werte.

⁹Bei den zwei autoregressiven Modellen handelt es sich zum einen um ein AR-Modell mit der maximalen Ordnung $n_a = 96$ („AR96“), zum anderen um ein AR-Modell mit der maximalen Ordnung $n_a = 672$ („AR672“). Die mathematische Beschreibung des AR-Modells ist im Anhang auf Seite 185 zu finden.

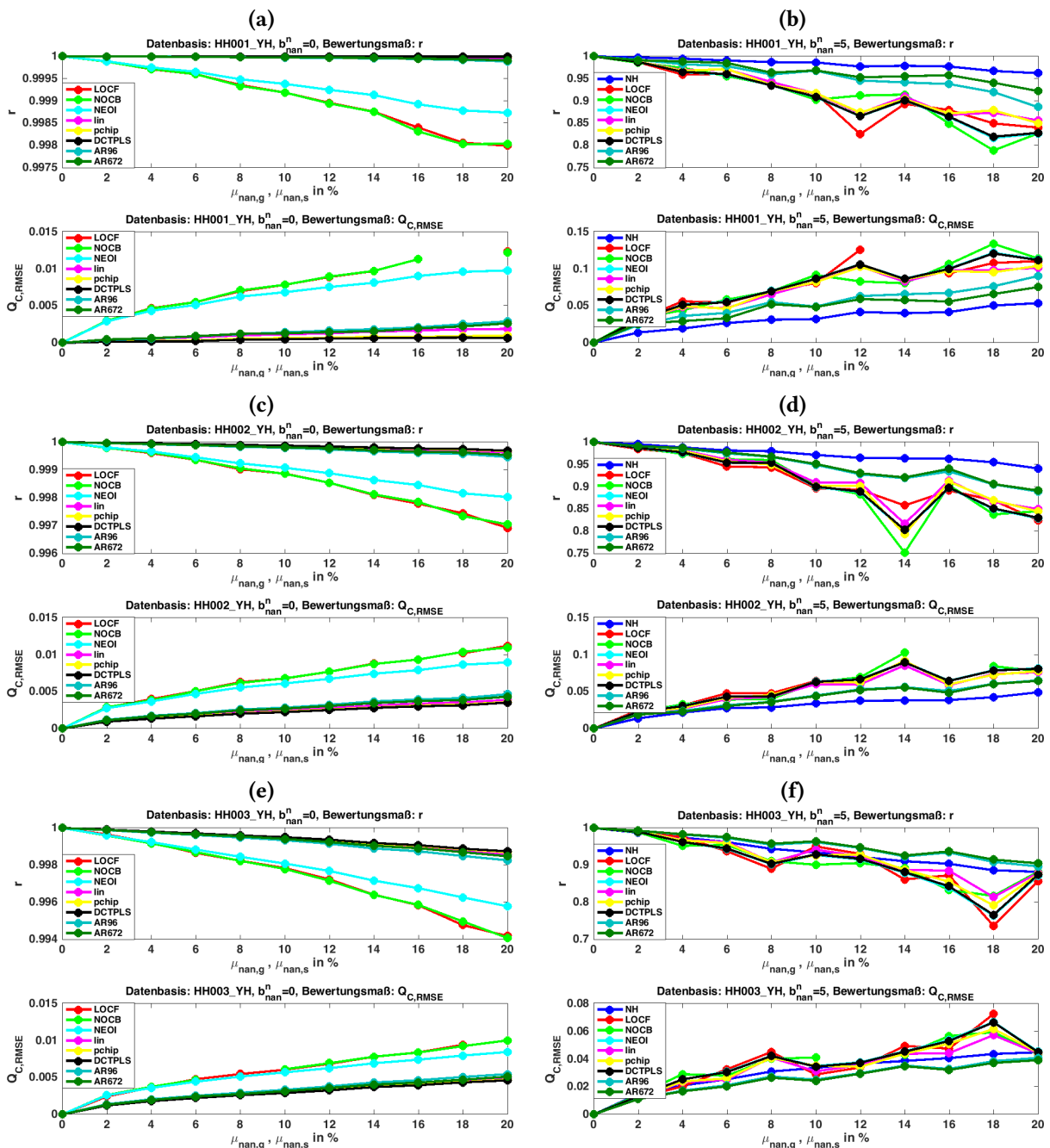


Abbildung 2.6: Vergleich verschiedener Imputationsverfahren bei fehlerbehafteten Lastzeitreihen für $b_{nan}^n = 0$ und $b_{nan}^n = 5$ (Werte aus Tabelle B.8, Ergebnisse für $b_{nan}^n = 1, 2, 3, 4$ in Abb. B.3 und B.4, Beschreibung der Imputationsverfahren in Tabelle 1.3); Abkürzungen: NH - Nearest Historical Average, LOCF - Last-Observation-Carried-Forward, NOCB - Next-Observation-Carried-Backward, NEOI - nächstgelegene Nicht-NaN-Wert, lin - lineare Interpolation zwischen benachbarten Nicht-NaN-Werten, pchip - Piecewise Cubic Hermite Interpolating Polynomial, DCTPLS - Penalized-Least-Square-Regression auf Basis einer diskreten Kosinustransformation, {AR96,AR672} - AR-Modell mit unterschiedlichen Ordnungen.

Zusammenfassend lässt sich sagen, dass bei einzelnen fehlenden Werten eines der fünf Verfahren „lin“, „pchip“, „DCTPLS“, „AR96“ und „AR672“ als Imputationsverfahren verwendet werden

sollte. Bei größeren Bereichen mit fehlenden Messwerten (> 30 min) ist es ratsam, Imputationsverfahren, wie das *Nearest-Historical-Average-Verfahren* oder ein autoregressives Modell zu verwenden. Im Rahmen der vorliegenden Arbeit wurde deshalb für die analysierten (realen) *Smart-Meter-Datensätze* ein autoregressives Modell („AR672“) zur Schätzung von Ersatzwerten für *NaN*-Werte verwendet.

Alternativ können die Gesamtzeitreihen in kürzere Zeitreihen mit der größtmöglichen Zeitreihenlänge ohne Bereiche mit fehlenden Messwerten unterteilt werden. Somit wird einerseits gewährleistet, dass möglichst lange Zeitreihen für die Datenanalyse (vgl. Abschnitt 2.5) und Modellierung (vgl. Abschnitt 2.6) zur Verfügung stehen. Andererseits wird dadurch verhindert, dass die größeren Bereiche mit fehlenden Messwerten (z.B. mehrere Tage) durch Imputationsverfahren zu stark verfälscht werden.

Das Löschen der gesamten Messzeitreihe bzw. des fehlerhaften Datentupels sollte als letzte der in Abschnitt 1.2.6 genannten Strategien gewählt werden. Der Löschschwelligwert gibt dabei das prozentuale Verhältnis zwischen der Anzahl an fehlenden Werten und der Gesamtlänge der Messzeitreihe an, bei dem das Löschen der gesamten Messzeitreihe bzw. des fehlerhaften Datentupels erfolgt. Unter Umständen sollte das Löschen der gesamten Messzeitreihe erst nach der Zeitreihenzerlegung erfolgen (siehe Beispiel 2 in Abschnitt 2.4.2).

2.4.1.2 Testverfahren zur Ermittlung von optimalen Hampel-Filterparametern

Nachdem im ersten Testverfahren verschiedene Imputationsverfahren verglichen wurden, befasst sich das zweite Testverfahren mit der Behandlung von Ausreißern in Lastzeitreihen. In der vorliegenden Arbeit wurde der robuste Hampel-Test oder Hampel-Filter als Methode zur Ausreißerererkennung angewendet [102, 232], der im Folgenden beschrieben wird.

Gegeben sei eine Zeitreihe $x[k]$ der Länge K mit $k = 1, \dots, K$. Nach Vorgabe der Hampel-Filterbreite h mit $2h + 1 \leq K$ wird der Median des Zeitreihenausschnitts $(x[k - h], \dots, x[k + h])$ berechnet. Liegt der Zeitreihenausschnitt an einem der beiden Enden der Zeitreihe muss eine wie in Gleichung (2.24) gezeigte Fallunterscheidung zur Berechnung des Medians getroffen werden.

$$\tilde{x} = \begin{cases} \text{median}(x[1], \dots, x[k + h]) & , \text{ für } k < h + 1 \\ \text{median}(x[k - h], \dots, x[k + h]) & , \text{ für } h + 1 \leq k \leq K - h \\ \text{median}(x[k - h], \dots, x[K]) & , \text{ für } k > K - h \end{cases} \quad (2.24)$$

Alternativ zur asymmetrischen Berechnung des Medians an den beiden Enden der Zeitreihe könnte dort auf die Ausreißerbehandlung verzichtet werden, was allerdings dazu führt, dass mögliche Ausreißer nicht erkannt und ersetzt werden. Die asymmetrische Berechnung wird in Kauf genommen, um möglichst viele Ausreißer erkennen zu können.

Im zweiten Schritt wird der Zentralwert der absoluten Abweichungen (engl.: *Median Absolute Deviation, MAD*) der einzelnen Messwerte des Zeitreihenausschnitts vom eben berechneten Zentralwert \tilde{x} gemäß

$$MAD = \begin{cases} \text{median}(|\tilde{x} - x[1]|, \dots, |\tilde{x} - x[k+h]|) & , \text{ für } k < h+1 \\ \text{median}(|\tilde{x} - x[k-h]|, \dots, |\tilde{x} - x[k+h]|) & , \text{ für } h+1 \leq k \leq K-h \\ \text{median}(|\tilde{x} - x[k-h]|, \dots, |\tilde{x} - x[K]|) & , \text{ für } k > K-h \end{cases} \quad (2.25)$$

bestimmt. Ein Messwert $x[k]$ wird nun als Ausreißer erkannt, wenn

$$|\tilde{x} - x[k]| \geq C_{h,\alpha}^{Ham} \cdot MAD, \quad (2.26)$$

gilt. $C_{h,\alpha}^{Ham}$ ist der kritische Wert nach Hampel, der von der Hampel-Filterbreite h und dem Signifikanzniveau α abhängt. Nach der Detektion des Ausreißers wird dieser durch den Zentralwert \tilde{x} des Zeitreihenausschnitts ersetzt. Der Hampel-Filter „wandert“¹⁰ anschließend um einen Abtastzeitpunkt weiter zum nächsten Zeitreihenausschnitt, wo eine neue Ausreißerdetektion beginnt.

Da in der Literatur keine konkreten Angaben über die Wahl der beiden Filterparameter h und $C_{h,\alpha}^{Ham}$ bei Lastzeitreihen zu finden sind, wurde im Rahmen der Arbeit ein Testverfahren zur Ermittlung von optimalen Hampel-Filterparametern durchgeführt. Die Basis des Testverfahrens stellen unbeeinflusste Jahreslastzeitreihen mit variierenden Abtastperiodendauern, variierenden Anteilen an Ausreißern und variierender Lastsignalcharakteristik dar. Die Parametereinstellungen des Benchmarkgenerators zur Erzeugung dieser Jahreslastzeitreihen sind in Tabelle B.6 dargestellt. Dabei wurde

- die Abtastperiodendauer in vier Ausprägungen ($T_S = 5 \text{ min}$, $T_S = 15 \text{ min}$, $T_S = 30 \text{ min}$, $T_S = 60 \text{ min}$),
- die Anteile an Ausreißern in vier Ausprägungen ($\mu_a = 2\%$, $\mu_a = 4\%$, $\mu_a = 6\%$, $\mu_a = 8\%$) und
- die Lastsignalcharakteristik in drei Ausprägungen,
 - mit einem Standardlastprofil ohne Rauschen (Musteranteil 90%, stochastischer Anteil 0%, Trendanteil 5%, Periodischer Anteil 5%, exogener Anteil 0%),
 - mit einem Standardlastprofil mit Rauschen (Musteranteil 63%, stochastischer Anteil 27%, Trendanteil 5%, Periodischer Anteil 5%, exogener Anteil 0%) und
 - mit einem Standardlastprofil mit starkem Rauschen (Musteranteil 36%, stochastischer Anteil 54%, Trendanteil 5%, Periodischer Anteil 5%, exogener Anteil 0%),

variiert und so insgesamt 48 verschiedene Jahreslastzeitreihen (siehe Tabelle B.9) generiert. Die konstanten Parameter $\mu_E = 3.2$, $\sigma_E^2 = 0$, $\mu_{nan} = 0$, $\sigma_{nan}^2 = 0$, $\sigma_a^2 = 0$, $\beta_T = 5\%$, $\beta_P = 5\%$ und $\beta_E = 0\%$ stehen aus Platzgründen nicht in Tabelle B.9. Die Lage der Ausreißer wurde mithilfe einer Zufallsvariablen gleichmäßig über der Zeitreihenlänge verteilt. Die Höhe der Ausreißer wurde gemäß Gleichung (2.22) mit $a = 1.5$ und $b = 5$ festgelegt.

¹⁰Der Hampel-Filter bewegt sich mit einer bestimmter Filterbreite abschnittsweise durch die Zeitreihe und wird somit im Englischen als *moving window algorithm* bezeichnet.

Für jede der 48 Jahreslastzeitreihen wurden der Hampel-Filterparameter h mit $h = 2, 3, \dots, 12$ und der Hampel-Filterparameter $C_{h,\alpha}^{Ham}$ mit $C_{h,\alpha}^{Ham} = 2, 3, \dots, 12$ variiert und die ausreißerbehafte Lastzeitreihe bereinigt. Anschließend wurde die Wurzel der mittleren quadratischen Abweichung $Q_{C, RMSE}$ und der Korrelationskoeffizient r zwischen der originalen (d.h., ohne Ausreißer) und der bereinigten Jahreslastzeitreihe berechnet. Die Formeln zur Berechnung von $Q_{C, RMSE}$ und r stehen in Tabelle 2.11. Aus den 121 Parameterkombinationen¹¹ je Jahreslastzeitreihe wurde nun diejenige Parameterkombination als die optimale ausgewählt, bei der der Wert für $Q_{C, RMSE}$ am geringsten ist. Die entsprechenden optimalen Werte für h und $C_{h,\alpha}^{Ham}$ und die Werte für $Q_{C, RMSE}$ (nach Hampel) und r (nach Hampel) sind in Tabelle B.9 dargestellt. Zudem sind zwei weitere Spalten mit der Wurzel der mittleren quadratischen Abweichung $Q_{C, RMSE}$ (vor Hampel) und dem Korrelationskoeffizienten r (vor Hampel) zwischen der originalen (d.h., ohne Ausreißer) und der ausreißerbehafte Jahreslastzeitreihe angegeben.

Die geringen Werte für $Q_{C, RMSE}$ (nach Hampel) zeigen, dass der Hampel-Filter bei allen Zeitreihen die meisten Ausreißer erkennt und geeignete Ersatzwerte für die Ausreißer liefert. Die Auswertung von Tabelle B.9 ergibt, dass in ca. 90% der 48 Jahreslastzeitreihen ein optimaler Hampel-Filterparameter $C_{h,\alpha}^{Ham}$ von $C_{h,\alpha}^{Ham} = 2$ zu wählen ist. Zudem stellt sich in 35 von 48 Fällen eine optimale Filterbreite h von $h = 3$ oder $h = 4$ ein. Die Ergebnisse in Tabelle B.9 lassen zudem den Schluss zu, dass bei einer Erhöhung des Anteils an Ausreißern die Filterbreite h des Hampel-Filters zu erhöhen ist. Abschließend gilt festzuhalten, dass für die Anwendung des neuen Konzeptes an realen und unbekanntem *Smart-Meter*-Datensätzen die beiden Filterparameter auf die Werte $h = 4$ und $C_{h,\alpha}^{Ham} = 2$ festgesetzt werden. Die Anwendung des Hampel-Filters (mit den gewählten Filterparametern) an unbekanntem *Smart-Meter*-Datensätzen ist insofern Erfolg versprechend, als die getesteten Jahreslastzeitreihen eine realistische Datenbasis hinsichtlich der Lastsignal- und Ausreißercharakteristik (vgl. Tabelle B.6) darstellen. Beispiel 2 verdeutlicht die Schritte der Datenvorverarbeitung an einem Beispiellastsignal.

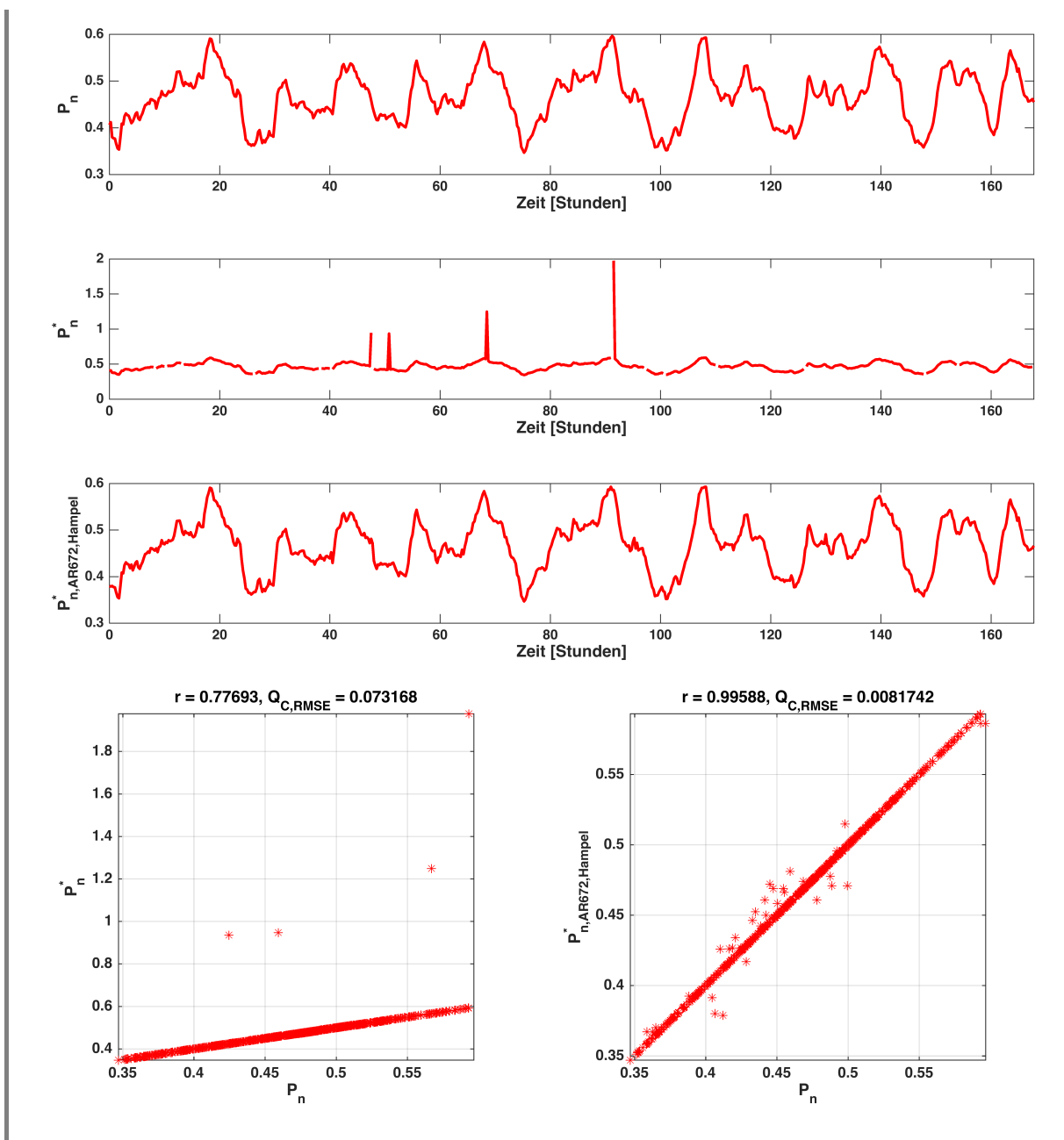
2.4.1.3 Sonstige Datenvorverarbeitungsschritte

Neben den beiden Fehlerarten „fehlender Messwert“ und „Ausreißer“ wurde in den untersuchten *Smart-Meter*-Datensätzen ein weitere Fehlerart registriert. Die Rohdaten beinhalteten negative Werte für die Haushaltslast, welche durch den Wert 0 ersetzt wurden.

Beispiel 2: Datenvorverarbeitung eines Beispiellastsignals

Zur Verdeutlichung der Schritte der Datenvorverarbeitung aus Abschnitt 2.4.1 wird die Datenvorverarbeitung (Imputationsverfahren: AR672; Ausreißerbehandlung: Hampel-Filter) an einem Beispiellastsignal ($T_S = 15$ min, $K = 35040$, $\beta_T = 5\%$, $\beta_M = 45\%$, $\beta_S = 40\%$, $\beta_P = 5\%$, $\beta_E = 5\%$, $\mu_E = 3.2$) angewendet. Aus dem originalen Lastsignal P_n wird durch den Messfehlergenerator ($b_{nan}^n = 2$, $\mu_{nan,g} = 3$, $\sigma_{nan,g}^2 = 0$, $\mu_{nan,s} = 3$, $\sigma_{nan,s}^2 = 0$, $\mu_{nan,a} = 1.5$, $\sigma_{nan,a}^2 = 1.5$, $a = 0.5$, $b = 4$) ein fehlerbehaftetes Lastsignal P_n^* berechnet. Die Datenvorverarbeitung erkennt Ausreißer und fehlende Werte und berechnet das Lastsignal $P_{n,AR672,Hampel}^*$.

¹¹Die Zahl 121 ergibt sich aus der Kombination der elf Werte für h mit den elf Werten für $C_{h,\alpha}^{Ham}$.



2.4.2 Zeitreihenzerlegung

Die Zeitreihen eines *Smart-Meter*-Datensatzes werden bei der Zeitreihenzerlegung anhand des Zeitstempels in kürzere Zeitreihenabschnitte unterteilt. Die Länge K_s der kürzeren Zeitreihenabschnitte ergibt sich aus der Multiplikation der Abtastfrequenz f_s [1/h] und der Länge ΔT [h] der Zeitreihenabschnitte. Die kürzeren Zeitreihen sind wie folgt definiert:

$$x_{j, n_D}[k], \text{ mit } k = 1, \dots, K_s \quad (2.27)$$

$$\text{und } n_D = 1, \dots, N_D \cdot \frac{K}{K_s}.$$

Grundsätzlich kann die Länge der Zeitreihenabschnitte beliebig gewählt werden. Typischerweise werden Zeitreihen in Abschnitte unterteilt, die einen periodischen Verlauf aufweisen, wie z.B. Tageskurven, Wochenkurven oder Jahreskurven. Dabei entscheidet die Aufgabenstellung über die Länge der Zeitreihenabschnitte. Für die Aufgabenstellungen A1 bis A3 (vgl. Seite 32) ist es von Vorteil, wenn die Lastkurven der Haushalte als Tageslastkurven vorliegen, da mit den analyseorientierten Methoden der Pipeline typische Muster für das preisbeeinflusste Verbrauchsverhalten identifiziert werden sollen [147]. So wird z.B. die Jahreslastkurve $P_n[k]$ eines Haushaltes n in D Tageskurven $P_{d,n}[k]$ mit $d = 1, \dots, D$ zerlegt.

Wie bereits in Abschnitt 2.4.1 angedeutet, sollte das Löschen von Messzeitreihen ab einer bestimmten Anzahl an fehlenden Messwerten unter Umständen erst nach der Zeitreihenzerlegung durchgeführt werden, da so einzelne Tageskurven erhalten bleiben. In Beispiel 3 wird anhand eines Minimalbeispiels die Problematik von ungleichmäßig verteilten fehlenden Werten bei der Datenvorverarbeitung und Zeitreihenzerlegung verdeutlicht.

Beispiel 3: Ungleichmäßig verteilte fehlende Werte

Es ist in Abb. 2.7 eine synthetische Lastkurve ($\beta_T = 5\%$, $\beta_M = 65\%$, $\beta_S = 20\%$, $\beta_P = 5\%$, $\beta_E = 5\%$, $\mu_E = 3.2$) über einen Zeitraum von drei Tagen ($K = 288$) bei einer Abtastfrequenz von $f_s = 4/\text{h}$ dargestellt. Die roten Kreise kennzeichnen 32 fehlende Messwerte, wobei alle im Zeitraum des zweiten Tages ($k = 97, \dots, 192$) liegen. Bezogen auf die Gesamtlänge der Zeitreihe fehlen somit $\frac{32}{288} \approx 11\%$ aller Messwerte. Bei Annahme eines Löschschwelliges für die gesamte Messzeitreihe von 10% hätte dies zu der Löschung der gesamten Messzeitreihe geführt. Wird die Löschung der fehlerbehafteten Messzeitreihe allerdings erst nach der Zeitreihenzerlegung durchgeführt, würden für das gezeigte Beispiel zwei der drei Tageszeitreihen erhalten bleiben und nur die mittlere Tageszeitreihe ($k = 97, \dots, 192$) mit 33% fehlenden Werten gelöscht werden. Es muss allerdings angemerkt werden, dass so keine lückenlose, chronologische Abfolge der Tageszeitreihen gewährleistet werden kann.

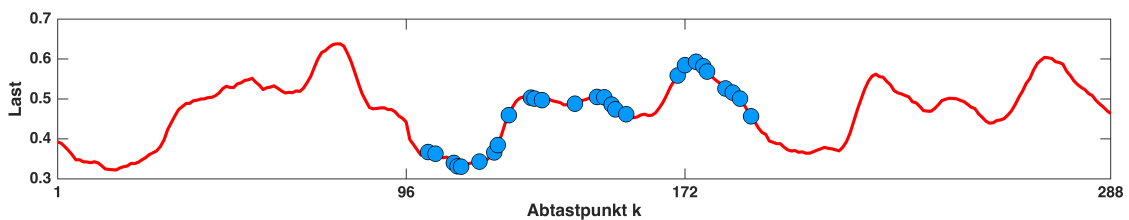


Abbildung 2.7: Veranschaulichung der Problematik von ungleichmäßig verteilten fehlenden Werten bei der Datenvorverarbeitung und Zeitreihenzerlegung: trotz $\frac{32}{288} \approx 11\%$ fehlender Messwerte (gekennzeichnet mit Kreisen) in der gesamten Messzeitreihe können die Tage 1 und 3 vollständig erhalten bleiben, wenn das Löschen der fehlerbehafteten Messzeitreihe erst nach der Zeitreihenzerlegung erfolgt.

2.4.3 Zeitreihennormierung

Nach der Beschreibung der Zeitreihenzerlegung im vorherigen Abschnitt wird in Abschnitt 2.4.3 auf die Normierung von Energiezeitreihen eingegangen. Unter dem Begriff der Normierung von Zeitreihen werden in der vorliegenden Arbeit alle Verfahren zusammengefasst, die das Ziel haben, Verzerrungen in Zeitreihen in *Smart-Meter*-Datensätzen zu erkennen und geeignet

zu beseitigen. Energiezeitreihen (z.B. Lastzeitreihen, Zeitreihen der volatilen Stromerzeugung) können eine Vielzahl von Verzerrungen aufweisen, wie beispielsweise

- die Verschiebung um einen Offset in y-Richtung,
- die Stauchung oder Streckung in y-Richtung oder
- das Vorhandensein eines Trends.

Vor der datengetriebenen Analyse (vgl. Abschnitt 2.5) und Modellierung (vgl. Abschnitt 2.6) des preisbeeinflussten Verbrauchsverhaltens müssen geeignete Normierungsmethoden ausgewählt und eingesetzt werden. Tabelle B.10 liefert eine Übersicht über gängige Verfahren der Normierung von Zeitreihen. Darüber hinaus bietet [235] eine umfassende Analyse von varianzstabilisierenden Zeitreihentransformationen. Die Herausforderung beim Einsatz der Normierungsverfahren besteht darin, dass einerseits die Bearbeitung der Aufgabenstellungen A1 bis A4 (vgl. Seite 32) unterstützt wird, und andererseits die potentiell interessanten Informationen (z.B. Zeitpunkt der Spitzenlast) aus den Zeitreihen nicht eliminiert werden. Die in *Smart-Meter*-Datensätzen zusammengefassten Haushaltslastkurven unterscheiden sich z.T. sehr stark hinsichtlich des Verlaufs und der Absolutwerte. Um aus einer Vielzahl von Zeitreihen Aussagen über unbeeinflusstes und beeinflusstes Verbrauchsverhalten treffen zu können, ist es unerlässlich, die Zeitreihen zu normieren. In der vorliegenden Arbeit wurden aus der Vielzahl an Normierungsverfahren die Verfahren „Mittelwertdivision“ (Index normmean) und „Z-Score-Normalisierung“ (Index norm) ausgewählt.

Die Normierung von Tageslastkurven $P_{d,n}[k]$ durch die Division des Mittelwerts $\bar{P}_{d,n}$ stellt sich wie folgt dar:

$$P_{d,n,\text{normmean}}[k] = \frac{P_{d,n}[k]}{\bar{P}_{d,n}}, \quad (2.28)$$

$$\text{mit } \bar{P}_{d,n} = \frac{1}{K} \sum_{k=1}^K P_{d,n}[k].$$

Es muss an dieser Stelle noch angemerkt werden, dass die Bezeichnungen „mittelwertnormierte Zeitreihe“ und „Normierung der Zeitreihe durch Division des Mittelwert“ synonym verwendet werden können.

Die Z-Score-Normalisierung eignet sich bei Tageslastzeitreihen mit unterschiedlichen Offsets und Amplituden und führt zu varianznormierten Tageslastzeitreihen. Die Z-Score-Normalisierung ist folgendermaßen definiert:

$$P_{d,n,\text{norm}}[k] = \frac{1}{\sigma_P} \cdot (P_{d,n}[k] - \bar{P}_{d,n}), \quad (2.29)$$

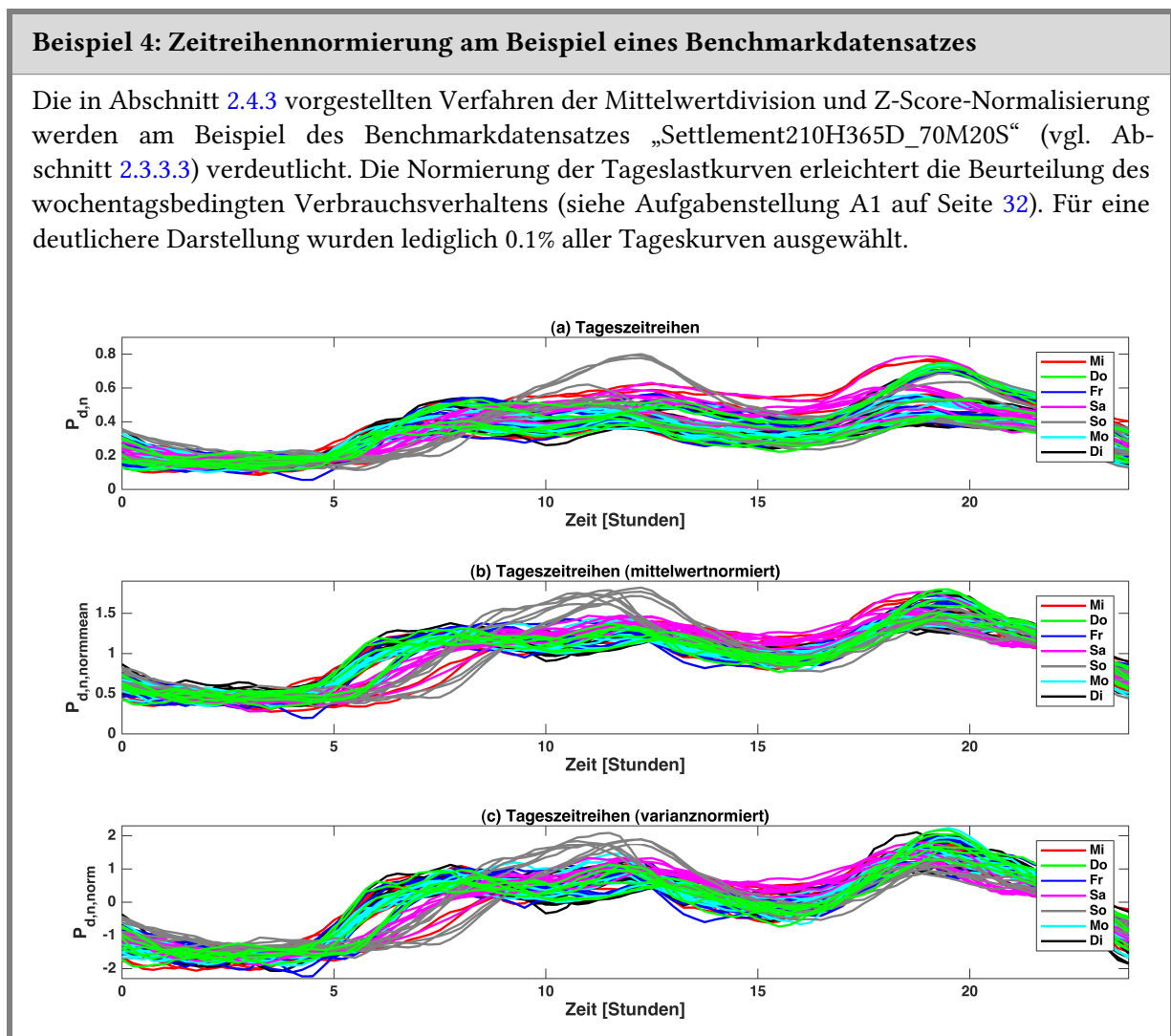
$$\text{mit } \sigma_P = \sqrt{\frac{1}{K-1} \sum_{k=1}^K (P_{d,n}[k] - \bar{P}_{d,n})^2}.$$

Dabei ist σ_P die Standardabweichung¹² einer Grundgesamtheit (engl. *population standard deviation*), die angibt, wie stark die Werte der Lastzeitreihe $P_{d,n}[k]$ um den Mittelwert $\bar{P}_{d,n}$ streuen.

Beide Normierungsmethoden eignen sich insbesondere, um in preisbeeinflussten Tageszeitreihen Lastverschiebungen (siehe *load shifting* in Abb. 1.4) zu erkennen, da der Einfluss des variierenden Tagesstromverbrauchs fast vollständig unterdrückt wird [76]. Der Verlauf der normierten Tageslastkurven lässt zudem Rückschlüsse auf das individuelle Verbrauchs- und Tagesmuster, wie Aufstehzeiten oder Abwesenheitszeiten, zu, was unter datenschutzrechtlichen und sicherheitstechnischen Gesichtspunkten kritisch zu sehen ist. In Beispiel 4 ist das Ergebnis der Normierung von Tageszeitreihen des Benchmarkdatensatzes „Settlement210H365D_60M30S“ dargestellt.

Beispiel 4: Zeitreihennormierung am Beispiel eines Benchmarkdatensatzes

Die in Abschnitt 2.4.3 vorgestellten Verfahren der Mittelwertdivision und Z-Score-Normalisierung werden am Beispiel des Benchmarkdatensatzes „Settlement210H365D_70M20S“ (vgl. Abschnitt 2.3.3.3) verdeutlicht. Die Normierung der Tageslastkurven erleichtert die Beurteilung des wochentagsbedingten Verbrauchsverhaltens (siehe Aufgabenstellung A1 auf Seite 32). Für eine deutlichere Darstellung wurden lediglich 0.1% aller Tageskurven ausgewählt.



¹²Wenn es sich bei den Daten um eine Stichprobe handelt, wird von der empirischen Standardabweichung (engl. *sample standard deviation*) gesprochen. In diesem Fall muss in Gleichung (2.29) der Nenner $K - 1$ durch K ersetzt werden.

2.4.4 Zeitreihenaggregation

Zeitreihenaggregation meint das Zusammenfassen von mehreren Zeitreihen zu einer Mittelwertzeitreihe. Zunächst werden für jede der m_y Klassen der Ausgangsgröße y_i alle $N_{D,b}$ Datentupel ausgewählt, die zur b -ten Klasse gehören. Die ausgewählten Datentupel werden anschließend zu einer Mittelwertzeitreihe $\bar{x}_b[k] = \bar{x}_b$ mit

$$\bar{x}_b[k] = \frac{1}{N_{D,b}} \sum_{n=1}^{N_{D,b}} x_{j,n}[k] \quad (2.30)$$

zusammengefasst, sprich aggregiert. Ein Beispiel für eine aggregierte Zeitreihe stellt das h0-Standardlastprofil dar. Dabei werden die Lastgänge einer Vielzahl von Haushalten zusammengefasst und als gemittelte Lastkurve angegeben, die das Verbrauchsverhalten möglichst vieler Haushalte repräsentieren soll.

Entscheidend bei der Zeitreihenaggregation ist die Wahl der Ausgangsgröße (z.B. „Haushalt“ oder „Wochentag“), nach deren Klassen (z.B. {„Haushalt01“, „Haushalt02“, . . . , „Haushalt99“} oder {„Montag“, „Dienstag“, . . . , „Sonntag“}) aggregiert wird. Im Rahmen der vorliegenden Arbeit wird die Zeitreihenaggregation sowohl für die Ausgangsgröße „Tarifgruppe“ als auch für die Ausgangsgröße „Haushalt“ durchgeführt. Im Folgenden wird nur auf die Zeitreihenaggregation für die Ausgangsgröße „Tarifgruppe“ eingegangen, wobei die Aggregation für die Ausgangsgröße „Haushalte“ in analoger Weise funktioniert.

Die mittlere Lastkurve $\bar{P}^{T_i}[k]$ aller N_{T_i} Haushalte der Tarifgruppe T_i ist definiert als

$$\bar{P}^{T_i}[k] = \frac{1}{N_{T_i}} \sum_{n=1}^{N_{T_i}} P_n[k], \quad \text{mit } k = 1, \dots, K. \quad (2.31)$$

Für die Identifikation von tarifbedingten Unterschieden für eine DR-Maßnahme (vgl. Aufgabenstellung A3 auf Seite 32) werden die Tageslastkurven $P_{d,n}[k]$ aller N_{T_i} Haushalte der Tarifgruppe T_i durch Mittelung aggregiert. Die mittlere Tageslastkurve $\bar{P}_d^{T_i}[k]$ aller Haushalte einer Tarifgruppe für einen bestimmten Tag d berechnet sich wie folgt:

$$\bar{P}_d^{T_i}[k] = \frac{1}{N_{T_i}} \sum_{n=1}^{N_{T_i}} P_{d,n}[k], \quad \text{mit } k = 1, \dots, K_s \quad (2.32)$$

und $d = 1, \dots, D$.

Die Gesamtlänge K einer Zeitreihe kann somit auch als Produkt $K = K_s \cdot D$ angegeben werden, wobei K_s die Länge der Tageszeitreihe ist. Durch eine weitere Aggregation von $\bar{P}_d^{T_i}[k]$ in Gleichung (2.32) über alle D Tage ergibt sich die mittlere Tageslastkurve aller Haushalte einer Tarifgruppe über alle Tage:

$$\bar{P}^{T_i}[k] = \frac{1}{D \cdot N_{T_i}} \sum_{d=1}^D \sum_{n=1}^{N_{T_i}} P_{d,n}[k], \quad \text{mit } k = 1, \dots, K_s. \quad (2.33)$$

Eine erste Abschätzung über das aggregierte preisbeeinflusste Verbrauchsverhalten einer Haushaltstarifgruppe liefert der Vergleich zwischen der aggregierten Zeitreihe $\bar{P}^{T_i}[k]$ der Tarifgruppe T_i und der aggregierten Zeitreihe $\bar{P}^{\text{FIXED}}[k]$ der Kontrollgruppe (FIXED), sofern vorhanden. Durch Umstellen von Gleichung (2.8) gemäß

$$\bar{P}_R^{T_i}[k] = \bar{P}^{T_i}[k] - \bar{P}_U^{T_i}[k] \quad (2.34)$$

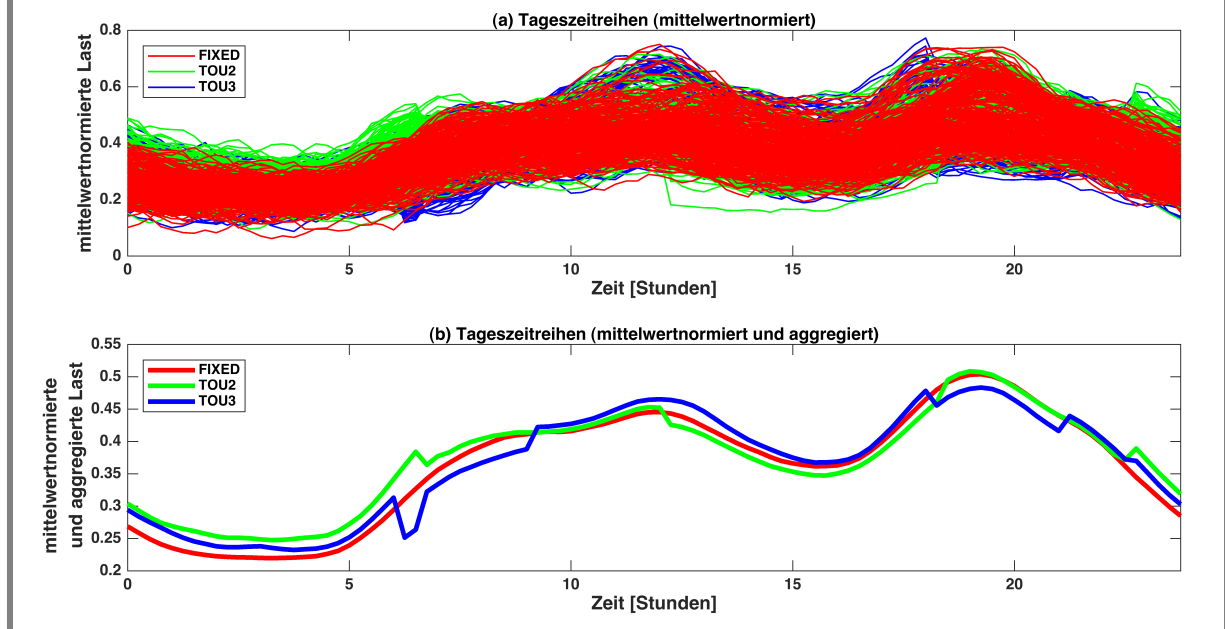
und mit $\bar{P}_U^{T_i}[k] = \bar{P}^{\text{FIXED}}[k]$ kann der beeinflusste Anteil der aggregierten Verbraucherlast in erster Näherung wie folgt geschätzt werden:

$$\hat{\bar{P}}_R^{T_i}[k] = \bar{P}^{T_i}[k] - \underbrace{\bar{P}^{\text{FIXED}}[k]}_{\bar{P}_U^{T_i}[k]}. \quad (2.35)$$

In Beispiel 5 wurde dieser Vergleich exemplarisch an einem Benchmarkdatensatz durchgeführt.

Beispiel 5: Zeitreihenaggregation am Beispiel eines Benchmarkdatensatzes

Im Benchmarkdatensatz „Settlement210H365D_50M40S“ (vgl. Abschnitt 2.3.3.3) ist das Verbrauchsverhalten von insgesamt 210 Haushalten repräsentiert, wobei 70 Haushalte der TOU2-Tarifgruppe, 70 Haushalte der TOU3-Tarifgruppe und 70 Haushalte der FIXED-Kontrollgruppe angehören. Für die Ausgangsgröße „Zugehörigkeit der Haushalte zu einer Tarifgruppe“ ergeben sich somit $m_y = 3$ Klassen („TOU2“, „TOU3“, „FIXED“). Die unten stehenden Abbildungen zeigen die Tageslastkurven vor und nach der Zeitreihenaggregation der Tarifgruppen „TOU2“, „TOU3“ und „FIXED“ des Benchmarkdatensatzes. Für eine deutlichere Darstellung wurden lediglich 1% aller Tageskurven ausgewählt.



Nachdem nun alle in der Arbeit verwendeten Datenvorverarbeitungsmethoden vorgestellt wurden, lassen sich *Smart-Meter*-Datensätze, wie in Tabelle 2.3 gezeigt, in Fälle mit unterschiedlicher Zeitreihenlänge und Aggregationsniveaus einteilen. Alle Fälle, bei denen eine

Zeitreihennormierung durchgeführt wurde, sind mit dem Index „norm“ (Varianznormierung) bzw. „normmean“ (Mittelwertnormierung) gekennzeichnet. Zudem ist eine weitere Differenzierung der in Tabelle 2.3 gezeigten Fälle denkbar, wie beispielsweise die Aufteilung des Falles DT in die Fälle DT^{wc} (alle Wochenendtage, *weekend days*) und DT^{wo} (alle Werktage, *working days*)

Tabelle 2.3: Strukturierung und Einteilung von *Smart-Meter*-Datensätzen in Fälle mit unterschiedlicher Zeitreihenlänge und Aggregationsniveaus (aus [76]): N_D - Anzahl der Datentupel, N - Anzahl der Haushalte, N_{T_i} - Anzahl an Haushalten der Tarifgruppe T_i , D - Anzahl der Tage, K - Länge der Gesamtzeitreihe als Anzahl an Abtastzeitpunkten. Die möglichen Ausgangsgrößen für die unterschiedlichen Aggregationen sind in Tabelle B.11 dargestellt.

	einzelne Haushalte (H)	alle Haushalte einer Tarifgruppe (T_i)
Gesamtzeitreihen (G)	Fall GH, GH _{norm} , GH _{normmean} ($N_D = N$)	Fall GT, GT _{norm} , GT _{normmean} ($N_D = N_{T_i}$)
Variable Zeitreihenlänge (V)	Fall VH, VH _{norm} , VH _{normmean} ($N_D = N$)	Fall VT, VT _{norm} , VT _{normmean} ($N_D = N_{T_i}$)
Tageszeitreihen (D)	Fall DH, DH _{norm} , DH _{normmean} ($N_D = N \cdot D$)	Fall DT, DT _{norm} , DT _{normmean} ($N_D = N_{T_i} \cdot D$)
Abtastzeitpunkte (S)	Fall SH, SH _{norm} , SH _{normmean} ($N_D = N \cdot D \cdot K$)	Fall ST, ST _{norm} , ST _{normmean} ($N_D = N_{T_i} \cdot D \cdot K$)

2.4.5 Zusammenfassung und Diskussion

In Abschnitt 2.4 wurden diejenigen Methoden der Methodenpipeline aus Abb. 2.1 beschrieben, die sowohl bei der analyseorientierten als auch bei der prognoseorientierten Pipeline zum Einsatz kommen und üblicherweise unter dem Begriff „Datenvorverarbeitung“ zusammengefasst werden. Im Einzelnen sind dies die Datenvorverarbeitung, die Zeitreihenzerlegung, die Zeitreihennormierung und die Zeitreihenaggregation.

Diese verfahrensübergreifenden Methoden sind für die datengetriebene Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens insofern unerlässlich, als reale Messdaten in den meisten Fällen nicht in der gewünschten Qualität und Form vorliegen. Aus diesem Grund bietet Abschnitt 2.4 eine Anleitung zum Umgang mit Energiezeitreihen. Insbesondere die Datenvorverarbeitung bei fehlerbehafteten Lastzeitreihen wurde intensiv in Abschnitt 2.4 behandelt. So dienten zwei Testverfahren in Abschnitt 2.4.1 der systematischen Analyse verschiedener Imputationsverfahren und der Ermittlung von optimalen Hampel-Filterparametern zur Erkennung von Ausreißern in Lastzeitreihen.

Im ersten Testverfahren konnte gezeigt werden, dass mehrere Imputationsverfahren zur Verfügung stehen, um sinnvolle Ersatzwerte für fehlende Messwerte zu liefern. Bei Verzicht auf diese Imputationsverfahren müssten fehlerhafte Datentupel, Teile von Messzeitreihen oder ganze Messzeitreihen gelöscht werden, was sich nachteilig auf den Umfang des Datensatzes

auswirkt. Das zweite Testverfahren wies nach, dass der Hampel-Filter ein geeignetes und robustes Verfahren ist, um Ausreißer in Energiezeitreihen zu erkennen und geeignete Ersatzwerte zu liefern.

Die Schritte der Zeitreihenzerlegung (Abschnitt 2.4.2) und der Zeitreihennormierung (Abschnitt 2.4.3) mit anschließender Zeitreihenaggregation (Abschnitt 2.4.4) führen dazu, dass die Rohdaten aus dem preisbeeinflussten Verbrauchsprozess als strukturierter *Smart-Meter*-Datensatz vorliegen. Die strukturierten *Smart-Meter*-Datensätze bilden die Ausgangsbasis für die datengetriebene Analyse des preisbeeinflussten Verbrauchsprozesses (Abschnitt 2.5) und die Modellierung des Verbrauchsverhaltens von Haushaltsstromkunden als Reaktion auf variable Stromtarife (Abschnitt 2.6).

2.5 Analyseorientierte Methoden der Pipeline

Ein Teilziel der vorliegenden Arbeit ist die datengetriebene Analyse des preisbeeinflussten Verbrauchsverhaltens (vgl. Abschnitt 1.4). Dazu sollen strukturierte *Smart-Meter*-Datensätze mithilfe von Data-Mining-Methoden analysiert werden, um Einsichten auf folgende Fragen zu gewinnen:

- Wie sieht die Reaktion eines Haushaltes oder einer Haushaltsgruppe auf ein bestimmtes Preissignal aus? Inwieweit verändern variable Stromtarife ein bekanntes Verbrauchsverhalten, wie beispielsweise ein Standardlastprofil?
- Gibt es saisonale und wochentagsbedingte Schwankungen für eine *Demand-Response*-Maßnahme, z.B. für Fälle mit ungewöhnlichen Preissignalen? (siehe Aufgabenstellung A1 auf Seite 32)
- Welche prozentualen Energieeinsparungen bzw. zeitliche Lastverschiebungen ergeben sich durch verschiedene Tarife? Inwiefern rufen verschiedene Stromtarife Unterschiede in der Verbrauchsreaktion hervor? Steigern bestimmte Stromtarife die Unsicherheiten und Schwankungen in der Verbrauchsreaktion? (siehe Aufgabenstellung A2 auf Seite 32)
- Reagieren alle Haushalte ähnlich auf bestimmtes Preissignal oder gibt es wesentliche Unterschiede? (siehe Aufgabenstellung A3 auf Seite 32)

Die Ermittlung von preisbeeinflussten Typtagen mithilfe von Clusterverfahren stellt eine entscheidende Komponente im Gesamtkonzept zur datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens dar und wird in Abschnitt 2.5.1 behandelt. Doch zuvor werden im Folgenden alternative Ansätze vorgestellt, die bei der Beantwortung der obigen Fragen helfen.

Die **visuelle Inspektion von Lastzeitreihen** (allgemein: *Visual Analytics*) und anschließende Gruppierung der Lastzeitreihen zu preisbeeinflussten Typtagen ist ein einfacher Ansatz, mit dem relativ schnell eine grobe Abschätzung über das preisbeeinflusste Verbrauchsverhalten gemacht werden kann. Der subjektive Eindruck bei der visuellen Inspektion von „Zeitreihenwolken“ ist jedoch per se fehleranfällig, vor allem wenn der Datensatz Lastzeitreihen von einer großen Anzahl an Haushalten enthält. Eine differenzierte Beantwortung der obigen Fragen ist mit diesem Ansatz nur bedingt möglich. Eine gute Beschreibung von *Visual Analytics* im Allgemeinen liefern [236, 237]. Ausgewählte Anwendungsbeispiele im Energiebereich sind der Einsatz von *Visual Analytics* zur dynamischen *Demand-Response*-Visualisierung [238] oder die Anomalieerkennung in Stromlastdaten mithilfe von *Visual-Analytics*-Methoden [239].

Die **Mittelung von Lastzeitreihen über alle Haushalte einer Tarifgruppe** ist ein Ansatz, um das aggregierte Verbrauchsverhalten ganzer Haushaltsgruppen (vgl. Gleichungen (2.34) und (2.35) in Abschnitt 2.4.4) zu ermitteln. Jedoch macht dieser Ansatz das Vorhandensein einer Kontrollgruppe notwendig, da es ohne Kontrollgruppe nicht möglich ist, den Unterschied zwischen unbeeinflusstem und beeinflusstem Verbrauchsverhalten auszumachen. Handelt es sich bei den untersuchten Haushalten um Haushalte, die individuell auf ein bestimmtes Preissignal reagieren (manuelles *Demand Response*), werden durch die Mittelung über Tarife Responder-, Semi-Responder- und Non-Responder-Haushalte vermengt. Somit kann nicht mehr eindeutig

beantwortet werden, welche wesentlichen Unterschiede es in der Verbrauchsreaktion zwischen unterschiedlichen Haushalten gibt.

Als weitere Möglichkeit zur Ermittlung von preisbeeinflussten Typtagen ist die **Trennung der Lastzeitreihen mithilfe einfacher Heuristiken** denkbar. So zeichnet sich eine Laständerung als Reaktion auf eine sprunghafte Änderung des Strompreises oftmals durch einen großen Lastgradienten aus. Mithilfe einfacher Heuristiken können diese Sprünge als Events detektiert werden und die entsprechenden Lastzeitreihen aus der Vielzahl an Lastzeitreihen im *Smart-Meter*-Datensatz getrennt werden. Dieser Ansatz liefert bei der Suche nach preisbeeinflussten Typtagen aus einer Vielzahl von Lastkurven in der Regel eine suboptimale Lösung, stellt aber einen Kompromiss zwischen Rechenaufwand und Güte der Lösung dar.

Die **Analyse einzelner Abtastzeitpunkte** stellt einen weiteren Ansatz dar, mit dem der Tarifeinflüsse auf die Last ermittelt werden kann und der in Abschnitt 2.5.5 genauer vorgestellt wird. Mit diesem Ansatz lassen sich prozentuale Energieeinsparungen und zeitliche Lastverschiebungen als Reaktion auf variable Strompreise erkennen. Jedoch kann der zeitliche Verlauf des preisbeeinflussten Verbrauchsverhaltens (z.B. „Die Haushalte der XYZ-Gruppe reagieren insofern auf das Preissignal, als es ca. 4 Stunden vor dem Hochpreissignal und ca. 3 Stunden nach dem Hochpreissignal zu einer Erhöhung der aggregierten Last kommt.“) mithilfe dieses Ansatzes nicht ermittelt werden.

Abschnitt 2.5 ist folgendermaßen strukturiert: Abschnitt 2.5.1 befasst sich damit, wie aus der Vielzahl an Lastzeitreihen in *Smart-Meter*-Datensätzen Informationen über typische Verhaltensmuster bei preisbeeinflussten Haushaltsstromkunden gewonnen werden. Die Bewertung einer gefundenen Clusterlösung wird in Abschnitt 2.5.2 behandelt. Die Zuordnung der gefundenen Clusterkurven zu Wochentagen, Monaten und Tarifen stellt einen wichtigen Baustein in der Gesamtpipeline dar und wird in Abschnitt 2.5.3 erläutert. In Abschnitt 2.5.4 werden drei verschiedene Analysen (Parameterstudien) vorgestellt, die die Sensitivität der Clusterergebnisse auf die Änderung bestimmter Einflussgrößen untersuchen sollen. Die Analyse einzelner Abtastzeitpunkte bezüglich der Tarifeinflüsse auf den Stromverbrauch wird in Abschnitt 2.5.5 beschrieben. Abschnitt 2.5 schließt mit einer Zusammenfassung und Diskussion der vorgestellten analyseorientierten Methoden der Pipeline.

2.5.1 Clustering zur Ermittlung von preisbeeinflussten Typtagen

Wie bereits in Abschnitt 1.2.7 beschrieben, wird der Lastgang von Haushaltsstromkunden heutzutage durch Standardlastprofile geschätzt. Der aggregierte Lastgang einer genügend großen Anzahl von Haushalten (mindestens 150 Haushalte [240]) entspricht in erster Näherung dem Standardlastprofil. Allerdings werden individuelle Tageslastgänge und kleinere Subgruppen von Haushalten durch das Standardlastprofil nur unzureichend repräsentiert. Werden Haushalte zudem durch variable Strompreise beeinflusst, versagen die Standardlastprofile, um das Verbrauchsverhalten zu schätzen. Aus diesem Grund müssen datengetriebene Verfahren eingesetzt werden, die exaktere Aussagen über das Verbrauchsverhalten von Haushaltsstromkunden zulassen. Deshalb wird im Folgenden ein datengetriebener Ansatz vorgestellt, der mithilfe von Clusterverfahren preisbeeinflusste Typtage identifiziert.

Ein entscheidender Vorteil beim Einsatz von Clusterverfahren für diese Aufgabenstellung liegt darin, dass in der Entwurfsphase des Verfahrens keinerlei Informationen über Wochentage, Jahreszeiten, Tarife und Haushaltsarten notwendig sind. Da die Lastzeitreihen aus *Smart-Meter*-Datensätzen eine hohe Variabilität und Überlagerungen untereinander aufweisen, ist davon auszugehen, dass keine scharfen und gut voneinander abgrenzbaren Cluster gefunden werden können. Die fehlende Eindeutigkeit bei der Zuordnung einer Lastzeitreihe zu einem Cluster erfordert die Verwendung von sogenannten unscharfen (englisch: *fuzzy*) Clustermethoden [85, 241]. Aus der Vielzahl von Clusterverfahren wird deshalb das Fuzzy-C-Means-Verfahren (FCM-Clustering) verwendet [178, 242, 243].

Im Rahmen der Arbeit wurde die Clusteranalyse an den preisbeeinflussten Lastzeitreihen der sechs Fälle DH , DH_{norm} , DH_{normmean} , DT , DT_{norm} und DT_{normmean} aus Tabelle 2.3 durchgeführt. Zudem wurde, um den Einfluss von unterschiedlichen Tagesroutinen an Werktagen und Wochenendtagen zu unterdrücken, die Clusteranalyse an den preisbeeinflussten Lastzeitreihen der sechs Fälle DH^{wo} , $DH_{\text{norm}}^{\text{wo}}$, $DH_{\text{normmean}}^{\text{wo}}$, DT^{wo} , $DT_{\text{norm}}^{\text{wo}}$ und $DT_{\text{normmean}}^{\text{wo}}$ durchgeführt. Der Clustering-Algorithmus versucht im Datensatz

$$\mathbf{X}_{K \times N_D} = \begin{pmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_{N_D} \end{pmatrix} = \begin{pmatrix} x_1[1] & \cdots & x_{N_D}[1] \\ \vdots & \ddots & \vdots \\ x_1[K] & \cdots & x_{N_D}[K] \end{pmatrix} \quad (2.36)$$

Clusterzentren

$$\mathbf{V}_{K \times C} = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_C \end{pmatrix} = \begin{pmatrix} v_1[1] & \cdots & v_C[1] \\ \vdots & \ddots & \vdots \\ v_1[K] & \cdots & v_C[K] \end{pmatrix} \quad (2.37)$$

durch Lösen des Optimierungsproblems

$$\min_{\boldsymbol{\mu}, \mathbf{V}} \left\{ J_q(\boldsymbol{\mu}, \mathbf{V}) = \sum_{n_D=1}^{N_D} \sum_{c=1}^C (\mu_{n_D,c})^q \cdot d_{\text{Euk},c}^2(\mathbf{x}_{n_D}, \mathbf{v}_c) \right\} \quad (2.38)$$

$$\text{u.d.N. } \mu_{n_D,c} \in [0, 1], \forall n_D = 1 \dots N_D, \forall c = 1 \dots C \quad (2.39)$$

$$\text{und } \sum_{c=1}^C \mu_{n_D,c} = 1, \forall n_D = 1 \dots N_D \quad (2.40)$$

$$\text{und } \sum_{n=1}^{N_D} \mu_{n,c} > 0, \forall c = 1 \dots C \quad (2.41)$$

zu finden. K ist dabei die Länge der Lastzeitreihe, N_D ist die Anzahl der Datentupel, C ist die Anzahl der Clusterzentren, q ist der Fuzzifizierungsgrad, $\mu_{n_D,c}$ ist die Zugehörigkeit eines Daten-

tupels $n_D = 1, \dots, N_D$ zu einem Clusterzentrum $c = 1, \dots, C$. $d_{Euk,c}(\mathbf{x}_{n_D}, \mathbf{v}_c)$ ist der euklidische Abstand der Punkte der Lastzeitreihe \mathbf{x}_{n_D} und des Clusterzentrums \mathbf{v}_c :

$$d_{Euk,c}(\mathbf{x}_{n_D}, \mathbf{v}_c) = \sqrt{\sum_{k=1}^K (x_{n_D}[k] - v_c[k])^2}. \quad (2.42)$$

Bei der Verwendung alternativer Distanzmaße in Fuzzy-Clusterverfahren sei auf [85, 244] verwiesen.

Anschließend folgt die iterative Berechnung der Cluster-Zugehörigkeiten

$$\mu_{n_D,c} = \frac{\left[d_{Euk,c}^2(\mathbf{x}_{n_D}, \mathbf{v}_c) \right]^{\frac{1}{1-q}}}{\sum_{i=1}^C \left[d_{Euk,i}^2(\mathbf{x}_{n_D}, \mathbf{v}_i) \right]^{\frac{1}{1-q}}}, \quad (2.43)$$

um neue Clusterzentren

$$\mathbf{v}_c = \frac{\sum_{n_D=1}^{N_D} (\mu_{n_D,c})^q \mathbf{x}_{n_D}}{\sum_{n_D=1}^{N_D} (\mu_{n_D,c})^q} \quad (2.44)$$

zu erhalten. Der Fuzzifizierungsgrad $q \in (1, \infty)$ sorgt für „harte“ ($q \rightarrow 1$) oder weichere Zugehörigkeiten $q \rightarrow \infty$. Die Literatur empfiehlt oftmals einen Fuzzifizierungsgrad von $q = 2$ [243], allerdings führt das zu nahezu gleichverteilten Zugehörigkeiten. In [76, 245] wird empfohlen, Werte zwischen $q = 1.05$ und 1.30 zu testen. Die divergierenden Aussagen bezüglich der Wahl des Fuzzifizierungsgrades werden in Abschnitt 2.5.4.1 in einer umfangreichen Sensitivitätsanalyse untersucht.

Die Datenvorverarbeitung (vgl. Abschnitt 2.4.1) sollte im Idealfall alle Ausreißer (allg.: Stördaten) in den *Smart-Meter*-Datensätzen erkennen und geeignete Ersatzwerte liefern. Da dies allerdings nicht garantiert werden kann, wird das Fuzzy-C-Means-Verfahren um das sogenannte Rausch-Clustering (englisch: *noise clustering*) ergänzt [246–248]. Die Idee ist, die Ausreißer in einem zusätzlichen Rauschcluster mit dem Clusterzentrum \mathbf{v}_{oc} (Index „oc“ - *outlier cluster*) und einem konstanten Rauschabstand

$$d_{Euk,oc}^2(\mathbf{x}_{n_D}, \mathbf{v}_{oc}) = \delta \quad (2.45)$$

zu allen Datentupeln \mathbf{x}_{n_D} zu gruppieren. Eine geeignete Schätzung des Rauschabstandes δ erfordert die visuelle Inspektion der Datentupel in den Clustern, welche für jeden Datensatz durchgeführt werden sollte. Grundsätzlich führen kleine Werte für δ zu falsch klassifizierten Datentupeln, da viele Datentupel als Rauschen dem Rausch-Cluster zugeordnet und somit „verschont“ werden. Auf der anderen Seite führt ein zu großer Rauschabstand dazu, dass Ausreißer nicht dem Rausch-Cluster, sondern einem „normalen“ Cluster zugewiesen wird, was fehlerhafte Cluster begünstigt. In [247, 249, 250] sind alternative Vorgehensweisen zur Schätzung eines sinnvollen Rauschabstandes δ beschrieben. Da der Rauschabstand δ ganz entscheidend die Wirksamkeit des vorgestellten Ansatzes beeinflusst, werden sinnvolle Werte des Rauschabstandes in Abschnitt 2.5.4.1 in einer umfangreichen Sensitivitätsanalyse ermittelt.

Es muss an dieser Stelle angemerkt werden, dass Lastzeitreihen, die durch seltene und ungewöhnliche Preissignale (z.B. *Critical Peak Pricing*) beeinflusst wurden, u.U. im Rausch-Cluster verschwinden. Wenn es darum geht, das *typische* preisbeeinflusste Verbrauchsverhalten zu bestimmen, ist diese Zuordnung zum Rausch-Cluster zulässig. Geht es jedoch darum, die Auswirkungen von seltenen und ungewöhnlichen Preissignalen aufzudecken, müssen die entsprechenden Lastzeitreihen aus dem Rausch-Cluster extrahiert werden.

Zur Verdeutlichung des eben vorgestellten Clusterverfahrens zur Ermittlung von Typtagen ist in Beispiel 6 das Ergebnis der Fuzzy-C-Means-Clusteranalyse beispielhaft gezeigt.

Beispiel 6: Clustering am Beispiel eines Benchmarkdatensatzes

Das in Abschnitt 2.5.1 vorgestellte Clusterverfahren zur Ermittlung von preisbeeinflussten Typtagen wird nun an einem Beispiel verdeutlicht. Als Datenbasis dient der Fall DH des Benchmarkdatensatzes „Settlement210H365D_80M10S“, der eigentlich aus $N_D = 76650$ Datentupeln ($N_D = 74674$ Datentupel nach der Datenvorverarbeitung) besteht, aber aus Gründen der Übersichtlichkeit auf 1% der Datentupel reduziert wurde. In Abb. 2.8 ist das Ergebnis des Fuzzy-Clusterings mit einer festgelegten Clusteranzahl von $C = 5$ dargestellt. Es sind klar voneinander unterscheidbare Cluster zu erkennen, was auch die Zugehörigkeitsfunktionen der einzelnen Datentupel in Abb. 2.10 belegen. Zudem sind in Abb. 2.9 die mittleren Clusterkurven der fünf Cluster dargestellt.

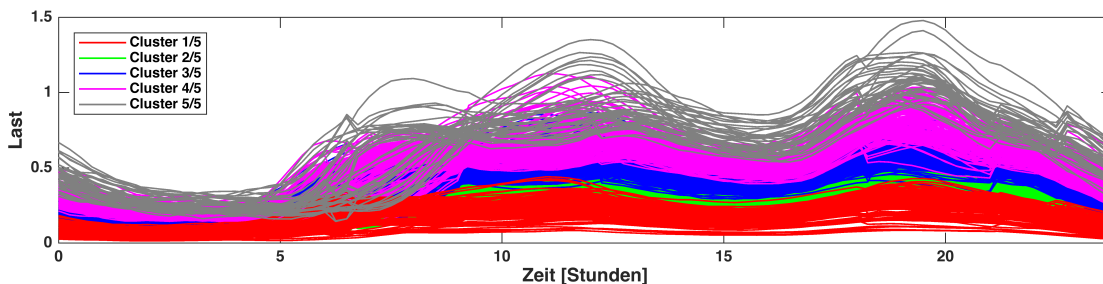


Abbildung 2.8: Zuordnung der Tageskurven zu den Clustern: geclustert wurden die Zeitreihen nach der Datenvorverarbeitung

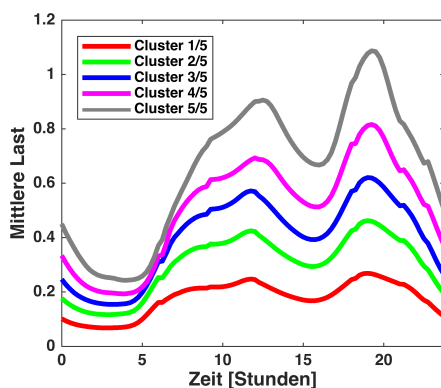


Abbildung 2.9: Mittlere Clusterkurven

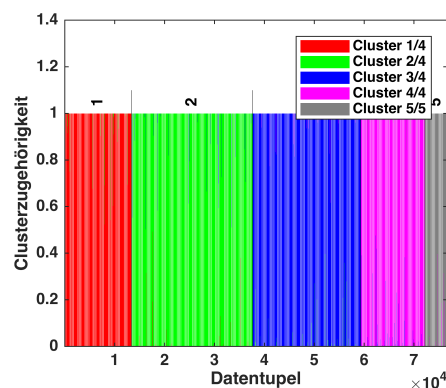


Abbildung 2.10: Clusterzugehörigkeiten der Datentupel zu den Clustern

2.5.1.1 Ablauf des Fuzzy-C-Means-Clusteralgorithmus

Der Fuzzy-C-Means-Clusteralgorithmus beinhaltet folgende Schritte, wobei i der Iterationsschritt ist:

- **Schritt 1:** Festlegung der Clusteranzahl C , des Fuzzifizierungsgrades q , des Rauschabstandes δ , der maximalen Anzahl an Iterationen i_{max} und der Konvergenzschwelle ϵ
- **Schritt 2:** $i = 0$: Initiale Festlegung der Startcluster $\mathbf{v}_c^{(0)}$
- **Schritt 3:** Berechnung der Fuzzy-Zugehörigkeiten $\mu_{n_D,c}^{(0)}$ für $\mathbf{v}_c^{(0)}$ gemäß Gleichung (2.43)
- **Schritt 4:** Berechnung neuer Clusterzentren $\mathbf{v}_c^{(i)}$ gemäß Gleichung (2.44)
- **Schritt 5:** Berechnung der Fuzzy-Zugehörigkeiten $\mu_{n_D,c}^{(i)}$ für $\mathbf{v}_c^{(i)}$ gemäß Gleichung (2.43)
- **Schritt 6:** Die Fuzzy-C-Means-Zielfunktion J_q erreicht ein lokales Minimum, wenn

$$\left\| \boldsymbol{\mu}_{n_D,c}^{(i)} - \boldsymbol{\mu}_{n_D,c}^{(i-1)} \right\| < \epsilon \quad (2.46)$$

gilt. Falls Gleichung (2.46) nicht erfüllt ist, wird i um Eins erhöht und die Schritte 4 bis 6 werden wiederholt.

2.5.1.2 Wahl der Startcluster

Die Konvergenzeigenschaften des Fuzzy-Clusteralgorithmus lassen zu, dass die Startcluster zufällig ausgewählt werden können [85]. Es kann allerdings nicht garantiert werden, dass das globale Minimum bei der Optimierung gefunden wird. Aus diesem Grund erscheint es sinnvoll, unterschiedliche Strategien zur Wahl der Startcluster auszuprobieren und der Clusterergebnisse anschließend zu vergleichen [244]. Im Rahmen der Arbeit werden folgende Strategien zur Bestimmung der initialen Clusterzentren getestet (siehe Abschnitt 2.5.4.3):

- **Strategie 1: zufällige Auswahl der Startcluster**
Die Startcluster werden zufällig im Bereich zwischen der unteren ($\min(\mathbf{P}_{d,n})$) und der oberen ($\max(\mathbf{P}_{d,n})$) Wertebereichsgrenzen aller Tageskurven ausgewählt.
- **Strategie 2: Auswahl gleichverteilter Startcluster**
Die Startcluster werden gleichverteilt im Bereich zwischen der unteren ($\min(\mathbf{P}_{d,n})$) und der oberen ($\max(\mathbf{P}_{d,n})$) Wertebereichsgrenzen aller Tageszeitreihen ausgewählt.
- **Strategie 3: Auswahl der Startcluster aus Fuzzy-Menge**
Die Wahl der Startcluster erfolgt gemäß folgender Berechnung: Es werden zunächst die Tagesmittelwerte $\bar{P}_{d,n}$ und Wertebereiche $P_{d,n,W}$ aller Tageskurven $P_{d,n}[k]$ berechnet. Den Einzelmerkmalen $\bar{P}_{d,n}$ und $P_{d,n,W}$ wird eine graduelle Zuordnung zu einem von fünf linguistischen Termen zugewiesen, was in Abb. 2.11 dargestellt ist. Die linguistischen Terme fungieren dabei als Ausgangsklassen. Anschließend werden die fünf Ausgangsklassen je Merkmal $\bar{P}_{d,n}$ und $P_{d,n,W}$ verknüpft (UND-Verknüpfung) und so im Idealfall 25 neue

Ausgangsklassen generiert. Es wird nun je Ausgangsklasse zufällig ein Datentupel ausgewählt, welches als Startcluster fungiert. In Beispiel 7 ist diese Strategie exemplarisch an einem Benchmarkdatensatz gezeigt.

- **Strategie 4: Aufteilung der Startcluster auf Wochentage**

Aus der Gesamtheit aller Datentupel werden manuell solche Datentupel als Startcluster ausgewählt, dass bei einer Clusteranzahl von $C \geq 7$ jeder Wochentag mindestens einmal als Startclustertageskurve vertreten ist. Diese Strategie verhindert jedoch nicht, dass ungünstige Startcluster ausgewählt werden.

Beispiel 7: Auswahl der Startcluster aus Fuzzy-Menge gemäß Strategie 3 am Beispiel eines Benchmarkdatensatzes

Als Datenbasis dient der Fall DH des Benchmarkdatensatzes „Settlement210H365D_60M30S“, der eigentlich aus $N_D = 76650$ Datentupeln besteht, aber aus Gründen der Übersichtlichkeit auf 1% der Datentupel reduziert wurde.

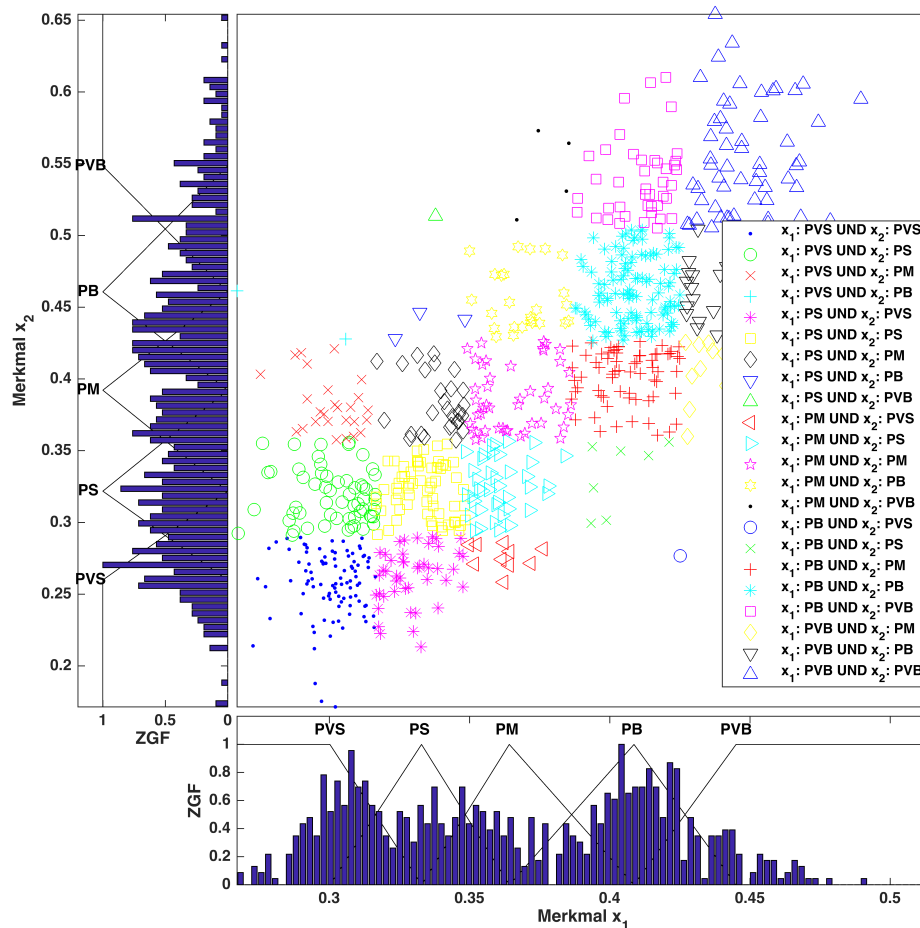


Abbildung 2.11: Wahl der Startcluster gemäß Strategie 3 auf Seite 66: Zugehörigkeitsfunktionen (ZGF) und Histogramme des Merkmals x_1 (Tagesmittelwerte $\bar{P}_{d,n}$, unten), des Merkmals x_2 (Wertebereiche $P_{d,n,w}$, links), und graduelle Zuordnung der Datentupel zu den vereinigten Klassen (mitte); PVS - Positive Very Small, PS - Positive Small, PM - Positive Mean, PB - Positive Big, PVB - Positive Very Big.

2.5.2 Bewertung der Clusterlösung

In Abschnitt 2.5.2 wird die Bewertung einer gefundenen Clusterlösung behandelt. In der Literatur wird vorgeschlagen, eine Clusterlösung anhand von Visualisierungen und Gütemaßen zu beurteilen [85, 178]. Die Visualisierung einer Clusterlösung, wie in Beispiel 6 exemplarisch gezeigt, kann helfen, eine qualitative Bewertung der gefundenen Lösung (etwa: „Es sind vier klar voneinander unterscheidbare Cluster zu erkennen.“) vorzunehmen. Jedoch eignet sich dieses Verfahren nicht, um aus einer Vielzahl an Clusterlösungen eine optimale Clusterlösung auszuwählen, da es sehr zeitaufwendig und zudem durch die subjektive Wahrnehmung fehlerbehaftet ist.

Bei den Gütemaßen wird zwischen externen und internen Gütemaßen unterschieden. Externe Gütemaße werden dann eingesetzt, wenn zwei Klasseneinteilungen verglichen werden, wobei eine Klasseneinteilung die „korrekte“ Klasseneinteilung im Sinne von gelabelten Klassen und die andere Klasseneinteilung die Clustereinteilung ist. Die externen Gütemaße machen das Vorhandensein einer sogenannten *ground truth* (künstliche Daten mit bekannten Eigenschaften und Klassen) notwendig, bei der die Struktur in den Daten bzw. die Klasseneinteilung bekannt ist. Der Vorteil bei der Verwendung von externen Gütemaßen liegt darin, Klassen in Datensätzen wiederzuentdecken und so eine quantitative Bewertung der Clustereinteilung (etwa: „Es werden 90% der Datentupel richtig zugeordnet.“) machen zu können, was einer Klassifikation ähnelt. In zahlreichen Veröffentlichungen wurden die externen Gütemaße *Cluster Similarity Measure*, *Jaccard-Index*, *Fowlkes–Mallows-Index* oder *F-Measure* verwendet [251–255]. Die Bewertung einer Clustereinteilung auf Basis von externen Gütemaßen erweist sich dann als problematisch, wenn das gelernte Modell auf reale Daten angewendet wird, die sich von denen in der *ground truth* unterscheiden. Neue, zuvor unbekannte Strukturen in den realen Daten werden bestraft, da sie keiner gelabelten Klasse entsprechen.

Interne Gütemaße beurteilen eine Clustereinteilung nur auf Basis der Daten und ohne Einbeziehung externer Informationen. Eine Bewertung auf Basis interner Gütemaße wird dann vorgenommen, wenn keine *ground truth* vorhanden ist und wenn verschiedene Clusterlösungen qualitativ miteinander verglichen werden sollen. Eine gute Clustereinteilung zeichnet sich durch eine große Ähnlichkeit der Datentupel innerhalb der Cluster (Kohäsion) und durch eine geringe Ähnlichkeit der Cluster untereinander (Separierung) aus. Die internen Gütemaße werden mit dem Ziel eingesetzt, Aussagen über die Kohäsion und/oder Separierung der Cluster zu treffen. Als prominente Beispiele für interne Gütemaße sind der *Calinski-Harabasz-Index* [256], der *Silhouette-Index* [257], der *Dunn-Index* [258], der *S_DbW-Index* [259], der *Xie-Beni-Index*¹³ [260] und der *Davies-Bouldin-Index* [261] zu nennen. Ein Vergleich verschiedener Gütemaße wurde in [262–266] vorgenommen. Eine mit internen Gütemaßen gefundene optimale Clusterlösung sollte in jedem Fall auf deren Sinnhaftigkeit geprüft werden, da Clusterverfahren Strukturen in den Daten erkennen, egal ob tatsächlich Strukturen existieren oder nicht. Eine umfangreiche Übersicht über externe und interne Clusterbewertungsmaße liefern [178, 265, 267–270].

¹³In [260] wird das Gütemaß als *compactness and separation validity index S* bezeichnet. Eine alternative deutsche Bezeichnung für *S* ist der Trennungsgrad.

Wie bereits in Abschnitt 2.5.1 beschrieben, ist das Auffinden von Gruppen ähnlicher Lastzeitreihen in *Smart-Meter*-Datensätzen das Ziel und die Clusterkurven das Ergebnis des Clusterings. Zur Bewertung der berechneten Clustereinteilung, und somit der gefundenen Clusterzentren, wurde in der vorliegenden Arbeit der *Xie-Beni*-Index verwendet, der im Folgenden als Trennungsgrad S bezeichnet wird. S wird gemäß

$$S(\boldsymbol{\mu}, C) = \frac{J_q(\boldsymbol{\mu}, \mathbf{V})}{C \cdot \min_{c,i=1\dots C, c \neq i} (d^2(\mathbf{v}_c, \mathbf{v}_i))} \quad (2.47)$$

berechnet. Dabei gibt S das Verhältnis zwischen dem mittleren Abstand der Datentupel zu den ihnen zugeordneten Clustern (Kompaktheit) und dem kleinsten Abstand zweier Clusterzentren (Separation) an. Je kleiner S , desto besser ist die Clustereinteilung. Da Voruntersuchungen [76, 209] mit konstantem Fuzzifizierungsgrad und Rauschabstand und aufsteigender Clusteranzahl ($C = 3, \dots, \bar{C}$) gezeigt haben, dass der geringste Wert für S oftmals bei $C = 3$ liegt, wird hier das erste lokale Minimum für $S(\boldsymbol{\mu}, C)$ mit $C > 3$ als optimale Clusteranzahl C_{opt} gewählt.

2.5.3 Clusterzuordnungen

Die in Abschnitt 2.5.1 gefundenen Clusterzentren lassen sich als preisbeeinflusste Typtage (alternative Bezeichnung: typische preisbeeinflusste Lastprofile, TLP) interpretieren, die preisinduzierte und saisonale Effekte auf das Verbrauchsverhalten von Haushaltsstromkunden erkennen lassen. Allerdings lässt die Clusteranalyse keine Aussagen darüber zu, wie stark die Zusammenhänge zwischen den preisbeeinflussten Typtagen und den Strompreistarifen sind. Auch zur Beantwortung der Frage, ob es saisonale oder wochentagsbedingte Schwankungen für eine *Demand-Response*-Maßnahme gibt, müssen geeignete automatisierte Clusterbewertungsverfahren eingesetzt werden, was im Folgenden vorgestellt wird.

2.5.3.1 Bewertung der Zusammenhänge von Clustern und Werktagen, Monaten und Tarifen mit Verbundverteilungen

Eine manuelle Analyse der Zusammenhänge zwischen den Clustern und Werktagen, Monaten oder Tarifen ist sehr aufwändig und schwer quantifizierbar. Grundsätzlich existieren zahlreiche Bewertungsmaße, um die Abhängigkeit bei kategorialen Variablen zu schätzen, wie z.B. der χ^2 -Unabhängigkeitstest. In der vorliegenden Doktorarbeit wurden für die Analyse des Zusammenhangs zwischen Clustern und anderen Ausgangsgrößen sowohl Tabellen für die Verbundverteilungen als auch informationstheoretische Maße eingesetzt.

Es sei mit y_1 die Ausgangsgröße „Cluster“ mit den Clusterzugehörigkeiten μ_c ($c = 1, \dots, C$) als Klassen der Ausgangsgröße gegeben. Zudem sei y_2 eine zweite Ausgangsgröße (z.B. „Werktag“, „Monat“ oder „Tarif“) mit den Klassen $b = 1, \dots, m_y$. Die **Verbundverteilung** zwischen y_1 und y_2 berechnet sich folgendermaßen:

$$P(y_1 = c \cap y_2 = b) = \frac{N_D(y_1 = c \cap y_2 = b)}{N_D}, \quad (2.48)$$

wobei $N_D(y_1 = c \cap y_2 = b)$ die Anzahl der Datentupel für das Auftreten des Ereignisses ($y_1 = c \cap y_2 = b$) und N_D die Gesamtanzahl der Datentupel ist. Wie in Abb. 2.15 gezeigt, kann die Verbundverteilung entweder als Tabelle oder Balkendiagramm dargestellt werden.

2.5.3.2 Bewertung der Zusammenhänge von Clustern und Werkzeugen, Monaten oder Tarifen mit informationstheoretischen Maßen

Zur Beurteilung der Abhängigkeit zweier wertediskreter Merkmale bzw. Ausgangsgrößen eignet sich auch die **Transinformation** (englisch: *mutual information*) bzw. die relative Transinformation, welche aus dem ID3-Verfahren bei Entscheidungsbäumen bekannt ist [271, 272]. Beide Maße gehen auf den informationstheoretischen Begriff der Entropie zurück, welcher von Shannon [273] begründet wurde. Die Entropie H (auch Unsicherheit oder mittlerer Informationsgehalt genannt) einer Ausgangsgröße y mit den Klassen $b = 1, \dots, m_y$ kann grundsätzlich mithilfe von

$$H(y) = - \sum_{b=1}^{m_y} P(y = b) \cdot \log_2 P(y = b) \quad (2.49)$$

berechnet werden, wobei $P(y = b)$ die Auftrittswahrscheinlichkeit der Merkmalsausprägung $y = b$ ist. $-\log_2 P(y = b)$ ist der Informationsgehalt der Merkmalsausprägung $y = b$, welcher von der Auftrittswahrscheinlichkeit abhängt. Bei einer geringen Auftrittswahrscheinlichkeit ist der Informationsgehalt hoch; tritt eine Merkmalsausprägung $y = b$ häufig auf, ist deren Informationsgehalt gering (siehe Abb. 2.12a). Die Entropie erreicht ihr Maximum bei gleich-

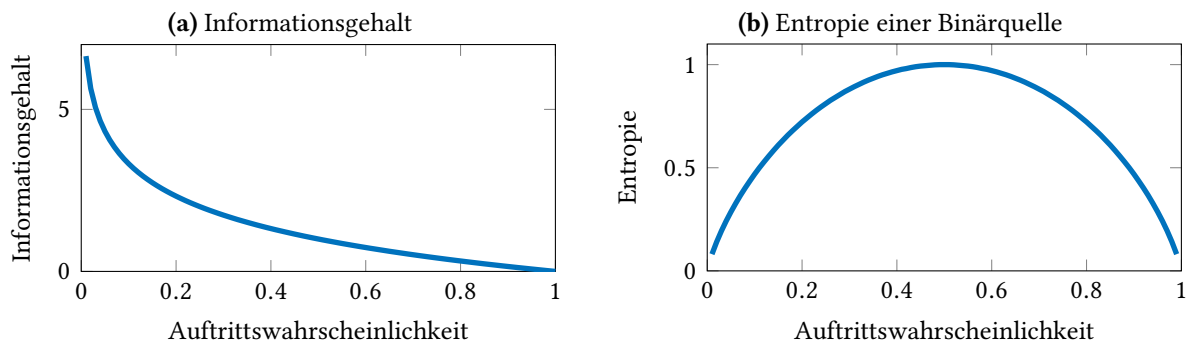


Abbildung 2.12: (a) Informationsgehalt einer Merkmalsausprägung $y = b$ in Abhängigkeit von der Auftrittswahrscheinlichkeit $P(y = b)$ und (b) Entropie einer Binärquelle (modifiziert nach [272])

wahrscheinlichen Merkmalsausprägungen (siehe Beispiel der Entropie einer Binärquelle in Abb. 2.12b).

Um nun die Abhängigkeit zweier Ausgangsgrößen y_1 und y_2 zu bestimmen, wird eine der Ausgangsgröße als Informationsquelle mit der Eingangsentropie $H(y_1)$ und die andere Ausgangsgröße als Informationssenke mit der Ausgangsentropie $H(y_2)$ interpretiert, was in Abb. 2.13a und Abb. 2.13b dargestellt ist. Die Eingangs-, Ausgangs- und Verbundentropie werden analog

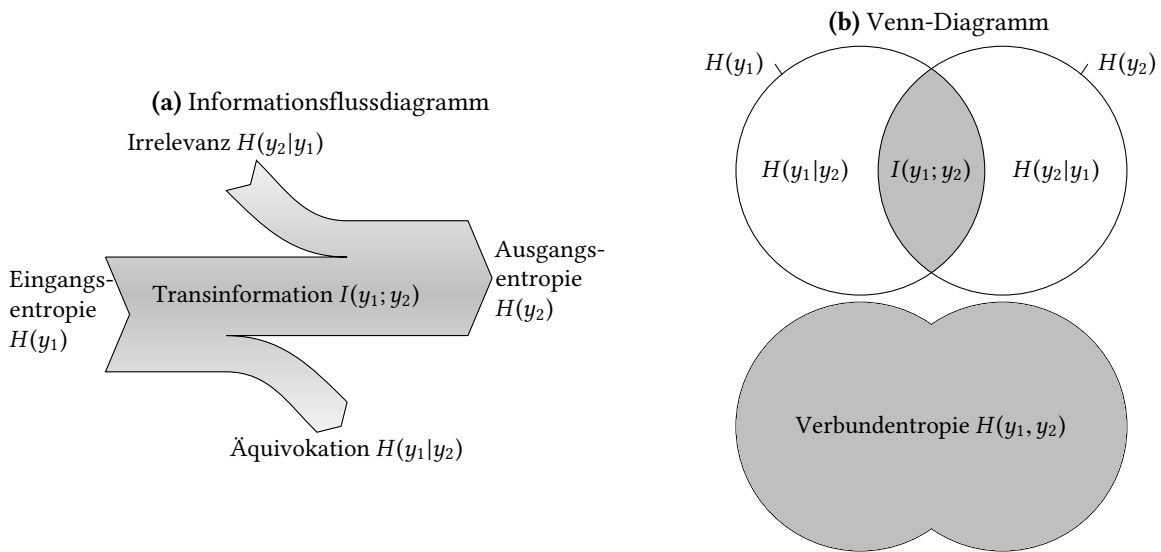


Abbildung 2.13: (a) Informationsflussdiagramm und (b) Venn-Diagramm zur Darstellung der Abhängigkeit von Eingangsentropie $H(y_1)$, Ausgangsentropie $H(y_2)$, Verbundentropie $H(y_1, y_2)$, Transinformation $I(y_1; y_2)$, Irrelevanz $H(y_2|y_1)$ und Äquivokation $H(y_1|y_2)$

zu Gleichung (2.49) folgendermaßen berechnet:

$$H(y_1) = - \sum_{c=1}^C P(y_1 = c) \cdot \log_2 P(y_1 = c), \quad (2.50)$$

$$H(y_2) = - \sum_{l=1}^{m_y} P(y_2 = l) \cdot \log_2 P(y_2 = l), \quad (2.51)$$

$$H(y_1, y_2) = - \sum_{c=1}^C \sum_{l=1}^{m_y} P(y_1 = c \cap y_2 = l) \cdot \log_2 P(y_1 = c \cap y_2 = l). \quad (2.52)$$

Da es in Abschnitt 2.5.3 konkret um die Abhängigkeit zwischen den gefundenen Clustern und weiteren Ausgangsgrößen (wie Werktag, Monat und Tarif) geht, wird in den Gleichungen (2.50) und (2.52) als Eingang y_1 immer das Cluster mit der maximalen Clusterzugehörigkeit $y_1 = c = \operatorname{argmax} \mu_c$ verwendet. Als Ausgang y_2 wird in den Gleichungen (2.51) und (2.52) in drei separaten Untersuchungen die Ausgangsgröße „Monat“ ($m_y = 12$: *Jan, Feb, ..., Dez*), die Ausgangsgröße „Werktag“ ($m_y = 2$: *ja* oder *nein*) und die Ausgangsgröße „Tarif“ (m_y abhängig vom Datensatz) eingesetzt.

Der Informationsfluss von der Quelle (Eingangsentropie) zur Senke (Ausgangsentropie) ist durch die Transinformation $I(y_1; y_2)$ beschrieben. Für den Zusammenhang von Transinformation und Eingangs- bzw. Ausgangsentropie gilt Folgendes:

$$I(y_1; y_2) = H(y_1) - H(y_1|y_2) = H(y_2) - H(y_2|y_1), \quad (2.53)$$

$$= H(y_1) + H(y_2) - H(y_1, y_2). \quad (2.54)$$

Die Äquivokation oder Verlustentropie $H(y_1|y_2)$ ist der Teil der Information, der beim Informationsfluss von Quelle zur Senke verloren geht. Die Irrelevanz oder Fehlinformation $H(y_2|y_1)$

ist der Teil der Information, der von der Senke empfangen wird, aber nicht von der Quelle gesendet wurde. Mithilfe der Transinformation kann nun die Abhängigkeit zweier Ausgangsgrößen bewertet werden: Die Transinformation hat den Wert Null, wenn y_1 und y_2 unabhängig sind und erreicht den maximalen Wert, wenn sich eine Ausgangsgröße vollkommen aus der anderen berechnen lässt. Zur besseren Vergleichbarkeit zwischen Datensätzen ist es ratsam, die Transinformation durch Division mit der Eingangs- oder Ausgangsentropie zu normieren. Die relative Transinformation pro Eingangsentropie Q_E bzw. Ausgangsentropie Q_A berechnet sich wie folgt:

$$Q_E = \frac{I(y_1; y_2)}{H(y_1)}, \quad Q_E \in [0, 1] \text{ und} \quad (2.55)$$

$$Q_A = \frac{I(y_1; y_2)}{H(y_2)}, \quad Q_A \in [0, 1]. \quad (2.56)$$

Die relative Transinformation ergibt somit ein relatives Maß zwischen 0 (keine Abhängigkeit) und 1 (aus den Clustern folgt eine eindeutige Zuordnung zu den Klassen der Ausgangsgröße), mit dem sich die jeweilige Ausgangsgröße durch die Cluster erklären lässt. Mithilfe von Q_E bzw. Q_A können unterschiedliche Fälle und Ausgangsgrößen miteinander quantitativ verglichen werden. Die Bewertung der Abhängigkeit zweier Ausgangsgrößen ist exemplarisch in Beispiel 8 gezeigt.

Beispiel 8: Bewertung der Abhängigkeit zweier Ausgangsgrößen

In diesem Beispiel bildet der Benchmarkdatensatz „Settlement210H365D_70M20S“ die Datenbasis, wobei der Datensatz auf 10% der Datentupel mit zufälliger Auswahl der Datentupel reduziert wurde. Nach erfolgtem Clustering des DH-Falls mit den Clusterparametern $C = 6$, $\delta = 1.5$ und $q = 1.05$ ergeben sich die in Abb. 2.14 gezeigten Clusterkurven.

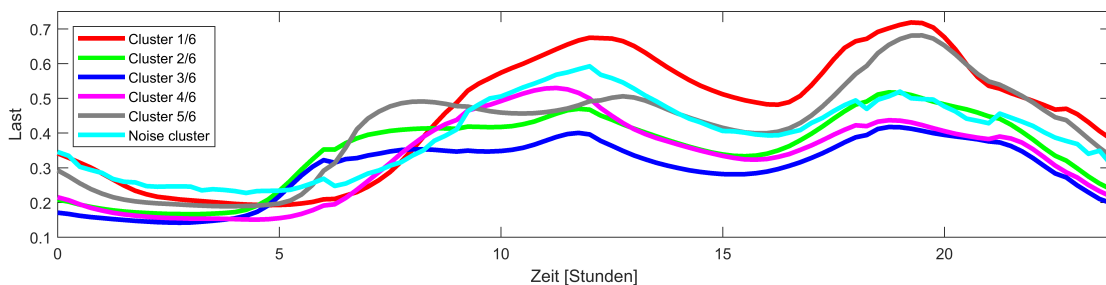


Abbildung 2.14: Mittlere Clusterkurven der sechs Cluster

Exemplarisch wird die Abhängigkeit der Ausgangsgrößen „Cluster 6“ und „Wochentag“ (Abb. 2.16) bzw. „Cluster 6“ und „Monat“ (Abb. 2.15) mithilfe von Verbundverteilungen und informationstheoretischen Maßen bewertet. Dazu werden die relativen Häufigkeiten der beiden Ausgangsgrößen und deren Verbundverteilung berechnet und als Balken in einem Balkendiagramm dargestellt. Die entsprechenden Entropie- und Transinformationswerte für das gezeigte Beispiel stehen auf der rechten Seite der Abb. 2.15 und 2.16.

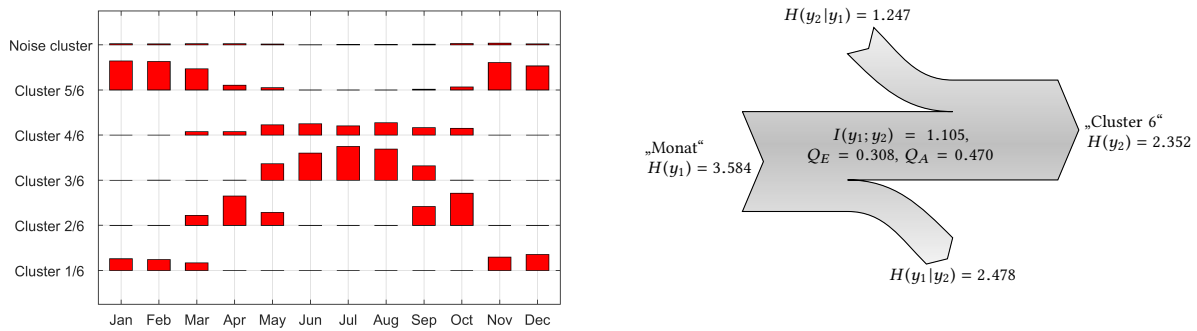


Abbildung 2.15: Ausgangsgrößen „Cluster 6“ und „Monat“: (links) Verbundverteilung der Ausgangsgrößen als Balkendiagramm, (rechts) Informationsflussdiagramm mit Entropie- und Transinformationswerten

Der relativ hohe Wert für Q_A für die Ausgangsgrößen „Monat“ und „Cluster 6“ bedeutet, dass aus den Klassen der Ausgangsgröße „Monat“ gut auf die Klassen der Ausgangsgröße „Cluster 6“ geschlossen werden kann. So ist beispielsweise Cluster 3/6 ein deutliches Sommercluster, wohingegen Cluster 5/6 aus Lastzeitreihen besteht, die in den Monaten November, Dezember, Januar und Februar aufgezeichnet wurden. Der kleinere Wert für Q_E im Vergleich zu Q_A bedeutet, dass in diesem Beispiel besser von den Klassen der Ausgangsgröße „Monat“ auf die Klassen der Ausgangsgröße „Cluster 6“ geschlossen werden kann als umgekehrt.

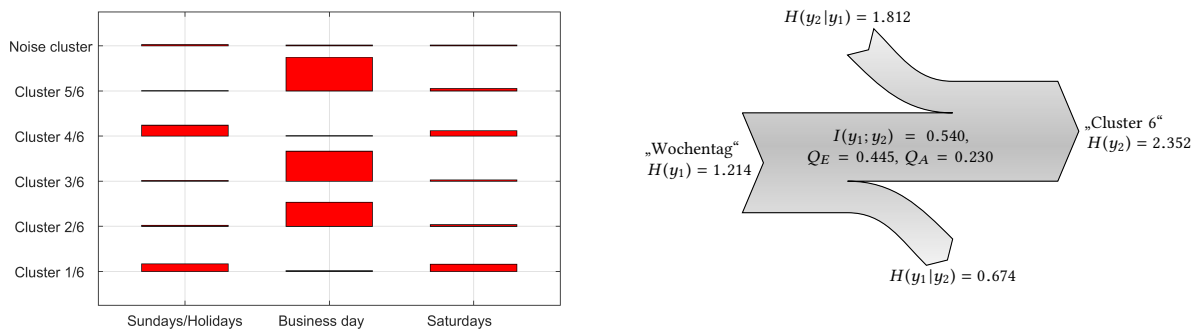


Abbildung 2.16: Ausgangsgrößen „Cluster 6“ und „Wochentag“: (links) Verbundverteilung der Ausgangsgrößen als Balkendiagramm, (rechts) Informationsflussdiagramm mit Entropie- und Transinformationswerten

Der Wert für Q_A für die Ausgangsgrößen „Wochentag“ und „Cluster 6“ liegt deutlich darunter. Hier fällt es schwerer, von den Klassen der Ausgangsgröße „Wochentag“ auf die Cluster zu schließen. Jedoch kann von den Klassen der Ausgangsgröße „Cluster 6“ gut auf die Klassen der Ausgangsgröße „Wochentag“ geschlossen werden. Der Wert für Q_E liegt hier über dem Wert für Q_A .

Die in den Abschnitten 2.5.1 und 2.5.3 vorgestellten Methoden helfen bei der Beantwortung der Fragen auf Seite 61. Dies soll nun am Beispiel des ersten Fragenblocks erläutert werden: „Wie sieht die Reaktion eines Haushaltes oder einer Haushaltsgruppe auf ein bestimmtes Preissignal aus? Inwieweit verändern variable Stromtarife ein bekanntes Verbrauchsverhalten, wie beispielsweise ein Standardlastprofil?“. Zur Beantwortung dieser Fragen müssen die Tageszeitreihen zunächst mit den in Abschnitt 2.5.1 präsentierten Methoden geclustert werden. Anschließend muss der Zusammenhang zwischen den gefundenen Clustern und den Tarifgruppen mithilfe von Verbundverteilungen und informationstheoretischen Maßen bewertet werden. Erst wenn die

Verbundverteilung (Cluster enthalten mehrheitlich Datentupel einer Tarif- oder Kontrollgruppe) und die informationstheoretischen Maße (durch hohe Werte für Q_A) eine große Abhängigkeit zwischen den Ausgangsgrößen „Cluster“ und „Tarifgruppe“ aufzeigen, entfalten die mittleren Clusterkurven ihre Aussagekraft. Cluster, die mehrheitlich Datentupel einer Kontrollgruppe enthalten, repräsentieren das unbeeinflusste Verbrauchsverhalten. Cluster, die mehrheitlich Datentupel einer Tarifgruppe enthalten, repräsentieren entsprechend das preisbeeinflusste Verbrauchsverhalten dieser Tarifgruppe. Durch visuelle Inspektion und Vergleich der gefundenen Clusterkurven lassen sich so qualitative Antworten auf die eben genannten Fragen geben. Es muss abschließend betont werden, dass die Kombination aus Clustering (Abschnitt 2.5.1) und Clusterzuordnungen (Abschnitt 2.5.3) grundsätzlich qualitative Aussagen über das preisbeeinflusste Verbrauchsverhalten ermöglicht, dies jedoch nicht für alle *Smart-Meter*-Datensätze garantiert werden kann.

2.5.4 Sensitivitätsanalysen

Bei den Sensitivitätsanalysen geht es darum, die Sensitivität (alternative Bezeichnung: Empfindlichkeit) eines Clusterergebnisses auf die Änderung bestimmter Einflussgrößen zu untersuchen. In der vorliegenden Arbeit wird die Sensitivität der Clusterergebnisse bezüglich folgender Einflussgrößen ermittelt:

- die Signalcharakteristik der Lastkurven in den (*Smart-Meter*-) Datensätzen (Aufbau des Lastsignals aus verschiedenen Signalkomponenten, die unterschiedlich stark gewichtet sind; siehe additives Komponentenmodell in Gleichung (1.1)),
- die Strategie zur Wahl der Startcluster (siehe Strategien 1 bis 4 in Abschnitt 2.5.1.2),
- die Clusteranzahl C , die im Wertebereich $[\underline{C}, \bar{C}]$ mit der Stufenbreite ΔC variiert wird,
- der Fuzzifizierungsgrad q , der im Wertebereich $[\underline{q}, \bar{q}]$ mit der Stufenbreite Δq variiert wird und
- der Rauschabstand δ , der im Wertebereich $[\underline{\delta}, \bar{\delta}]$ mit der Stufenbreite $\Delta \delta$ variiert wird.

Die in Abschnitt 2.5.4 vorgestellten Sensitivitätsanalysen erweitern die Untersuchungen in [76, 209] insofern, als zusätzlich die beiden Clusterparameter q und δ variiert werden, um die Sensitivität bezüglich dieser Parameter zu analysieren. Neben dem bereits erwähnten Gütemaß S und der optimalen Clusteranzahl C_{opt} werden die Clusterergebnisse noch hinsichtlich des prozentualen Anteils τ_{oc} der Datentupel im Rauschcluster bezogen auf die Gesamtanzahl der Datentupel bewertet. τ_{oc} berechnet sich folgendermaßen:

$$\tau_{oc} = \frac{N_{D,oc}}{N_D}, \quad (2.57)$$

wobei $N_{D,oc}$ die Anzahl an Datentupeln im Rauschcluster und N_D die Gesamtanzahl der Datentupel ist.

Jeder Sensitivitätsanalyse in Abschnitt 2.5.4 liegt ein Versuchsplan zugrunde, der alle möglichen Kombinationen der Ausprägungen der Einflussgrößen testet und das Clustering N_{CP} -fach mit

anderen Ausprägungen der Einflussgrößen durchführt. Die Anzahl an Clusterdurchläufen N_{CP} berechnet sich wie folgt:

$$N_{CP} = N_{BD} \cdot N_{SC} \cdot (N_P + N_{P,norm} + N_{P,normmean}), \quad (2.58)$$

wobei N_{BD} die Anzahl an Datensätzen mit unterschiedlicher Lastsignalcharakteristik, N_{SC} die Anzahl an Strategien zur Wahl der Startcluster, N_P die Anzahl an unterschiedlichen Parameterkombinationen für nichtnormierte Tageszeitreihen (siehe Fall DH in Tabelle 2.3), $N_{P,norm}$ die Anzahl an unterschiedlichen Parameterkombinationen für varianznormierte Tageszeitreihen (siehe Fall DH_{norm} in Tabelle 2.3) und $N_{P,normmean}$ die Anzahl an unterschiedlichen Parameterkombinationen für mittelwertnormierte Tageszeitreihen ist (siehe Fall $DH_{normmean}$ in Tabelle 2.3). N_P , $N_{P,norm}$ und $N_{P,normmean}$ berechnen sich gemäß

$$N_i = \left(\frac{\bar{C} - \underline{C}}{\Delta C} + 1 \right) \cdot \left(\frac{\bar{q} - \underline{q}}{\Delta q} + 1 \right) \cdot \left(\frac{\bar{\delta} - \underline{\delta}}{\Delta \delta} + 1 \right), \quad (2.59)$$

mit $i = \{P, P_{norm}, P_{normmean}\}$. In Abb. 2.17 ist der Aufbau und Ablauf der Sensitivitätsanalysen schematisch dargestellt.

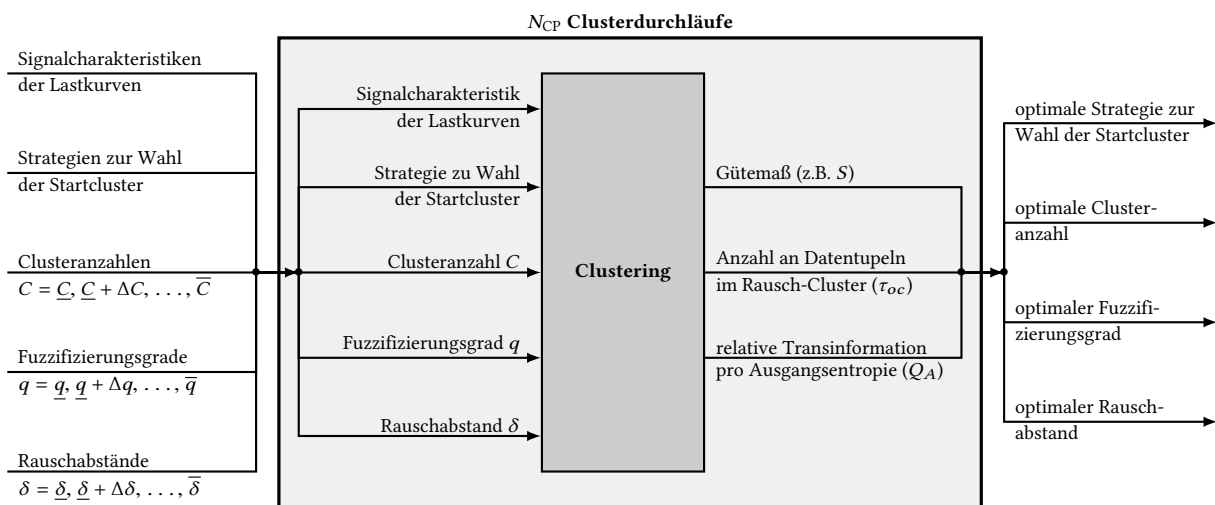


Abbildung 2.17: Schematische Darstellung des Ablaufs der Sensitivitätsanalysen in Abschnitt 2.5.4

Zur schnellen Bewertung der Sensitivitäten eignet sich die Darstellung der Werte des Xie-Beni-Index, der optimalen Clusteranzahl, der Anzahl an Datentupeln im Rausch-Cluster und der Werte der relativen Transinformation pro Ausgangsentropie in Abhängigkeit von C , q und δ als Heatmap oder 3D-Oberflächengrafik (siehe Abschnitte 4.1.2 und 4.2.2). Im Folgenden werden drei Analysen vorgestellt, die die Sensitivität der Clusterergebnisse bezüglich der Clusterparameter (siehe Abschnitt 2.5.4.1), bezüglich der Lastsignalcharakteristik (siehe Abschnitt 2.5.4.2) und bezüglich der Wahl der Startcluster (siehe Abschnitt 2.5.4.3) untersuchen sollen.

2.5.4.1 Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern

Der auf Seite 66 vorgestellte Clusteralgorithmus geht davon aus, dass der Fuzzifizierungsgrad q und der Rauschabstand δ für eine bestimmte Clusteranzahl C manuell festgelegt wird. Des Weiteren werden die Startcluster mithilfe von einer der vier Strategien aus Abschnitt 2.5.1.2 initial festgelegt. Auf Basis eines Datensatzes, der Lastkurven mit einer speziellen Signalcharakteristik enthält, berechnet der Clusteralgorithmus anschließend die Clusterzentren. Allerdings kann so nicht gewährleistet werden, dass die gewählten Clusterparameter zu bestmöglichen Ergebnissen führen. Eine Sensitivitätsanalyse kann helfen, den Einfluss von Clusterparametern auf das Ergebnis des Clusterings besser zu verstehen. Aus diesem Grund wird in Abschnitt 2.5.4.1 die Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparameter vorgestellt, wobei Tabelle 2.4 die Konfiguration der Sensitivitätsanalyse enthält.

Tabelle 2.4: Konfiguration der Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern

Einflussgröße	Ausprägung	Anzahl an Ausprägungen
Lastsignalcharakteristik	Fälle DH, DH_{norm} und DH_{normmean} des Benchmarkdatensatzes „Settlement210H365D_70M20S“ aus Abschnitt 2.3.3.3	$N_{BD} = 1$
Strategien zur Wahl der Startcluster	Strategie 2: Auswahl gleichverteilter Startcluster aus Abschnitt 2.5.1.2	$N_{SC} = 1$
Clusteranzahl C Fuzzifizierungsgrad q Rauschabstand δ	$\underline{C} = 3, \bar{C} = 15, \Delta C = 1$ $\underline{q} = 1.05, \bar{q} = 1.25, \Delta q = 0.05$ $\underline{\delta} = 0.5, \bar{\delta} = 1.5, \Delta \delta = 0.25$	$N_P = 325$
Clusteranzahl C Fuzzifizierungsgrad q Rauschabstand δ	$\underline{C} = 3, \bar{C} = 15, \Delta C = 1$ $\underline{q} = 1.05, \bar{q} = 1.25, \Delta q = 0.05$ $\underline{\delta} = 0.5, \bar{\delta} = 1.5, \Delta \delta = 0.25$	$N_{P, \text{norm}} = 325$
Clusteranzahl C Fuzzifizierungsgrad q Rauschabstand δ	$\underline{C} = 3, \bar{C} = 15, \Delta C = 1$ $\underline{q} = 1.05, \bar{q} = 1.25, \Delta q = 0.05$ $\underline{\delta} = 0.5, \bar{\delta} = 1.5, \Delta \delta = 0.25$	$N_{P, \text{normmean}} = 325$
Anzahl an Clusterdurchläufen		$N_{CP} = 975$

Die Ergebnisse der ersten Sensitivitätsanalyse für die Fälle DH, DH_{norm} und DH_{normmean} sind als 3D-Oberflächengrafiken in Abb. 2.18, B.5 und B.7 und als Heatmaps in Abb. 2.19, B.6 und B.8 dargestellt. Die optimale Clusteranzahl liegt in allen drei Fällen im Bereich von $C = \{4, 5, 6, 7, 8, 9, 14\}$, weswegen in Abb. 2.18, B.5 und B.7 keine Oberflächengrafiken für $C = \{10, 11, 12, 13, 15\}$ aufgeführt sind. Die roten (alle Cluster enthalten Datentupel) und weißen (mindestens ein Cluster enthält keine Datentupel) Kreise kennzeichnen diejenigen δ, q -Parameterkombinationen, für die sich eine optimale Clusteranzahl einstellt.

Die Ergebnisse der ersten Sensitivitätsanalyse lassen sich wie folgt zusammenfassen, wobei nur der Fall DH berücksichtigt wird:

- Abb. 2.18a und 2.19a: Es stellt sich insgesamt eine homogene Verteilung der optimalen Clusteranzahlen ($C_{opt} = \{4, 5, 6, 7\}$) für verschiedene δ, q -Parameterkombinationen ein. Jedoch sorgen Parameterkombinationen mit großen Werten des Fuzzifizierungsgrades ($q = 1.25$) und großen Werten des Rauschabstandes ($\delta = \{1, \dots, 1.5\}$) dafür, dass die optimale Clusteranzahl mit $C_{opt} = 8$ etwas von dieser homogenen Verteilung abweicht. Einen deutlichen Ausreißer stellt die ($\delta = 1, q = 1.1$)-Parameterkombination mit einer optimalen Clusteranzahl von $C_{opt} = 14$ dar.
- Abb. 2.18b und 2.19c: Es wird deutlich, dass der Fuzzifizierungsgrad q bei einer geringen Clusteranzahl ($C = 4$) einen geringen Einfluss auf die Anzahl an Datentupeln im Rauschcluster hat. Bei mittleren und hohen Rauschabständen ($\delta = \{1, 1.25, 1.5\}$) und hohen Clusteranzahlen ($C = \{6, 7, 8, 14\}$) landen insgesamt sehr wenige Datentupel¹⁴ im Rauschcluster.
Bei geringen Rauschabständen ($\delta = \{0.5, 0.75\}$) und hohen Clusteranzahlen ($C = \{6, 7, 8\}$) gewinnt der Fuzzifizierungsgrad q an Einfluss. Ein geringer Fuzzifizierungsgrad ($q = 1.1$) führt hier im Vergleich zu einem hohen Fuzzifizierungsgrad ($q = \{1.2, 1.25\}$) dazu, dass vermehrt Datentupel im Rauschcluster landen.
- Abb. 2.19d bis 2.19g: Die Clusterlösungen für bestimmte δ, q -Parameterkombinationen sind zunächst als gleichwertige Lösungen anzusehen. Um aus der Menge an Clusterlösungen eine Empfehlung für bestimmte Clusterlösungen und damit für bestimmte δ, q -Parameterkombinationen geben zu können, ist es ratsam, die Werte der relativen Transinformation pro Ausgangsentropie in Abb. 2.19d bis 2.19g zu betrachten.
Der Maximalwert der relativen Transinformation pro Ausgangsentropie für die Ausgangsgrößen „Tarif“ und „Cluster C_{opt} “ bzw. für die Ausgangsgrößen „Responder“ und „Cluster C_{opt} “ liegt bei 0,0063. Sowohl die Klassen der Ausgangsgröße „Tarif“ („TOU2“, „TOU3“ und „FIXED“) als auch die Klassen der Ausgangsgröße „Responder“ (NR-, SR- und R-Haushalte) lassen sich nicht durch die gefundenen Cluster erklären.
Allerdings zeigen die hohen Werte der relativen Transinformation pro Ausgangsentropie in Abb. 2.19d, dass die gefundenen DH-Cluster sehr gut das saisonal geprägte Verbrauchsverhalten beschreiben.
Das unterschiedliche Verbrauchsverhalten an Werk- und Wochenendtagen kann mit den gefundenen Clustern nur eingeschränkt erklärt werden. Die Werte der relativen Transinformation pro Ausgangsentropie in Abb. 2.19e liegen unter den Werten in Abb. 2.19d. Am besten lassen sich die drei Klassen der Ausgangsgröße „Werktag“ („Werktag“, „Samstag“, „Sonntag/Feiertag“) durch die acht Cluster der ($\delta = 1.5, q = 1.25$)-Parameterkombination erklären.
- Es kann keine eindeutige Empfehlung für eine optimale δ, q -Parameterkombination gegeben werden. Die Sensitivitätsanalyse macht jedoch deutlich, dass bei der Wahl der Clusterparameter sehr sorgsam umgegangen werden muss.

¹⁴Der Minimalwert für die Anzahl an Datentupeln im Rauschcluster liegt bei 0.04% für $C_{opt} = 7, \delta = 1.5$ und $q = 1.1$.

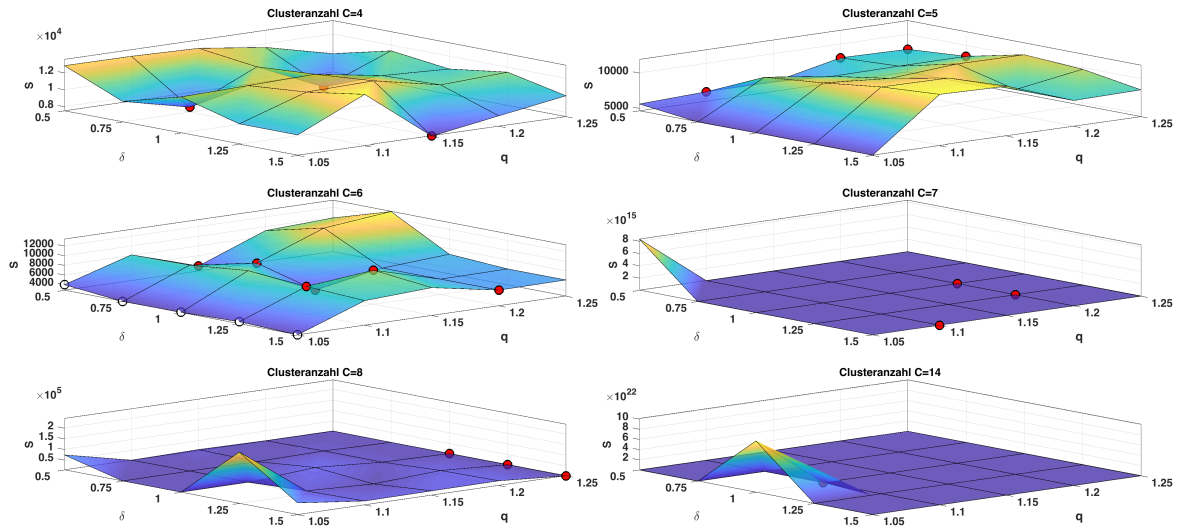
Die Normierung der Tageszeitreihen bewirkt, dass das saisonal geprägte Verbrauchsverhalten in den Hintergrund und das unterschiedliche Verbrauchsverhalten an Werk- und Wochenendtagen in den Vordergrund rückt. Die Klassen der Ausgangsgröße „Werktag“ lassen sich durch die DH_{norm} - bzw. DH_{normmean} -Cluster besser erklären als durch die DH-Cluster.

In der zweiten (siehe Abschnitt 2.5.4.2) und dritten Sensitivitätsanalyse (siehe Abschnitt 2.5.4.3) wird der Einfluss der Lastsignalcharakteristik bzw. der Einfluss der Strategie zur Wahl der Startcluster untersucht. Dazu werden die Clusterparameter C , δ und q konstant gehalten. Somit müssen für diese Analysen vernünftige Werte für C , δ und q festgelegt werden, die aus der ersten Sensitivitätsanalyse mit dem Benchmarkdatensatz „Settlement210H365D_70M20S“ gewonnen werden.

Die Bestimmung der optimalen Clusterparameter lässt sich als multikriterielles Optimierungsproblem begreifen, bei dem die Teilziele „Maximierung von Q_A (Monat, Werktag, Tarifgruppe)“, „akzeptable Anzahl an Datentupeln im Rauschcluster“ und „möglichst geringe Anzahl an Clusterzentren“ durch eine Lösung gleichzeitig zu erfüllen sind. Grundsätzlich kann ein solches Optimierungsproblem dadurch gelöst werden, dass die Teilziele mithilfe von Gewichtungsfaktoren zu einer Zielfunktion zusammengefasst werden, die anschließend minimiert wird. Hier erfolgt die Auswahl einer optimalen Lösung jedoch nicht nach streng mathematischen Gesichtspunkten, sondern durch die subjektive Einschätzung aller Lösungen und durch Auswahl einer „optimalen Kompromisslösung“ bezüglich der drei Teilziele.

Für den DH-Fall werden die Clusterparameter $C = 6$, $\delta = 1.25$ und $q = 1.1$ gewählt. Eine gute Clusterlösung des DH_{norm} -Falls kann mit den Clusterparametern $C = 8$, $\delta = 1.0$ und $q = 1.05$ erzielt werden. Für den DH_{normmean} -Fall ergeben die Clusterparameter $C = 7$, $\delta = 1.0$ und $q = 1.15$ kompakte und gut voneinander unterscheidbare Cluster.

(a) Sensitivität des Trennungsgrades



(b) Sensitivität des prozentualen Anteils der Datentupel im Rauschcluster

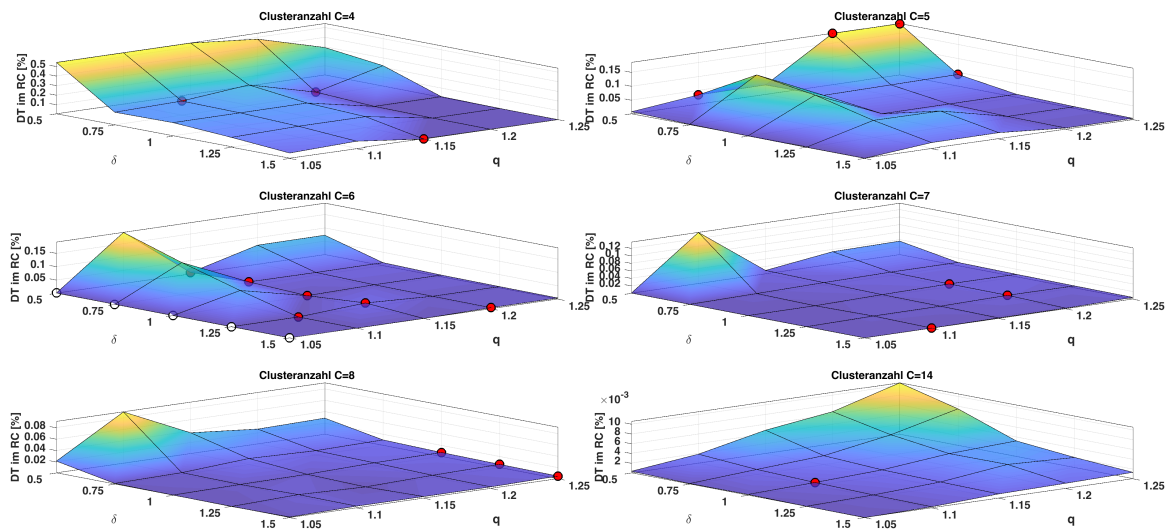


Abbildung 2.18: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH des Benchmarkdatensatzes „Settlement210H365D_70M20S“ ((a) Trennungsgrad, (b) prozentualer Anteil an Datentupeln im Rauschcluster). Für Fall DH_{norm} siehe Abb. B.5, für Fall $DH_{normmean}$ siehe Abb. B.7. Ergänzende Visualisierungen in Abb. 2.19. Die roten (alle Cluster enthalten Datentupel) und weißen (mindestens ein Cluster enthält keine Datentupel) Kreise kennzeichnen diejenigen δ, q -Parameterkombinationen, für die sich eine optimale Clusteranzahl einstellt.

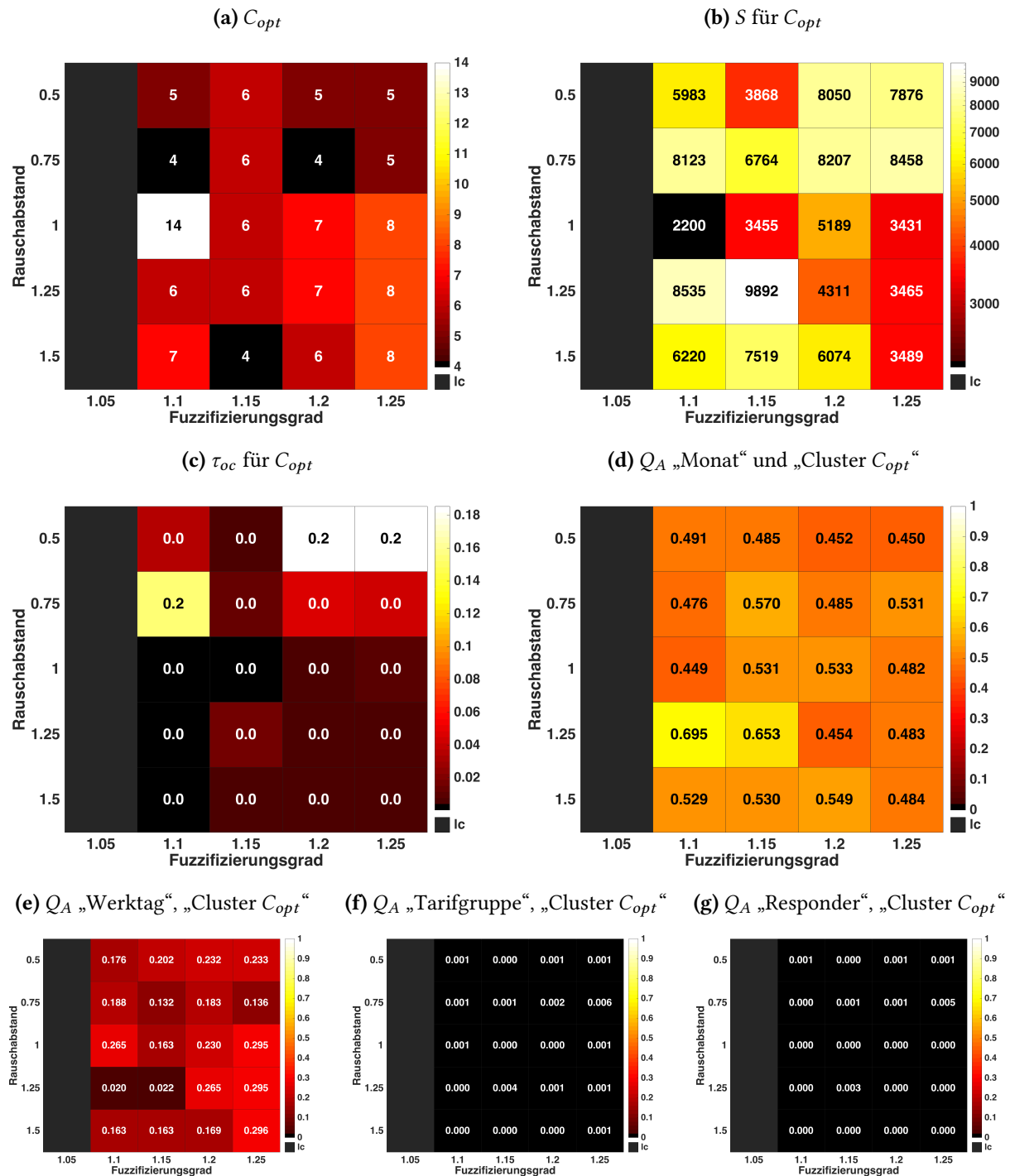


Abbildung 2.19: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH des Benchmarkdatensatzes „Settlement210H365D_70M20S“ ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(g) relative Transinformation pro Ausgangsentropie). Für Fall DH_{norm} siehe Abb. B.6, für Fall $DH_{normmean}$ siehe Abb. B.8. Ergänzende Visualisierungen zu Abb. 2.18. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

2.5.4.2 Sensitivität der Clusterergebnisse bezüglich der Lastsignalcharakteristik

Mit der zweiten Analyse soll herausgefunden werden, wie sich das Clusterergebnis bei variierender Lastsignalcharakteristik verhält. Die Analyse soll die Frage klären, bis zu welchen Anteilen der Musterkomponente der Clusteralgorithmus mit konstanten Clusterparametern ähnliche, sinnvolle Clusterergebnisse liefert. Die Datenbasis für die Parameterstudie bilden die Benchmarkdatensätze aus Abschnitt 2.3.3.3. Die Parametereinstellungen des Benchmarkgenerators zur Erzeugung der Benchmarkdatensätze stehen in Tabelle B.7. Die Konfiguration der Sensitivitätsanalyse ist in Tabelle 2.5 dargestellt. Dabei gilt zu beachten, dass die Clusterparameter C , δ und q und die Strategie zur Wahl der Startcluster für alle 19 Clusterdurchläufe je Fall konstant gehalten wurden.

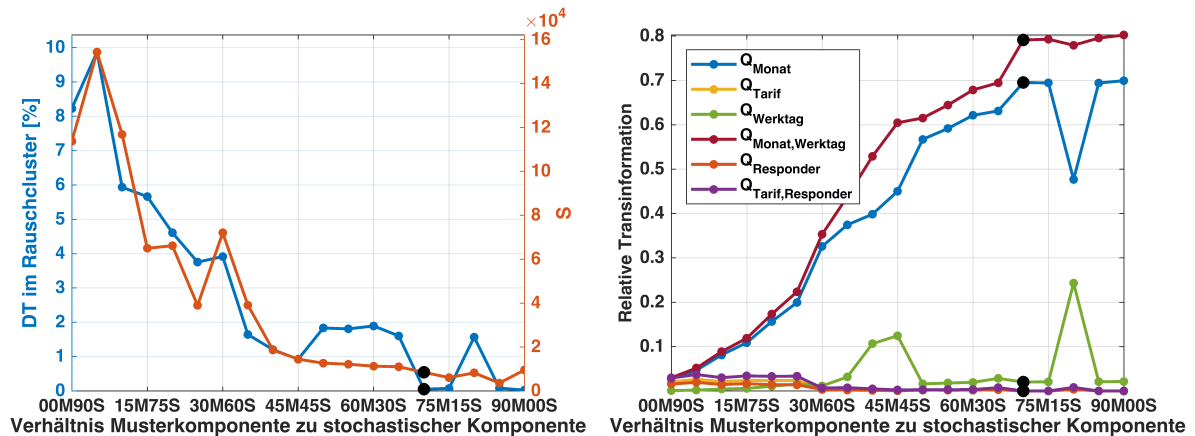
Tabelle 2.5: Konfiguration der Sensitivitätsanalyse zur Bestimmung des Einflusses der Lastsignalcharakteristik

Einflussgröße	Ausprägung	Anzahl an Ausprägungen
Lastsignalcharakteristik	Fälle DH, DH_{norm} und $DH_{normmean}$ aller Benchmarkdatensätze aus Abschnitt 2.3.3.3	$N_{BD} = 19$
Strategie zur Wahl der Startcluster	Strategie 2: Auswahl gleichverteilter Startcluster aus Abschnitt 2.5.1.2	$N_{SC} = 1$
Clusteranzahl C Fuzzifizierungsgrad q Rauschabstand δ	$C = 6$ $q = 1.1$ $\delta = 1.25$	$N_P = 1$
Clusteranzahl C Fuzzifizierungsgrad q Rauschabstand δ	$C = 8$ $q = 1.05$ $\delta = 1.0$	$N_{P,norm} = 1$
Clusteranzahl C Fuzzifizierungsgrad q Rauschabstand δ	$C = 7$ $q = 1.15$ $\delta = 1.0$	$N_{P,normmean} = 1$
Anzahl an Clusterdurchläufen		$N_{CP} = 57$

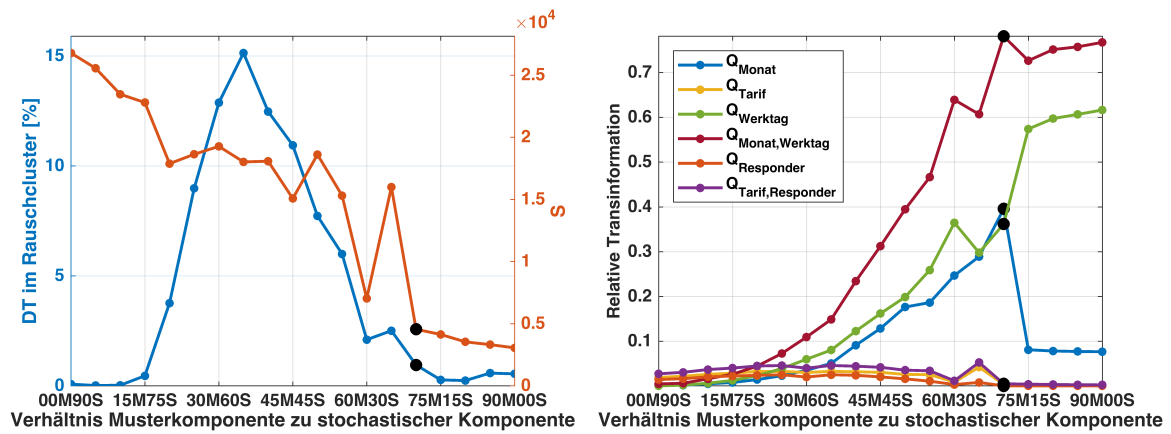
Die Ergebnisse der zweiten Sensitivitätsanalyse sind für die DH-Fälle in Abb. 2.20a, für die DH_{norm} -Fälle in Abb. 2.20b und für die $DH_{normmean}$ -Fälle in Abb. 2.20c grafisch dargestellt. Aus Abb. 2.20 lassen sich folgende Beobachtungen ableiten:

- Abb. 2.20a: Liegt der Anteil der Musterkomponente im Bereich von 50-90% (stochastische Komponente im Bereich von 40-0%), ergibt das Clustering der DH-Fälle mit den Clusterparametern $C = 6$, $q = 1.1$ und $\delta = 1.25$ nur geringfügig unterschiedliche Clusterergebnisse. Sowohl die Werte für S als auch die Anzahl der Datentupel im Rauschcluster liegen in diesem Bereich sehr nahe beieinander. Für die Werte der relativen Transinformation gilt ähnliches.
Ab einem Anteil der stochastischen Komponente von ca. 45% steigen die Werte für S und

(a) DH-Fälle: Clusterparameter $C = 6$, $q = 1.1$ und $\delta = 1.25$



(b) DH_{norm} -Fälle: Clusterparameter $C = 8$, $q = 1.05$ und $\delta = 1.0$



(c) $DH_{normmean}$ -Fälle: Clusterparameter $C = 7$, $q = 1.15$ und $\delta = 1.0$

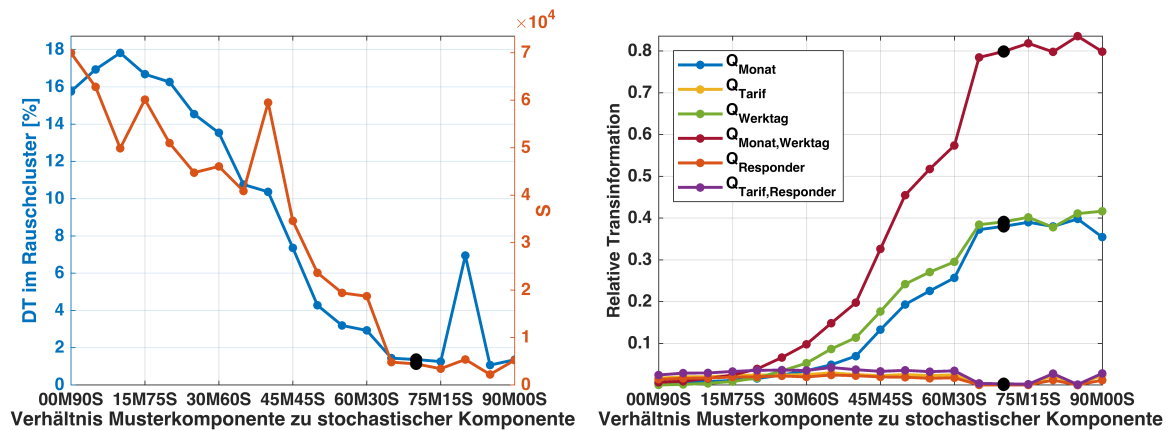


Abbildung 2.20: Sensitivität der Clusterergebnisse bezüglich der Lastsignalcharakteristik für (a) DH-Fälle, (b) DH_{norm} -Fälle und (c) $DH_{normmean}$ -Fälle. Die schwarzen Punkte kennzeichnen den Benchmarkdatensatz, der in der ersten Sensitivitätsanalyse in Abschnitt 2.5.4.1 verwendet wurde und für den die gewählten Clusterparameter eine gute Clusterlösung ergaben.

die Anzahl der Datentupel im Rauschcluster stark an. Der erhöhte Anteil der stochastischen Komponente sorgt dafür, dass keine kompakten und klar voneinander unterscheidbaren Cluster gefunden werden können.

- Abb. 2.20b: Auch beim Clustering der varianznormierten Tageskurven mit den Clusterparametern $C = 8$, $q = 1.05$ und $\delta = 1.0$ weisen die Datensätze mit hohen Anteilen der Musterkomponente kompaktere Cluster auf. Die geringsten Werte für S ergeben sich bei Datensätzen mit einem Anteil der Musterkomponente im Bereich von 70-90%. Bei Datensätzen mit einem sehr hohen Anteil der stochastischen Komponente (75-90%) können mit den Clusterparametern $C = 8$, $q = 1.15$ und $\delta = 1$ keine sinnvollen Cluster gefunden werden. Zudem enthalten die Rauschcluster so gut wie keine Datentupel. Der Verlauf der Werte der relativen Transinformation ist ein ähnlicher wie in Abb. 2.20a.
- Abb. 2.20c: Die Werte für S und die Anzahl an Datentupeln im Rauschcluster nehmen einen ähnlichen Verlauf wie die Werte in Abb. 2.20a. Bei Anteilen der Musterkomponente im Bereich von 50-90% werden kompakte Cluster durch den Clusteralgorithmus entdeckt. Bei Anteilen der stochastischen Komponente im Bereich von 50-90% nehmen die Werte für S und die Anzahl an Datentupeln im Rauschcluster stark zu. Eine Ausnahme bildet das Clustering des Benchmarkdatensatzes „Settlement210H365D_80M10S“. Hier liegt die Anzahl an Datentupeln im Rauschcluster ungewöhnlich hoch.
- Die Werte der relativen Transinformation auf der rechten Seite der Abb. 2.20a bis 2.20c zeigen einen weiteren interessanten Effekt. Der saisonale Einfluss auf das Verbrauchsverhalten kann am ehesten mit den DH-Clustern bewertet werden, da hier die Tageszeitreihen nicht normiert wurden und, bedingt durch unterschiedliche Temperaturen, unterschiedliche Stromverbräuche aufweisen. Die hohen Werte für Q_{Monat} in Abb. 2.20a belegen dies. Die geringen Werte für Q_{Werktag} zeigen dagegen, dass ein wochentagsbedingtes Verbrauchsverhalten nicht mit DH-Clustern erklärt werden kann.
- Die Normierung der Tageskurven führt dazu, dass der Clusteralgorithmus Clusterzentren findet, die den Einfluss des Wochentags besser erklären. Die hohen Werte für Q_{Werktag} in Abb. 2.20b und 2.20c belegen dies.
- Die sehr geringen Werte für Q_{Tarif} und $Q_{\text{Responder}}$ in Abb. 2.20a bis 2.20c lassen erkennen, dass der Clusteralgorithmus Schwierigkeiten hat, Cluster zu finden, die sich durch unterschiedliche Tarife oder Verbrauchsreaktionen erklären lassen.

Die haushaltsindividuellen Schwankungen in den Lastzeitreihen verhindern, dass das Clustering DH-/DH_{norm}-/DH_{normmean}-Cluster findet, die den Tarifeinfluss gut erklären. Hingegen zeigen die Ergebnisse der Sensitivitätsanalyse, dass unterschiedliche Tagesverläufe an Werk- und Wochenendtagen und unterschiedliche Monatsverläufe gut durch das Clustering der DH-/DH_{norm}-/DH_{normmean}-Fälle aufgedeckt werden.

Um eine qualitative Aussage über die Clusterergebnisse für die aggregierten DT-/DT_{norm}-/DT_{normmean}-Fälle treffen zu können, wurde abschließend eine weitere Sensitivitätsanalyse durchgeführt. Die Ergebnisse der Sensitivitätsanalyse für die DT-Fälle sind in Abb. B.9a, für die DT_{norm}-Fälle in Abb. B.9b und für die DT_{normmean}-Fälle in Abb. B.9c dargestellt. Es muss angemerkt werden, dass die hier verwendeten optimalen Clusterparameter nicht neu

ermittelt, sondern direkt aus Tabelle 2.5 übernommen wurden. Die Ergebnisse aus Abb. B.9 zeigen, dass auch hier unterschiedliche Tagesverläufe an Werk- und Wochenendtagen und unterschiedliche Monatsverläufe gut durch das Clustering der aggregierten Fälle aufgedeckt werden. Ein deutlicher Unterschied zwischen den DH -/ DH_{norm} -/ $DH_{normmean}$ -Clustern und den DT -/ DT_{norm} -/ $DT_{normmean}$ -Clustern zeigt sich vor allem darin, dass DT -/ DT_{norm} -/ $DT_{normmean}$ -Cluster den Tarifeinfluss besser erklären als DH -/ DH_{norm} -/ $DH_{normmean}$ -Cluster, insbesondere nach der Normierung.

2.5.4.3 Sensitivität der Clusterergebnisse bezüglich der Wahl der Startcluster

Die dritte Analyse soll die Robustheit der Clusterergebnisse bei anderen Startclustern untersuchen. Dabei werden alle vier in Abschnitt 2.5.1.2 vorgestellten Strategien zur Wahl der Startcluster getestet. Als Datenbasis dieser Analyse dienen die Fälle DH , DH_{norm} und $DH_{normmean}$ des Benchmarkdatensatzes „Settlement210H365D_70M20S“ aus Abschnitt 2.3.3.3. Die Konfiguration der Sensitivitätsanalyse zur Bestimmung des Einflusses der initialen Startcluster ist in Tabelle 2.6 zusammengefasst. Da sich die Clusterergebnisse trotz gleicher Clusterparameter unterscheiden können, wurde das Clustering 10-fach wiederholt.

Tabelle 2.6: Konfiguration der Sensitivitätsanalyse zur Bestimmung des Einflusses der initialen Startcluster

Einflussgröße	Ausprägung	Anzahl an Ausprägungen
Lastsignalcharakteristik	DH -Fall des Benchmarkdatensatzes „Settlement210H365D_70M20S“ aus Abschnitt 2.3.3.3	$N_{BD} = 1$
Strategie zur Wahl der Startcluster	Strategien 1 (zufällige Auswahl der Startcluster), 2 (Auswahl gleichverteilter Startcluster), 3 (Auswahl der Startcluster aus Fuzzy-Menge) und 4 (Aufteilung der Startcluster auf Wochentage) aus Abschnitt 2.5.1.2	$N_{SC} = 4$
Clusteranzahl C Fuzzifizierungsgrad q Rauschabstand δ	$\underline{C} = 3, \bar{C} = 15, \Delta C = 1$ $q = 1.1$ $\delta = 1.25$	$N_P = 13$
Clusteranzahl C Fuzzifizierungsgrad q Rauschabstand δ	$\underline{C} = 3, \bar{C} = 15, \Delta C = 1$ $q = 1.05$ $\delta = 1.0$	$N_{P,norm} = 13$
Clusteranzahl C Fuzzifizierungsgrad q Rauschabstand δ	$\underline{C} = 3, \bar{C} = 15, \Delta C = 1$ $q = 1.15$ $\delta = 1.0$	$N_{P,normmean} = 13$
Anzahl an Clusterdurchläufen		$N_{CP} = 156 \cdot 10$

Das Ergebnis des Vergleichs der vier Strategien für den DH-Fall ist in Abb. 2.21, das Ergebnis für den DH_{norm} -Fall in Abb. B.10 und das Ergebnis für den DH_{normmean} -Fall in Abb. B.10 dargestellt.

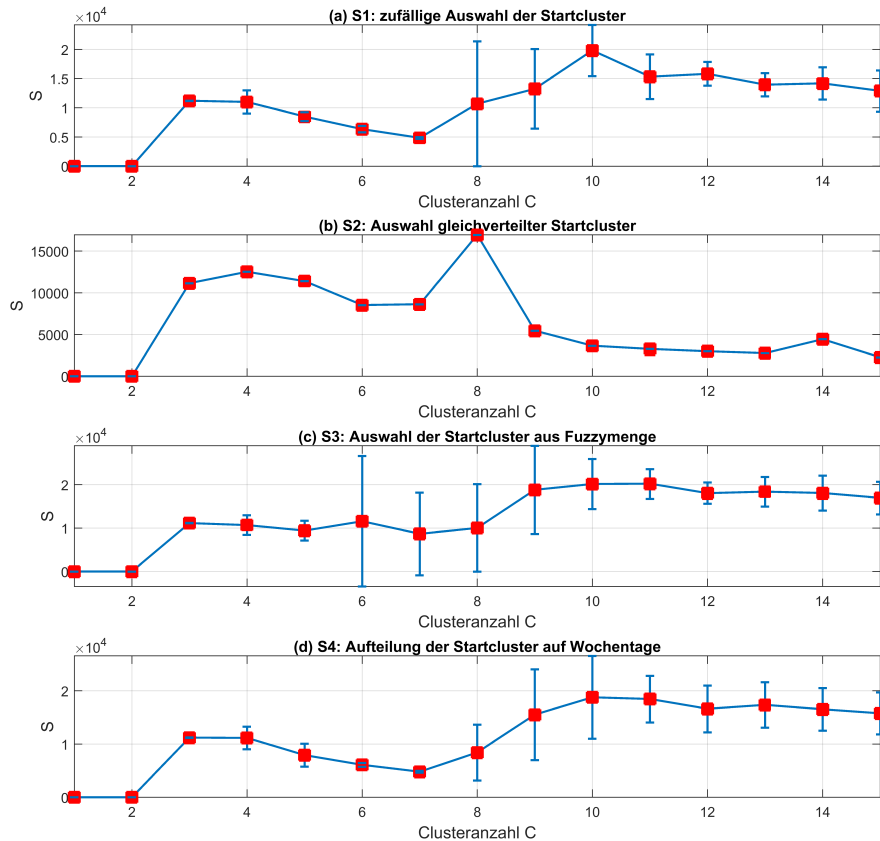


Abbildung 2.21: Sensitivität der Clusterergebnisse bezüglich der Wahl der Startcluster für den DH-Fall des Benchmarkdatensatzes „Settlement210H365D_70M20S“ (für Fall DH_{norm} siehe Abb. B.10, für Fall DH_{normmean} siehe Abb. B.11): das Clustering wurde 10-fach mit den Clusterparametern $C = 3, \dots, 15$, $q = 1.1$ und $\delta = 1.25$ durchgeführt. Die blaue Linie kennzeichnet den mittleren Verlauf der Werte für S bei 10-facher Wiederholung des Clusterings. Die Balken geben die Streuung der Werte für S um den Mittelwert an.

Es lassen sich folgende Beobachtungen aus den Abb. 2.21, B.10 und B.11 bezüglich der Sensitivität der Clusterergebnisse bezüglich der Wahl der Startcluster ableiten:

- Abb. 2.21: Für geringe Clusteranzahlen ($C = 3, 4, 5$) liegen die Werte für S je Strategie dicht beieinander. Beim Vergleich der Strategien fällt zudem auf, dass die Werte für S einen ähnlichen Verlauf nehmen. Ab einer Clusteranzahl von $C = 6$ steigt die Streuung für S für die Strategien 3 und 4 stark an. Ab einer Clusteranzahl von $C = 7$ führen ungünstige Startwerte bei Strategie 1 zu weniger sinnvollen Clustern. Strategie 2 mit der Auswahl gleichverteilter Startcluster führt als einzige der vier Strategien bei gleichen Clusterparametern zu gleichen Clusterergebnissen.
- Abb. B.10 und B.11: Der Vergleich der Clusterergebnisse bei normierten Zeitreihen zeigt, dass die Strategien 1, 3 und 4 keine reproduzierbaren Ergebnisse liefern. Zudem unterscheidet sich der Verlauf der Werte für S von Strategie zu Strategie sehr stark. Analog zu

den Ergebnissen in Abb. 2.21 führt Strategie 2 mit der Auswahl gleichverteilter Startcluster als einzige Strategie bei gleichen Clusterparametern zu gleichen Clusterergebnissen.

Für alle nachfolgenden Clusteranalysen wird Strategie 2 aus Abschnitt 2.5.1.2 zur Wahl der Startcluster ausgewählt.

2.5.5 Analyse einzelner Abtastzeitpunkte bezüglich der Tarifeinflüsse auf die Last

Prozentuale Energieeinsparungen, hervorgerufen durch variable Strompreise, lassen sich durch die Analyse der Fälle SH und ST (vgl. Tabelle 2.3) bewerten. Sollen dagegen zeitliche Lastverschiebungen beurteilt werden, eignen sich eher deren normierte Fälle SH_{normmean} und ST_{normmean} , da eine (eventuelle) prozentuale Energieeinsparung bereits herausgerechnet ist. Die Fälle SH, ST, SH_{normmean} und ST_{normmean} resultieren aus der Umwandlung jedes Abtastzeitpunktes der Zeitreihen von DH, DT, DH_{normmean} und DT_{normmean} in einzelne Datentupel. So wird z.B. die Preiszeitreihe $p^{T_i}[k]$ des Tarifs T_i mit $k = 1, \dots, K$ in das Einzelmerkmal $p_{n_D}^{T_i}$ mit den Datentupeln $n_D = 1, \dots, K$ umgewandelt. Anschließend werden, ähnlich wie in Abb. 2.11 gezeigt, die unterschiedlichen Preise eines Tarifs in Klassen eingeteilt und die korrespondierenden Leistungswerte untersucht. Dabei werden die Mittelwerte und Streuungen der Leistungswerte aller Datentupel, die zu einer bestimmten Preisklasse gehören, berechnet und mit denen einer anderen Tarifgruppe verglichen. Zur Bewertung der Mittelwertunterschiede kommt der Zweistichproben-t-Test zum Einsatz [274–276]. Die Eigenschaften der *Smart-Meter*-Daten verletzen jedoch stellenweise die Vorbedingungen für den Einsatz dieses Hypothesentests (Datenumfang, identische Varianz in den Stichproben). Deswegen werden Sonderformen des Zweistichproben-t-Tests, wie Welch-Test und exakter Test nach Fisher [277], verwendet. Dabei wird die Verschlechterung des Signifikanzniveaus in Kauf genommen, um verlässlichere Ergebnisse zu erhalten. In Beispiel 9 ist die Analyse einzelner Abtastzeitpunkte bezüglich des Tarifeinflusses auf die Last musterhaft gezeigt.

An dieser Stelle muss unbedingt noch angemerkt werden, dass diese Analyseergebnisse nur dann vergleichbar sind, wenn entweder die gleichen Haushalte bzgl. ihres DR-Verhaltens bei unterschiedlichen Tarifen über einen längeren Zeitraum (z.B. ein Jahr) beobachtet werden, oder die gemessenen Leistungswerte von Haushalten stammen, die durch eine geeignete Studienplanung direkt vergleichbar sind (vgl. Eigenschaft E10 und E11 in den Tabellen 2.1 und 2.2). Ist diese Voraussetzung nicht erfüllt, kann ein kausaler Zusammenhang zwischen Stromtarifen und Lastunterschieden (Mittelwertunterschieden) nur eingeschränkt bewertet werden.

Beispiel 9: Analyse des Tarifeinflusses auf die Last

In diesem Beispiel bildet der Fall SH_{normmean} des Benchmarkdatensatzes „Settlement210H365D_70M20S“ die Datenbasis. In Tabellen 2.7 und 2.8 ist die Auswertung der Preisklassen für den TOU2-Tarif bzw. TOU3-Tarif gegenüber der Kontrollgruppe (FIXED) dargestellt. Hierfür wurde die statistische Signifikanz von Mittelwertunterschieden mithilfe des Zweistichproben-Welch-Tests bewertet. Die Höhe der Signifikanz der Ergebnisse ist in beiden Tabellen mit Sternen gekennzeichnet: si-

gnifikante Ergebnisse ($\alpha < 0.05$) sind mit *, hochsignifikante Ergebnisse ($\alpha < 0.01$) mit ** und höchstsignifikante Ergebnisse ($\alpha < 0.001$) mit *** gekennzeichnet. Die neben den Tabellen gezeichneten Histogramme veranschaulichen grafisch die Verteilung der Datentupel je Preisklasse und Tarifgruppe. Die Analyseergebnisse zeigen, dass Haushalte der TOU2- bzw. TOU3-Tarifgruppe im Vergleich zur Kontrollgruppe die Last zu Hochpreiszeiten verringert und zu Niedrigpreiszeiten erhöht haben, was auf Lastverschiebung schließen lässt.

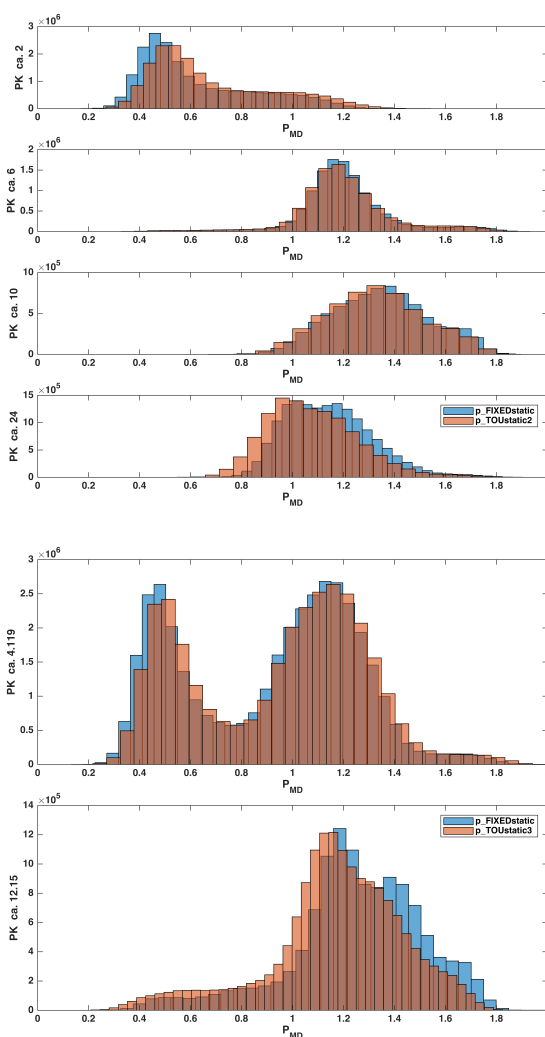
Ein Vergleich zwischen den TOU2-Haushalten und den TOU3-Haushalten ohne Berücksichtigung der Kontrollgruppe führt zu verzerrten Ergebnissen, da zu Hochpreiszeiten (TOU2-Tarif: 6-9 Uhr und 18-21 Uhr) in der Regel ein hoher Stromverbrauch vorliegt. Dies ist jedoch nicht auf das Preissignal zurückzuführen, sondern auf das bekannte Stromverbrauchsverhalten mit ausgeprägten Lastspitzen am Morgen und Abend.

Tabelle 2.7: Fall $SH_{normmean}$: Auswertung der Preisklassen (PK in ct/kWh) für den TOU2-Tarif gegenüber der Kontrollgruppe; N_D : Anzahl der Datentupel pro Gruppe; Angabe der mittleren normierten Leistung, Standardabweichung und Datentupelanzahl je Preisklasse und Gruppe; Zweistichproben-Welch-Test, zweiseitig: FIXED vs. TOU2

PK	N_D	FIXED	TOU2
2	871325	0.621+-0.227	0.679+-0.238***
6	497900	1.205+-0.184	1.203+-0.189***
10	398320	1.349+-0.193	1.328+-0.192***
24	622375	1.144+-0.173	1.077+-0.179***

Tabelle 2.8: Fall $SH_{normmean}$: Auswertung der Preisklassen (PK in ct/kWh) für den TOU3-Tarif gegenüber der Kontrollgruppe; N_D : Anzahl der Datentupel pro Gruppe; Angabe der mittleren normierten Leistung, Standardabweichung und Datentupelanzahl je Preisklasse und Gruppe; Zweistichproben-Welch-Test, zweiseitig: FIXED vs. TOU3

PK	N_D	FIXED	TOU3
4.119	1792440	0.917+-0.343	0.943+-0.342***
12.15	597480	1.250+-0.263	1.171+-0.268***



2.5.6 Zusammenfassung und Diskussion

Das Ziel von Abschnitt 2.5 war es, eine systematische, datenbasierte Analyse der Auswirkungen verschiedener Preissignale auf das Verbrauchsverhalten von Haushaltsstromkunden

vorzustellen. In Abschnitt 2.4 wurden hierfür alle notwendigen Schritte erläutert, die aus den *Smart-Meter*-Rohdaten einen strukturierten *Smart-Meter*-Datensatz mit unterschiedlichen Fällen (siehe Tabelle 2.3) erzeugen. Für die systematische Analyse wird ein Schema mit vier Schritten empfohlen, an dessen Ende eine qualitative und quantitative Bewertung des preisbeeinflussten Verbrauchsverhaltens steht:

- **Schritt 1:** das Zeitreihen-Clustering zur Ermittlung von preisbeeinflussten Typtagen (siehe Abschnitt 2.5.1),
- **Schritt 2:** die Bewertung der Clusterlösung (siehe Abschnitt 2.5.2),
- **Schritt 3:** die Zuordnung der gefundenen Clusterkurven zu Wochentagen, Monaten und Tarifen (siehe Abschnitt 2.5.3) und
- **Schritt 4:** die Analyse einzelner Abtastzeitpunkte bezüglich der Tarifeinflüsse auf die Stromlast (siehe Abschnitt 2.5.5).

Die Mittelwertkurven der gefundenen Cluster stellen als Ergebnis des Clusterings die preisbeeinflussten Typtage dar. Um die Robustheit der Clusterergebnisse bewerten zu können, wurden in Abschnitt 2.5.4 verschiedene Sensitivitätsanalysen anhand der Benchmarkdatensätze aus Abschnitt 2.3.3 durchgeführt. In der ersten Sensitivitätsanalyse in Abschnitt 2.5.4.1 konnte gezeigt werden, dass bei der Wahl der Clusterparameter sehr sorgsam vorgegangen werden muss. Schon kleine Änderungen der Clusterparameter können zu stark veränderten Clusterergebnissen führen. In der zweiten Analyse in Abschnitt 2.5.4.2 wurde der Einfluss der Lastsignalcharakteristik auf die gefundenen Cluster untersucht. Der Einfluss der Lastsignalcharakteristik spielt insofern eine Rolle, als eine Zeitreihenaggregation (z.B. die Umwandlung des Falls DH in den Fall DT) zu Zeitreihen mit einem höheren Musteranteil führt und so der Clusteralgorithmus kompaktere und besser voneinander unterscheidbare, d.h. sinnvollere, Cluster findet. Die dritte Sensitivitätsanalyse in Abschnitt 2.5.4.3 setzte sich das Ziel, den Einfluss der Startcluster beim Clustering von Tageszeitreihen qualitativ zu bewerten, was durch einen Vergleich von vier verschiedenen Strategien zur Wahl der Startcluster erreicht wurde. Der Vergleich zeigt zum einen, dass sich die Clusterergebnisse je nach Strategie z.T. erheblich voneinander unterscheiden. Zum anderen kann eine randomisierte (siehe Strategie 1 in Abschnitt 2.5.1.2) und pseudo-randomisierte (siehe Strategie 3 und 4 in Abschnitt 2.5.1.2) Auswahl der Startcluster keine optimalen und wiederholbaren Clusterergebnisse garantieren. Die Auswahl gleichverteilter Startcluster (siehe Strategie 2 in Abschnitt 2.5.1.2) ist die einzige geeignete Strategie zur Wahl der Startcluster.

Mithilfe der Methoden aus Abschnitt 2.5.3 und der Sensitivitätsanalysen in Abschnitt 2.5.4 konnte demonstriert werden, dass sich die Clusterkurven der normierten Fälle besser eignen, um strukturelle Unterschiede zwischen den Wochentagen zu erklären. Sollen dagegen saisonale Unterschiede in der Verbrauchsreaktion aufgedeckt werden, eignet sich das Clustering der nicht-normierten Fälle besser als das Clustering der normierten Fälle. Strukturelle Unterschiede zwischen den Tarifgruppen lassen sich grundsätzlich mit den gleichen Verfahren aufdecken. Eine eindeutige Zuordnung der Cluster zu den Tarifgruppen konnte jedoch mit den Benchmarkdatensätzen aus Abschnitt 2.3.3 nicht gemacht werden. Möglicherweise würde eine Modifikation des *Demand-Response-Simulators* deutlichere Ergebnisse diesbezüglich liefern. Der Analyseschritt

in Abschnitt 2.5.5 ergänzt die Clusterergebnisse insofern, als dadurch prozentuale Energieeinsparungen erkannt und unterschiedliche Reaktionen auf Preissignale (z.B. für ungewöhnliche Preissignale) aufgezeigt werden können.

Zur Beantwortung der Fragen auf Seite 61 wurde ein datengetriebener Analyseansatz gewählt, der aus den oben beschriebenen vier Schritten besteht und dessen Kern das Zeitreihen-Clustering ist. Alternative Ansätze, wie die visuelle Inspektion von Lastzeitreihen, die Mittelung von Lastzeitreihen über alle Haushalte einer Tarifgruppe oder die Trennung von Lastzeitreihen mithilfe einfacher Heuristiken, wurden bereits zu Beginn von Abschnitt 2.5 vorgestellt und bewertet. Das Zeitreihen-Clustering zeichnet sich deshalb als eine geeignete Methode zur Beantwortung der Fragen auf Seite 61 aus, da das tatsächliche Verbrauchsverhalten eines preisbeeinflussten Haushaltes in den aufgezeichneten Leistungsdaten enthalten ist. Das Verfahren kann ohne umfangreiche Analyse des zugrunde liegenden Systems (z.B. technische Geräteausstattung des Haushaltes) angewendet werden, um typische Verbrauchsmuster aus den *Smart-Meter*-Messdaten zu finden. Das Zeitreihen-Clustering bietet zudem den Vorteil, unterschiedliche Verbrauchsreaktionen auf ein und dasselbe Preissignal ausfindig machen zu können. Mit möglichen Fehlern in den Messdaten, die die Beantwortung der Fragen auf Seite 61 erschweren, wird im vorgestellten Konzept an zwei Stellen umgegangen. Zum einen versucht die Datenvorverarbeitung (siehe Abschnitt 2.4.1) einen Großteil der fehlenden Werte und der Ausreißer zu beseitigen. Zum anderen werden Ausreißer, die die Datenvorverarbeitung nicht bereinigt, durch eine geschickte Wahl der Clusterparameter im Rauschcluster gruppiert. Es muss an dieser Stelle noch angemerkt werden, dass die mittleren Clusterkurven als Clusterergebnis hinsichtlich ihrer Einsatzmöglichkeiten limitiert sind, da sie einen beschreibenden Charakter haben und nicht als Prognosemodelle eingesetzt werden können. Dazu wird in Abschnitt 2.6 ein neues Konzept zur Modellierung des preisbeeinflussten Verbrauchsverhaltens vorgestellt.

In Tabelle 2.9 ist abschließend zusammengefasst, wie die Fragen auf Seite 61 durch die analyseorientierten Methoden der Pipeline beantwortet werden können.

Tabelle 2.9: Bewertung der analyseorientierten Methoden zur Beantwortung der Fragen auf Seite 61

Frage	Beantwortung
<p>Wie sieht die Reaktion eines Haushaltes oder einer Haushaltsguppe auf ein bestimmtes Preissignal aus? Inwieweit verändern variable Stromtarife ein bekanntes Verbrauchsverhalten, wie z.B. ein SLP?</p>	<p>Visuelle Inspektion und Vergleich der gefundenen Clusterkurven der Kontrollgruppe und der Tarifgruppen; Voraussetzung: Abhängigkeit zwischen den Clustern und den Tarifgruppen → Abschnitte 2.5.1 und 2.5.3</p>
<p>Gibt es saisonale und wochentagsbedingte Schwankungen für eine DR-Maßnahme, z.B. für Fälle mit ungewöhnlichen Preissignalen? (siehe Aufgabenstellung A1 auf Seite 32)</p>	<p>Saisonale und wochentagsbedingte Schwankungen können durch visuelle Inspektion der gefundenen Clusterkurven aufgedeckt werden; Voraussetzung: Abhängigkeit zwischen den Clustern und der entsprechenden Ausgangsgröße („Monat“, „Werktag“) → Abschnitt 2.5.3</p>
<p>Welche prozentualen Energieeinsparungen bzw. zeitlichen Lastverschiebungen ergeben sich durch verschiedene Tarife? Inwiefern rufen verschiedene Stromtarife Unterschiede in der Verbrauchsreaktion hervor? Steigern bestimmte Stromtarife die Unsicherheiten und Schwankungen in der Verbrauchsreaktion? (siehe Aufgabenstellung A2 auf Seite 32)</p>	<ul style="list-style-type: none"> - Erkennung von prozentualen Energieeinsparungen durch Analyse der Fälle ST, SH → Abschnitt 2.5.5; - Erkennung von zeitlichen Lastverschiebungen durch Analyse der Fälle $ST_{normmean}$, $SH_{normmean}$ → Abschnitt 2.5.5; - Erkennung von zeitlichen Lastverschiebungen durch visuelle Inspektion der gefundenen Clusterkurven der Kontrollgruppe und der Tarifgruppen → Abschnitte 2.5.1 und 2.5.3
<p>Reagieren alle Haushalte ähnlich auf bestimmtes Preissignal oder gibt es wesentliche Unterschiede? (siehe Aufgabenstellung A3 auf Seite 32)</p>	<p>Visuelle Inspektion der gefundenen Clusterkurven der Kontrollgruppe und der Tarifgruppen; Voraussetzung: Abhängigkeit zwischen den Clustern und den Tarifgruppen → Abschnitte 2.5.1 und 2.5.3</p>

2.6 Prognoseorientierte Methoden der Pipeline

Im vorigen Abschnitt wurde das neue Konzept zur datengetriebenen Analyse des preisbeeinflussten Verbrauchsverhaltens bei Haushaltsstromkunden vorgestellt. Mithilfe von Clusterverfahren werden typische preisbeeinflusste Lastprofile extrahiert und mit geeigneten Methoden bewertet. Des Weiteren wird durch die Analyse einzelner Abtastzeitpunkte der Tarifeinfluss auf die Stromlast beurteilt. Die Ergebnisse der Verfahren aus Abschnitt 2.5 haben analytischen Charakter und eignen sich aus diesem Grund nicht als Prognosemodell für die Modellierung des preisbeeinflussten Verbrauchsverhaltens von Haushaltsstromkunden.

Deswegen wird in Abschnitt 2.6 ein neues Konzept zur Modellierung des preisbeeinflussten Verbrauchsverhaltens mithilfe von systemtechnischen *Demand-Response*-Modellen vorgestellt. In Abschnitt 2.1 wurde bereits in den Gleichungen (2.4) bis (2.6) gezeigt, dass beim *Demand-Response*-Prozess zwischen der Verbrauchsreaktion $P_{R,n}[k]$ eines einzelnen Haushaltes n , der aggregierten Verbrauchsreaktion $\bar{P}_R^{T_i}[k]$ aller N_{T_i} Haushalte einer Tarifgruppe T_i und der aggregierten Verbrauchsreaktion $\bar{P}_R[k]$ aller N Haushalte unterschieden werden muss. Das zu entwickelnde DR-Modell soll idealerweise in der Lage sein, sowohl die aggregierte Verbrauchsreaktion als auch die haushaltsindividuelle Verbrauchsreaktion auf ein Preissignal zu modellieren.

Ferner soll das Prognosemodell in der Lage sein, die analytischen Ergebnisse der Aufgabenstellung A1 bis A3 zu berücksichtigen. Dies wird dadurch realisiert, dass die Analyseergebnisse aus Abschnitt 2.5 strukturelle Annahmen über das neue DR-Modell zulassen und als *a-priori*-Wissen in die Entwicklung des neuen DR-Modells einfließen (siehe Abschnitt 2.6.1). Falls beispielsweise eine Analyse gemäß Abschnitt 2.5 auf einem *Smart-Meter*-Datensatz ergibt, dass Haushalte eher Lastverschiebungen als reine Lastsenkungen als Verbrauchsreaktion auf ein bekanntes Strompreissignal zeigen, kann das zu entwickelnde DR-Modell auf diese Verbrauchsreaktion zugeschnitten werden. Als wichtige Limitierung muss an dieser Stelle angemerkt werden, dass die Analyseergebnisse aus Abschnitt 2.5 nicht eingesetzt werden, um die Modellparameter des neuen DR-Modells zu schätzen. Die Modellparameter des neuen DR-Modells werden stattdessen aus den *Smart-Meter*-Messdaten (GT-, VT- oder DT-Fälle) mit einem Trust-Region-Verfahren geschätzt, was in Abschnitt 2.6.4 erläutert wird.

Abschnitt 2.6 ist folgendermaßen aufgebaut: In Abschnitt 2.6.1 wird zunächst die Einbeziehung von *a-priori*-Wissen in die Entwicklung des neuen DR-Modells diskutiert. Der Aufbau und die Parameter der neu entwickelten DR-Modelle werden in Abschnitt 2.6.2 vorgestellt. Abschnitt 2.6.3 befasst sich mit der Strategie zur Modellierung des preisbeeinflussten Verbrauchsverhaltens. Dabei wird vorgestellt, wie die beeinflussten und unbeeinflussten Anteile der Verbraucherlast aus der gemessenen Verbraucherlast einer Tarifgruppe gewonnen werden. Die Entwurfs- (Modellselektion, Modellparameterschätzung) und Anwendungsphase (Modellanwendung, Modellbewertung) bei der Modellierung des preisbeeinflussten Verbrauchsverhaltens werden in Abschnitt 2.6.4 und Abschnitt 2.6.5 behandelt. Abschnitt 2.6 schließt mit einer Zusammenfassung und Diskussion der vorgestellten Methoden.

2.6.1 Integration von *a-priori*-Wissen bei der Entwicklung des neuen Demand-Response-Modells

Die Ergebnisse von zahlreichen *Demand-Response*-Feldstudien (eTelligence [64], Intelliekon [65], RESIDENS [78], Modellstadt Mannheim [66], *Olympic Peninsula Project* [73]) und die Ergebnisse der datengetriebenen DR-Analyse [76, 209] lassen zu, dass bestimmte Annahmen über das haushaltsindividuelle bzw. aggregierte Verbrauchsverhalten getroffen werden können. Diese Annahmen fließen als *a-priori*-Wissen in die Entwicklung des neu entwickelten DR-Modells (Vorgriff: das neue DR-Modell trägt den Name „Virtueller Speicher (VS)“) ein und stellen sich wie folgt dar:

- a) Das VS-Modell sollte quantitative Lasterhöhungen und -senkungen als Reaktion auf variable Strompreise abbilden. Diese Modellannahme sollte sowohl für Haushalte gelten, für die der zukünftige Strompreis (z.B. im Fall eines RTP-Tarifs, vgl. Abb. 1.5c) bekannt ist, als auch für Haushalte gelten, für die der zukünftige Strompreis (z.B. im Fall eines TOU-Tarifs, vgl. Abb. 1.5a) unbekannt ist.
- b) Wenn ein Haushalt im Betrachtungszeitraum einen konstanten Strompreis ($p_n[k] = konst.$) erhält, hat der Haushalt keinen Anreiz, Lasten aufgrund des Strompreises zu senken bzw. zu erhöhen (entspricht einem unbeeinflussten Stromverbrauch bzw. einer unbeeinflussten Lastkurve). Das VS-Modell sollte diesem Sachverhalt Rechnung tragen.
- c) Die aggregierte Last $P[k]$ bzw. die Last $P_n[k]$ eines Haushaltes n darf keine negativen Werte annehmen. Es sollte hier allerdings angemerkt werden, dass es im Falle von Haushalten mit lokaler Stromerzeugung durch Photovoltaik, lokalen Erzeugern mit Kraft-Wärme-Kopplung sowie fest oder in Elektrofahrzeugen installierten elektrischen Speichern durchaus zu negativen Lasten und somit zu einer Einspeisung in übergeordnete Netzebenen kommen kann.
Die VS-Modelle, die in der vorliegenden Doktorarbeit vorgestellt werden, gehen mit dieser Feststellung insofern um, als die Laständerung und damit auch die (aggregierte) Last auf $P[k] \geq 0$ bzw. $P_n[k] \geq 0$ begrenzt wird. In zukünftigen Arbeit könnte die Beschränkung aufgehoben werden. Jedoch müssten in diesem Fall im *Smart-Meter*-Datensatz neben der Haushaltslast auch die Leistungswerte der lokalen Erzeugung enthalten sein.
- d) In datengetriebenen DR-Analysen [76, 209] konnte gezeigt werden, dass Haushalte eher Lastverschiebungen als reine Lastsenkungen als Verbrauchsreaktion zeigen. D.h., dass der Stromverbrauch insgesamt fast nicht gesenkt, sondern tageszeitlich verschoben wird. Das VS-Modell sollte daher ein Modell sein, welches Lastverschiebungen abbildet.
- e) Das VS-Modell sollte berücksichtigen, dass sich Lastverlagerungen im Bereich von Haushaltsstromkunden auf einen Zeitraum von höchstens 24 Stunden beschränken [278].
- f) Im *Smart Grid* wird die Partizipation von Haushalten in DR-Programmen als wichtiger Bestandteil gesehen, um das Flexibilitätspotenzial auf der Verbrauchsseite zu heben [25]. In Zukunft ist mit einer Erhöhung des Lastverlagerungspotenzials zu rechnen, da vermehrt intelligente Stromgeräte (vgl. Anhang B.1.3) autonom innerhalb bestimmter Rahmenbedingungen über eine Lastverschiebung entscheiden und das Lastverlagerungspotenzial

in den Nachtstunden nutzen. Das VS-Modell sollte sowohl *Demand Response* als auch *Automated Demand Response* modellieren können.

- g) Um das reale preisbeeinflusste Verbrauchsverhalten abzubilden, sollte das VS-Modell ein *Gray-Box-Modell* mit einer verständlichen inneren Modellstruktur und freien Modellparametern sein, die anhand realer *Smart-Meter*-Daten geschätzt werden können.
- h) Das entworfene VS-Modell sollte in einem umfangreichen Modellvalidierungsverfahren auf seine Gültigkeit überprüft werden. Dazu ist es notwendig, dass in den experimentellen Daten auch Leistungsmesswerte von Haushalten einer genau konzipierten Kontrollgruppe enthalten sind.
- i) In komplexen Simulationsszenarien soll das Systemverhalten in einem rückgekoppelten System unter Verwendung der entwickelten VS-Modelle untersucht werden (siehe Abschnitt 1.4 und Abb. 2.2).

Im Rahmen der Arbeit wurden mehrere Basismodelle entwickelt: das VSA-Modell, das VSA*-Modell, das VSB-Modell und das VSB*-Modell. Der Modellaufbau und die Modellparameter der vier Modelle werden in Abschnitt 2.6.2 beschrieben.

2.6.2 Virtueller Speicher als adaptierbares Verbrauchermodell

Das grundlegende zeitdiskrete Virtuelle-Speicher-Modell besteht aus einem System von Differenzgleichungen, das Lastsenkungen/-erhöhungen als Reaktion auf Preissignale beschreibt. Im VS-Modell werden sowohl Systemdynamiken als auch minimale oder maximale Laständerungen berücksichtigt [279]. Zudem ist im VS-Modell eine rückwärts und vorwärts gerichtete Bewertung des Strompreises in die Modellstruktur integriert. Die Reaktion von Haushaltsstromkunden auf variable Strompreise wurde in Abschnitt 2.5 mithilfe von Data-Mining-Verfahren analysiert. Da die Reaktion der Haushaltsstromkunden auf ein von realen Messdaten abgeleitetes Preissignal (z.B. Daten aus *Olympic Peninsula Project* [73]) eher Lastverschiebungen als Energieeinsparungen zeigt [76], wird diese Tatsache als Einschränkung in noch genaueren VS-Modellen (VSB- und VSB*-Modell) berücksichtigt.

Die Modellbezeichnung „Virtueller Speicher“ setzt sich aus den Begriffen „virtuell“ und „Speicher“ zusammen, wobei der Begriff des Speichers verwendet wird, um die grundlegende Charakteristik der Verbrauchsreaktion auf ein Preissignal zu beschreiben. Lastverschiebungen als Reaktion auf ein Preissignal sind vergleichbar mit dem Be- und Entladen eines realen Speichersystems. Energiemengen, die aufgrund eines hohen Strompreises nicht verbraucht werden, werden aus einem „virtuellen“ Speicher geladen (Speicherentladung) und zu einem anderen Zeitpunkt verbraucht (Speicherbeladung). Der Begriff „virtuell“ soll verdeutlichen, dass der Speicher kein reales physikalisches Speichersystem ist (z.B. ein thermischer Energiespeicher oder eine Batterie).

Vorbemerkung zur mathematischen Beschreibung der VS-Modelle: Alle VS-Modelle in Abschnitt 2.6.2 eignen sich von ihrer Modellstruktur sowohl zur Modellierung des haushaltsindividuellen Verbrauchsverhaltens (siehe Gleichung (2.4)), als auch zur Modellierung des aggregierten Verbrauchsverhaltens einer gesamten Tarifgruppe (siehe Gleichung (2.5)).

Im Fall der Modellierung des haushaltsindividuellen Verbrauchsverhaltens müssten die Modellschätzvariable $\hat{P}_R[k]$, die Modellparameter a_{VS} , $a_{VS,S}[k]$, $E_{VS}[k]$, E_{VS}^{min} , E_{VS}^s , E_{VS}^{max} und die Strompreisvariablen $p[k]$, p_f , p_p korrekterweise mit dem tiefgestellten Haushaltsindex n versehen sein.

Im Fall der Modellierung des aggregierten Verbrauchsverhaltens aller N_{T_i} Haushalte der Tarifgruppe T_i müssten die Modellparameter a_{VS} , $a_{VS,S}[k]$, $E_{VS}[k]$, E_{VS}^{min} , E_{VS}^s , E_{VS}^{max} und die Strompreisvariablen $p[k]$, p_f , p_p korrekterweise mit dem hochgestellten Tarifgruppenindex T_i versehen sein. Zudem müsste die Modellschätzvariable $\hat{P}_R[k]$ korrekterweise mit einem Überstrich und dem hochgestellten Tarifgruppenindex T_i gemäß $\hat{\overline{P}}_R^{T_i}[k]$ geschrieben werden. Aus Gründen der Übersichtlichkeit wird in der mathematischen Beschreibung aller VS-Modelle auf die beiden Indizes und auf den Überstrich verzichtet.

VSA-Modell: Es sei mit $p[k]$ der aktuelle Strompreis und mit p_f der bekannte zukünftige Strompreis (vgl. Tabelle 2.10) gegeben. Die Schätzung des beeinflussten Anteils der Verbraucherlast $\hat{P}_R[k]$ (vgl. Gleichung (2.4)) berechnet sich wie folgt:

$$\hat{P}_R[k] = \begin{cases} a_{VS} \cdot (E_{VS}^{min} - E_{VS}[k]) & , \text{ für } p[k] > p_f \\ a_{VS} \cdot (E_{VS}^s - E_{VS}[k]) & , \text{ für } p[k] = p_f \\ a_{VS} \cdot (E_{VS}^{max} - E_{VS}[k]) & , \text{ für } p[k] < p_f \end{cases} \quad (2.60)$$

wobei a_{VS} die Speicherbe- bzw. -entladegeschwindigkeit, $E_{VS}[k]$ der Speicherfüllstand zum Zeitpunkt k , E_{VS}^{min} der minimale Füllstand des virtuellen Speichers, E_{VS}^{max} der maximale Füllstand des virtuellen Speichers, und E_{VS}^s der Sollwert des Speicherfüllstands des virtuellen Speichers ist. Die Höhe des Sollwertes wird so festgelegt, dass sie zeitkonstant ist und zwischen E_{VS}^{min} und E_{VS}^{max} liegt. Grundsätzlich kann jeder Füllstand für E_{VS}^s mit $E_{VS}^s \in (E_{VS}^{min}, E_{VS}^{max})$ gewählt werden.

Gleichung (2.60) kann nicht zur Modellierung des preisbeeinflussten Verbrauchsverhaltens verwendet werden, wenn die Höhe des zukünftigen Strompreises, wie im Fall eines RTP-Tarifs, unbekannt ist. Statt einer vorwärts gerichteten Preisbewertung muss hierfür eine rückwärts gerichtete Bewertung des Strompreises vorgenommen werden. Die Prognose des preisbeeinflussten Anteils der Verbraucherlast $\hat{\overline{P}}_R[k]$ stellt sich dann folgendermaßen dar:

$$\hat{\overline{P}}_R[k] = \begin{cases} a_{VS} \cdot (E_{VS}^{min} - E_{VS}[k]) & , \text{ für } p_p < p[k] \\ a_{VS} \cdot (E_{VS}^s - E_{VS}[k]) & , \text{ für } p_p = p[k] \\ a_{VS} \cdot (E_{VS}^{max} - E_{VS}[k]) & , \text{ für } p_p > p[k] \end{cases} \quad (2.61)$$

wobei p_p der vergangene Strompreis ist. Für die Berechnung der Strompreise p_f und p_p stehen die in Tabelle 2.10 gezeigten Ansätze zur Verfügung. Darüber hinaus existieren noch weitere

Berechnungsvorschriften für den vergangenen oder zukünftigen Strompreis, die in [147, 159, 280] aufgeführt sind.

Wie bereits in Modellannahme c) auf Seite 92 beschrieben, darf die aggregierte Last $P[k]$ bzw. die Last $P_n[k]$ eines Haushaltes n keine negativen Werte annehmen. Zusammen mit den Gleichungen (2.7) und (2.8) gilt folgende Einschränkung:

$$\hat{P}_R[k] \geq -P_U[k], \quad (2.62)$$

$$\hat{P}_{R,n}[k] \geq -P_{U,n}[k]. \quad (2.63)$$

Eine Lasterhöhung/-reduzierung als Reaktion auf ein Preissignal führt dazu, dass sich der Füllstand des Virtuellen Speichers gemäß

$$E_{VS}[k+1] = E_{VS}[k] + \hat{P}_R[k] \cdot T_S \quad (2.64)$$

ändert. Das grundlegende Verhalten des Speicherfüllstandes lässt sich abschließend folgendermaßen beschreiben:

- Das Einschalten von elektrischen Verbrauchern bzw. die Erhöhung der Last ($\hat{P}_R[k] > 0$) bewirkt das Füllen des virtuellen Speichers ($E_{VS}[k] < E_{VS}[k+1]$).
- Keine Änderung des Verbrauchs bzw. der Last ($\hat{P}_R[k] = 0$) hat auch keine Änderung des Speicherfüllstandes zur Folge ($E_{VS}[k] = E_{VS}[k+1]$).
- Das Abschalten elektrischer Verbraucher bzw. die Lastabsenkung ($\hat{P}_R[k] < 0$) führt zum Entleeren des virtuellen Speichers ($E_{VS}[k] > E_{VS}[k+1]$).

VSA*-Modell: Das VSA*-Modell erweitert das grundlegende VSA-Modell insofern, als ein zusätzlicher Differenzterme zu den Gleichungen (2.60) und (2.61) hinzugefügt wird, um eine graduelle Preisbewertung zu ermöglichen. Die Schätzung des beeinflussten Anteils der Verbraucherlast $\hat{P}_R[k]$ für bekannte zukünftige Strompreise (Gleichung (2.65)) und unbekannte zukünftige Strompreise (Gleichung (2.66)) berechnet sich folgendermaßen:

$$\hat{P}_R[k] = \begin{cases} a_{VS} \cdot |p[k] - p_f| \cdot (E_{VS}^{min} - E_{VS}[k]) & , \text{ für } p[k] > p_f \\ a_{VS} \cdot |p[k] - p_f| \cdot (E_{VS}^s - E_{VS}[k]) & , \text{ für } p[k] = p_f \\ a_{VS} \cdot |p[k] - p_f| \cdot (E_{VS}^{max} - E_{VS}[k]) & , \text{ für } p[k] < p_f \end{cases} \quad (2.65)$$

$$\hat{P}_R[k] = \begin{cases} a_{VS} \cdot |p[k] - p_p| \cdot (E_{VS}^{min} - E_{VS}[k]) & , \text{ für } p_p < p[k] \\ a_{VS} \cdot |p[k] - p_p| \cdot (E_{VS}^s - E_{VS}[k]) & , \text{ für } p_p = p[k] \\ a_{VS} \cdot |p[k] - p_p| \cdot (E_{VS}^{max} - E_{VS}[k]) & , \text{ für } p_p > p[k] \end{cases} \quad (2.66)$$

VSB-Modell: Im Vergleich zum VSA- bzw. VSA*-Modell kann das VSB-Modell nur dann eingesetzt werden, wenn der vergangene Strompreis bekannt ist. Eine Preisbewertung für bekannte zukünftige Strompreise wie in Gleichungen (2.60) und (2.65) findet hier nicht statt. Im VSB-Modell wird explizit die Modellannahme e) auf Seite 92 berücksichtigt, indem eine zusätzlich Speichervariable - die Speicherausgleichsfunktion $a_{VS,S}[k]$ - eingeführt wird. Die Schätzung des preisbeeinflussten Anteils der Verbraucherlast $\hat{P}_R[k]$ ist definiert als

$$\hat{P}_R[k] = \begin{cases} a_{VS} \cdot (1 - a_{VS,S}[k]) \cdot (E_{VS}^{min} - E_{VS}[k]) + \\ \quad a_{VS,S}[k] \cdot (E_{VS}^s - E_{VS}[k]) & , \text{ für } p_p < p[k] \\ a_{VS} \cdot (1 - a_{VS,S}[k]) \cdot (E_{VS}^s - E_{VS}[k]) + \\ \quad a_{VS,S}[k] \cdot (E_{VS}^s - E_{VS}[k]) & , \text{ für } p_p = p[k] \\ a_{VS} \cdot (1 - a_{VS,S}[k]) \cdot (E_{VS}^{max} - E_{VS}[k]) + \\ \quad a_{VS,S}[k] \cdot (E_{VS}^s - E_{VS}[k]) & , \text{ für } p_p > p[k] \end{cases} \quad (2.67)$$

Die Speicherausgleichsfunktion $a_{VS,S}[k]$ sorgt dafür, dass sich der Füllstand des Virtuellen Speichers im Laufe des Tages dem Sollfüllstand E_{VS}^s annähert. Die Annäherung wird durch zwei Varianten von $a_{VS,S}[k]$ erreicht. Zum einen ist das die lineare Speicherausgleichsfunktion $a_{VS,S}^{lin}[k]$, die folgendermaßen definiert ist:

$$a_{VS,S}^{lin}[k] = \frac{1}{K_s} \cdot k, \quad \text{mit } k = 1, \dots, K_s, \quad (2.68)$$

wobei K_s die Anzahl an Abtastzeitpunkten pro Tag ist. Als zweite Variante steht die exponentielle Speicherausgleichsfunktion $a_{VS,S}^{exp}[k]$ zur Verfügung, die wie folgt beschrieben ist:

$$a_{VS,S}^{exp}[k] = \beta_{VS} \cdot e^{\gamma_{VS} \cdot k}, \quad \text{mit } k = 1, \dots, K_s. \quad (2.69)$$

β_{VS} und γ_{VS} stellen in Gleichung (2.69) die Parameter der Exponentialfunktion dar und werden mithilfe der drei Stützstellen ($k = 1, a_{VS,S}^{exp} \approx 0$)¹⁵, ($k = \frac{K_s}{2}, a_{VS,S}^{exp} = 0.1$) und ($k = K_s, a_{VS,S}^{exp} = 1$) geschätzt, was zu $\beta_{VS} = 0.0096$ und $\gamma_{VS} = 0.0484$ führt. Die lineare und exponentielle Speicherausgleichsfunktion sind in Abb. 2.22 dargestellt und müssen *a priori* festgelegt werden.

In Ergänzung zu den Preisberechnungsvorschriften für die VSA- und VSA*-Modelle (siehe Tabelle 2.10) steht für das VSB-Modell die Bewertung des vergangenen Strompreises p_p mithilfe des exponentiell geglätteten Mittelwerts (engl. *exponential moving average*, EMA [281]) zur Verfügung. Dabei wird aus der Strompreiszeitreihe $p[k]$ die exponentiell geglättete Strompreiszeitreihe $\hat{p}_{EMA}[k]$ geschätzt:

$$\begin{aligned} \hat{p}_{EMA}[k] &= \alpha_{EMA} \cdot p[k] + (1 - \alpha_{EMA}) \cdot \hat{p}_{EMA}[k - 1], \\ &\text{mit } k = 2, \dots, K \\ &\text{und } \hat{p}_{EMA}[1] = p[1]. \end{aligned} \quad (2.70)$$

$\alpha_{EMA} \in [0, 1]$ ist der Glättungsfaktor. Für $\alpha_{EMA} = 1$ findet keine Glättung statt und der Schätzwert $\hat{p}_{EMA}[k]$ ist gleich dem Messwert $p[k]$. Für $\alpha_{EMA} = 0$ ändert sich der Schätzwert $\hat{p}_{EMA}[k]$ nicht. Durch die Verwendung des EMAs werden im Vergleich zum SMA jüngere Abtastzeitpunkte noch stärker gewichtet als weiter zurückliegende.

¹⁵Der Wert $a_{VS,S}^{exp} = 0$ ist nur approximativ erreichbar, da β_{VS} sonst den Wert Null besitzen würde.

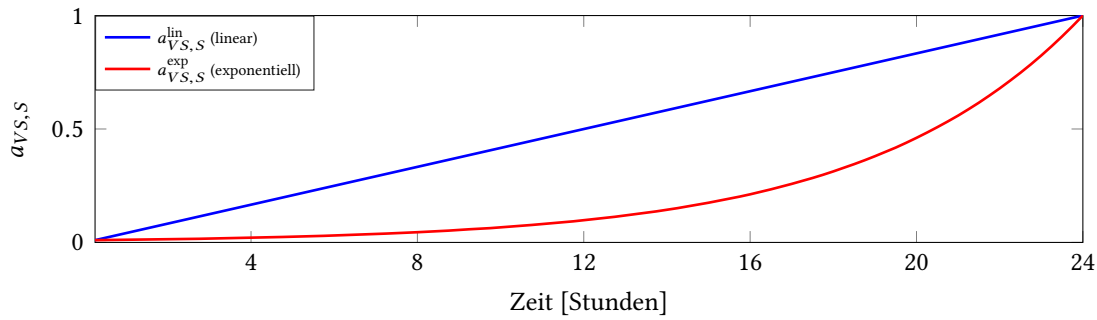


Abbildung 2.22: Lineare und exponentielle Speicherausgleichsfunktion

VS^B*-Modell: Das VS^B*-Modell hat, analog zum VS^A*-Modell, eine graduelle Preisbewertung. Aus Gleichung (2.67) folgt somit

$$\hat{P}_R[k] = \begin{cases} a_{VS} \cdot |p[k] - p_p| \cdot (1 - a_{VS,S}[k]) \cdot (E_{VS}^{min} - E_{VS}[k]) + \\ \quad a_{VS,S}[k] \cdot (E_{VS}^s - E_{VS}[k]) & , \text{ für } p_p < p[k] \\ a_{VS} \cdot |p[k] - p_p| \cdot (1 - a_{VS,S}[k]) \cdot (E_{VS}^s - E_{VS}[k]) + \\ \quad a_{VS,S}[k] \cdot (E_{VS}^s - E_{VS}[k]) & , \text{ für } p_p = p[k] \\ a_{VS} \cdot |p[k] - p_p| \cdot (1 - a_{VS,S}[k]) \cdot (E_{VS}^{max} - E_{VS}[k]) + \\ \quad a_{VS,S}[k] \cdot (E_{VS}^s - E_{VS}[k]) & , \text{ für } p_p > p[k] \end{cases} \quad (2.71)$$

Die Begrenzung von $\hat{P}_R[k]$ in den Gleichungen (2.62) und (2.63) und der Zusammenhang von $\hat{P}_R[k]$ und $E_{VS}[k]$ in Gleichung (2.64) gelten für alle VS-Modelle. Die zuvor eingeführten VS-Modelle und deren Modellspezifikationen sind abschließend in Tabelle 2.10 zusammengefasst. Des Weiteren verdeutlicht Beispiel 10 das Modellverhalten verschiedener VS-Modelle.

Inverses VS-Modell: Der beeinflusste Verbrauchsprozess setzt sich aus dem unbeeinflussten Verbrauchsprozess und dem *Demand-Response*-Prozess zusammen (siehe Abb. 2.2). Das VS-Modell kommt dann zum Einsatz, wenn der unbeeinflusste Anteil der Verbraucherlast $P_U[k]$ gegeben und die beeinflusste Verbraucherlast $P[k]$ geschätzt werden soll:

$$\hat{P}[k] = P_U[k] + \underbrace{\hat{P}_R[k]}_{\text{VS-Modell}} . \quad (2.72)$$

$\hat{P}_R[k]$ ist dabei die Modellschätzung des beeinflussten Anteils der Verbraucherlast $P_R[k]$ und wird je nach VS-Modell entsprechend den Gleichungen (2.60), (2.61), (2.65) bis (2.67) und (2.71) berechnet.

Das inverse VS-Modell wird verwendet, wenn die beeinflusste Verbraucherlast $P[k]$ gegeben und der unbeeinflusste Anteil der Verbraucherlast $P_U[k]$ prognostiziert werden soll:

$$\hat{P}_U[k] = P[k] + \underbrace{(-\hat{P}_R[k])}_{\text{Inverses VS-Modell}} . \quad (2.73)$$

Tabelle 2.10: Berechnungsvorschriften für die vorwärts und rückwärts gerichtete Bewertung des Strompreises für Virtuelle-Speicher-Modelle: H - Betrachtungshorizont des Strompreises, Index „ val “ - englisch *value* für einzelner Wert, Index „ med “ - Median, Index „ SMA “ - Abkürzung für *simple moving average*, Index „ EMA “ - Abkürzung für *exponential moving average*.

Basismodell	Spezifisches VS-Modell	Formel zur Berechnung von p_f bzw. p_p	Speicherausgleichsfunktion a_S
VSA-Modell Gleichungen (2.60) und (2.61)	VSA_{val}^f	$p_f = p[k + H]$	–
	VSA_{val}^p	$p_p = p[k - H]$	–
	VSA_{med}^f	$p_f = \text{median}(p[k + 1], \dots, p[k + H])$	–
	VSA_{med}^p	$p_p = \text{median}(p[k - H], \dots, p[k - 1])$	–
	VSA_{SMA}^f	$p_f = \text{mean}(p[k + 1], \dots, p[k + H])$	–
	VSA_{SMA}^p	$p_p = \text{mean}(p[k - H], \dots, p[k - 1])$	–
VSA*-Modell Gleichungen (2.65) und (2.66)	VSA_{val}^{*f}	$p_f = p[k + H]$	–
	VSA_{val}^{*p}	$p_p = p[k - H]$	–
	VSA_{med}^{*f}	$p_f = \text{median}(p[k + 1], \dots, p[k + H])$	–
	VSA_{med}^{*p}	$p_p = \text{median}(p[k - H], \dots, p[k - 1])$	–
	VSA_{SMA}^{*f}	$p_f = \text{mean}(p[k + 1], \dots, p[k + H])$	–
	VSA_{SMA}^{*p}	$p_p = \text{mean}(p[k - H], \dots, p[k - 1])$	–
VS_B-Modell Gleichung (2.67)	VS_{SMA}^{lin}	$p_p = \text{mean}(p[k - H], \dots, p[k - 1])$	$a_{VS,S} \rightarrow \text{linear}$
	VS_{SMA}^{exp}	$p_p = \text{mean}(p[k - H], \dots, p[k - 1])$	$a_{VS,S} \rightarrow \text{exponentiell}$
	VS_{EMA}^{lin}	$p_p = \hat{p}_{EMA}[k - 1]$	$a_{VS,S} \rightarrow \text{linear}$
	VS_{EMA}^{exp}	$p_p = \hat{p}_{EMA}[k - 1]$	$a_{VS,S} \rightarrow \text{exponentiell}$
VS_B*-Modell Gleichung (2.71)	VS_{SMA}^{*lin}	$p_p = \text{mean}(p[k - H], \dots, p[k - 1])$	$a_{VS,S} \rightarrow \text{linear}$
	VS_{SMA}^{*exp}	$p_p = \text{mean}(p[k - H], \dots, p[k - 1])$	$a_{VS,S} \rightarrow \text{exponentiell}$
	VS_{EMA}^{*lin}	$p_p = \hat{p}_{EMA}[k - 1]$	$a_{VS,S} \rightarrow \text{linear}$
	VS_{EMA}^{*exp}	$p_p = \hat{p}_{EMA}[k - 1]$	$a_{VS,S} \rightarrow \text{exponentiell}$

Wie in den Gleichungen (2.72) und (2.73) zu sehen, unterscheiden sich das VS-Modell und das inverse VS-Modell lediglich durch das negative Vorzeichen. Außerdem haben das VS-Modell und das inverse VS-Modell die gleiche Modellstruktur.

Beispiel 10: Modellverhalten bei synthetischen Strompreissignalen

Um das Modellverhalten verschiedener VS-Modelle zu verdeutlichen, beschäftigt sich dieses Beispiel mit dem Verhalten der VS-Modelle als Reaktion auf zwei synthetische Preissignale: ein zweistufiger TOU-Tarif (Abb. 2.23(a)) und ein RTP-Tarif (Abb. 2.23(f)) mit beliebig vielen Tarifstufen. Die Abb. 2.23(b)-(e) stellen das Modellverhalten als Reaktion auf den zweistufigen TOU-Tarif dar, während die Abb. 2.23(g)-(j) das Modellverhalten als Reaktion auf den RTP-Tarif zeigen.

Im Fall des TOU-Tarifs haben die Haushaltsstromkunden im Voraus Kenntnis über die Höhe des Strompreises, so dass eine vorwärts gerichtete Preisbewertung für das VSA-Modell verwendet werden kann, wie in Abb. 2.23(b) gezeigt. Die Reaktion des Haushaltes wird als Lastanstieg wenige Stunden vor dem Hochtarif modelliert, gefolgt von einer Lastreduktion zu Zeiten des Hochtarifs und einem geringen Lastanstieg nach dem Hochtarif. Der Vollständigkeit halber ist in Abb. 2.23(c) der geschätzte beeinflusste Anteil der Verbraucherlast unter Verwendung des rückwärts gerichteten VSA-Modells dargestellt. Für alle VSB-Modelle gilt die Einschränkung, dass Laständerungen auf einen Zeitraum von maximal 24 Stunden beschränkt sind. Diese Modelleinschränkung ist in Abb. 2.23(d) und (e) als abgeschwächte Laständerung als Reaktion auf den Hochtarif am Abend zu erkennen. Der Füllstand der VSB-Modelle soll am Ende des Tages den Sollwert des Speicherfüllstands einnehmen, was zu einer reinen Lastverschiebung ohne Änderung des täglichen Energieverbrauchs führt.

Im Fall des RTP-Tarifs in Abb. 2.23(f) ist der Strompreis im Voraus für die Hauskunden unbekannt, so dass die vorwärts gerichteten VSA-Modelle (VSA^f_{SMA}) ungeeignet sind. Aus diesem Grund ist in Abb. 2.23 nur das rückwärts gerichtete VSA-Modell dargestellt. Die Modellschätzungen der VSB-Modelle mit der SMA- und EMA-Preisbewertung sind in den Abb. 2.23(h)-(j) dargestellt.

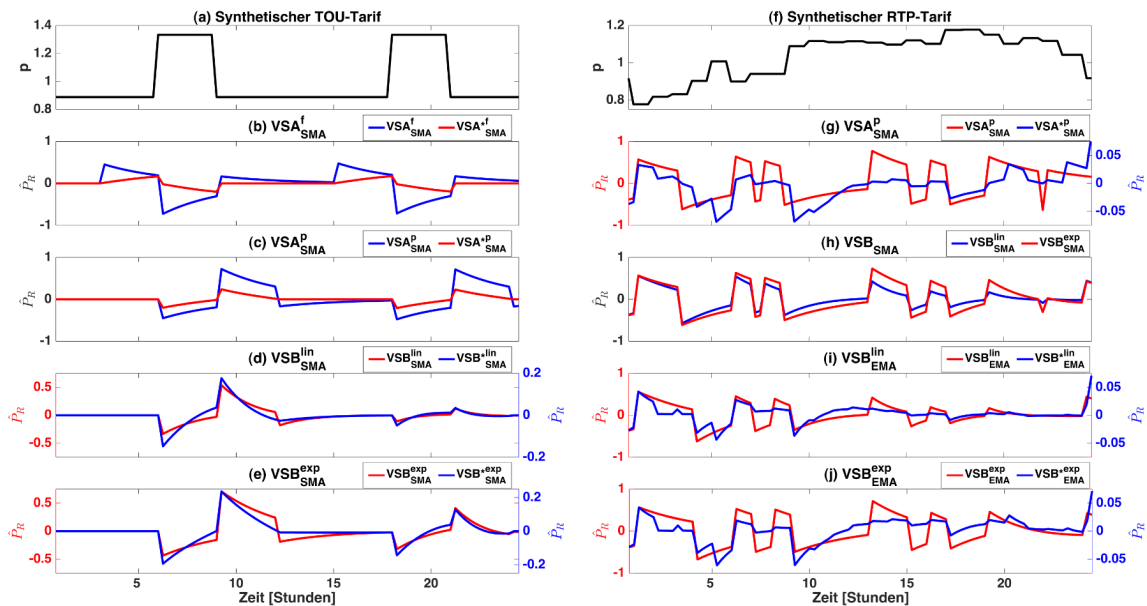


Abbildung 2.23: Modellverhalten verschiedener Virtueller-Speicher-Modelle als Reaktion auf synthetische Strompreise: (a) Synth. TOU-Tarif, (b)-(e) Modellverhalten als Reaktion auf synth. TOU-Tarif aus (a); (f) synth. RTP-Tarif, (g)-(j) Modellverhalten als Reaktion auf synth. RTP-Tarif aus (f). Die blauen Linien in (b), (c) und (h) haben den gleichen Wertebereich, daher ist nur eine y-Achsenbeschriftung vorhanden. Die blauen und roten Linien in (d), (e), (g), (i) und (j) haben einen unterschiedlichen Wertebereich, daher sind zwei y-Achsenbeschriftungen auf der linken und rechten Seite vorhanden.

2.6.3 Strategie zur Berechnung des beeinflussten Anteils der Verbraucherlast

Die Abschnitte 2.6.4 und 2.6.5 befassen sich mit der Modellierung des preisbeeinflussten Verbrauchsverhaltens. Oberstes Ziel des neuen Modellierungskonzeptes ist die Entwicklung eines DR-Modells, welches die Zusammenhänge in den Gleichungen (2.4) bis (2.6) modelliert. Allerdings ist der zu schätzende beeinflusste Anteil der Verbraucherlast P_R nicht direkt messbar, und

somit auch nicht als Zeitreihe in einem *Smart-Meter*-Datensatz vorhanden. In Abschnitt 2.6.3 wird deshalb vorgestellt, wie die beeinflussten und unbeeinflussten Anteile der Verbraucherlast aus der gemessenen Verbraucherlast gewonnen werden.

Eine erste Abschätzung des preisbeeinflussten Verbrauchsverhaltens einer Tarifgruppe wurde bereits kurz in Abschnitt 2.4.4 gemacht. Um den Einfluss des Preissignals bei einer preisbeeinflussten Verbrauchergruppe T_i (z.B. TOU-Gruppe oder RTP-Gruppe) zu ermitteln, ist das Vorhandensein einer Kontrollgruppe (FIXED) ohne Preiseinfluss im *Smart-Meter*-Datensatz erforderlich. Aus Gleichung (2.8) folgt, dass die beeinflusste Verbraucherlast einer Kontrollgruppe wie folgt definiert ist:

$$\bar{P}^{\text{FIXED}}[k] = \bar{P}_U^{\text{FIXED}}[k] + \underbrace{\bar{P}_R^{\text{FIXED}}[k]}_{=0} = \bar{P}_U^{\text{FIXED}}[k], \quad (2.74)$$

wobei $\bar{P}_U^{\text{FIXED}}[k]$ der unbeeinflusste Anteil der aggregierten Verbraucherlast und $\bar{P}_R^{\text{FIXED}}[k]$ der beeinflusste Anteil der aggregierten Verbraucherlast der Kontrollgruppe ist. Da die Haushalte der Kontrollgruppe keine Preisanreize durch variable Strompreise erfahren, ist $\bar{P}_R^{\text{FIXED}}[k]$ gleich Null. Im Fall einer preisbeeinflussten Verbrauchergruppe T_i (z.B. TOU-Gruppe oder RTP-Gruppe) gilt für die beeinflusste Verbraucherlast Folgendes:

$$\bar{P}^{T_i}[k] = \bar{P}_U^{T_i}[k] + \bar{P}_R^{T_i}[k], \quad (2.75)$$

wobei $\bar{P}_U^{T_i}[k]$ der unbeeinflusste Anteil der aggregierten Verbraucherlast und $\bar{P}_R^{T_i}[k]$ der beeinflusste Anteil der aggregierten Verbraucherlast der Tarifgruppe T_i ist.

Gemäß Gleichung (2.5) ist $\bar{P}_R^{T_i}[k]$ der beeinflusste Anteil der Verbraucherlast, den es durch das DR-Modell vorherzusagen gilt. Da allerdings nur die beeinflusste Verbraucherlast $\bar{P}^{T_i}[k]$ gemessen werden kann, muss $\bar{P}_R^{T_i}[k]$ berechnet werden. Durch Umstellen von Gleichung (2.75) nach $\bar{P}_R^{T_i}[k]$ und mit der Annahme in Gleichung (2.76), dass die unbeeinflussten Anteile der Kontroll- und Tarifgruppe gleich sind,

$$\bar{P}_U^{T_i}[k] = \bar{P}_U^{\text{FIXED}}[k] \quad (2.76)$$

kann der beeinflusste Anteil der Verbraucherlast wie folgt berechnet werden:

$$\bar{P}_R^{T_i}[k] = \bar{P}^{T_i}[k] - \bar{P}_U^{\text{FIXED}}[k]. \quad (2.77)$$

Für die Modellvalidierung wäre eine Zusammensetzung der Kontrollgruppe und der preisbeeinflussten Verbrauchergruppe aus gleichen Haushalten von Vorteil, da in diesem Fall die Annahme in Gleichung (2.76) zutrifft (siehe Eigenschaft E11 in Tabelle 2.2). Allerdings ist dieser Idealfall in den meisten DR-Feldstudien nicht gegeben. Aufgrund einer heterogenen Zusammensetzung der Gruppen (z.B. unterschiedlicher Tagesabläufe für Haushalte innerhalb einer Gruppe) und den damit einhergehenden großen Abweichungen in den täglichen Lastprofilen wird die Annahme in Gleichung (2.76) verletzt, was zu einer höheren Modellierungsunsicherheit führt.

2.6.4 Modellentwurf

Die Abschnitte 2.6.4 und 2.6.5 befassen sich mit dem Ablauf der Modellierung des preisbeeinflussten Verbrauchsverhaltens, der zudem in Abb. 2.24 skizziert ist. Das Ziel bei der Modellierung des preisbeeinflussten Verbrauchsverhaltens ist, das Verhalten realer Haushaltsstromkunden als Reaktion auf Preissignale möglichst genau zu schätzen. In Abschnitt 2.6.2 wurden dafür zahlreiche Modellstrukturen des VS-Modells (VSA-, VSA*-, VSB- und VSB*-Modell) vorgestellt, deren Modellparameter anhand realer *Smart-Meter*-Daten geschätzt werden können. Die Herausforderung besteht nun darin, je nach *Smart-Meter*-Datensatz aus der Vielzahl an VS-Modellen das beste Modell auszuwählen und dieses mit etablierten dynamischen Regressionsmodellen (siehe Gleichung (1.4) auf Seite 19) hinsichtlich Genauigkeit und Generalisierungsfähigkeit zu vergleichen.

Ausgehend von der Problemstellung (siehe Aufgabenstellung A4 auf Seite 32) lässt sich die Modellierung in zwei Phasen einteilen: in die **Modellentwurfsphase** und die **Modellanwendungsphase**. In Abschnitt 2.6.4 wird zunächst auf die Aufteilung des *Smart-Meter*-Datensatzes in Lern- und Testdaten eingegangen. Des Weiteren werden die einzelnen Prozesse des Modellentwurfs genauer erläutert.

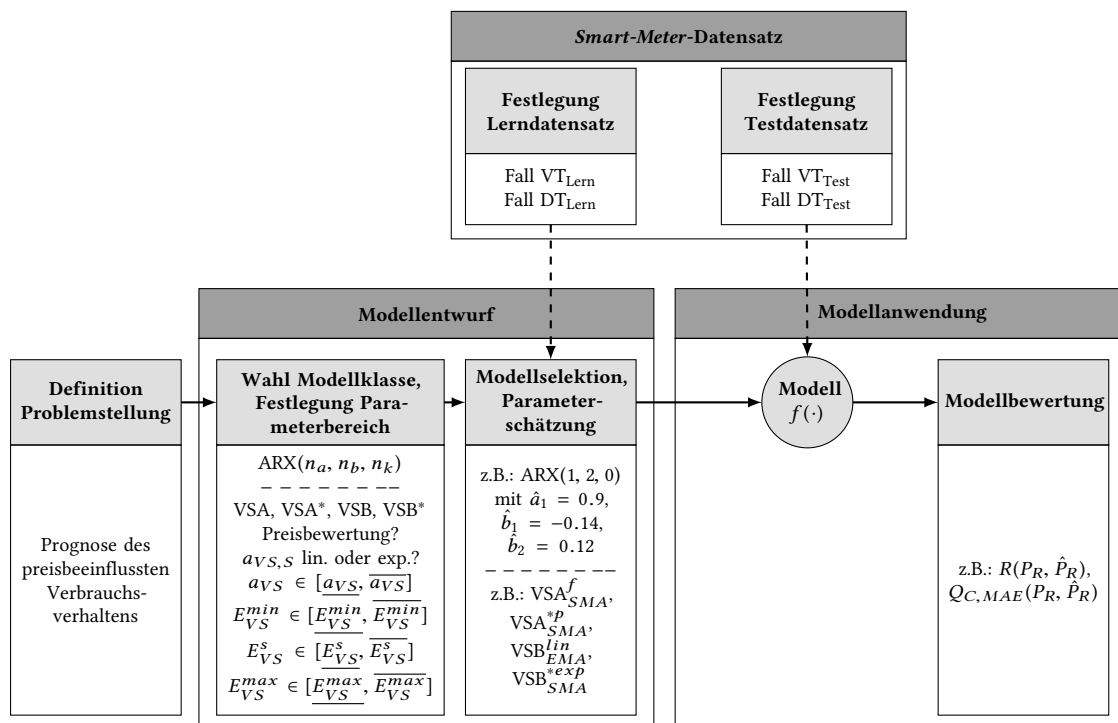


Abbildung 2.24: Entwurfs- und -anwendungsphase bei der Modellierung des preisbeeinflussten Verbrauchsverhaltens: DT_{Lern} - DT-Datensatz zum Lernen der Modelle, DT_{Test} - DT-Datensatz zum Testen der Modelle, $\{n_a, n_b\}$ - Modellordnungen des ARX-Modells, n_k - Totzeit des ARX-Modells, $\{\hat{a}_1, \hat{b}_1, \hat{b}_2\}$ - Schätzparameter des $ARX(1, 2, 0)$ -Modells, $Q_{C,MAE}$ - mittlerer absoluter Fehler, R - Korrelationskoeffizient, VT_{Lern} - VT-Datensatz zum Lernen der Modelle, VT_{Test} - VT-Datensatz zum Testen der Modelle.

Festlegung des Lern- und Testdatensatzes: Grundsätzlich kann die Modellierung des preisbeeinflussten Verbrauchsverhaltens anhand verschiedener Datensätze und Aggregationsniveaus durchgeführt werden. Um jedoch die Generalisierungsfähigkeit eines datengetriebenen Modells bewerten zu können, muss eine geeignete Aufteilung des Datensatzes in Lerndatensatz für den Modellentwurf und Testdatensatz für die Modellanwendung gewählt werden. Ein gutes und universell einsetzbares Verfahren hierfür ist die Kreuzvalidierung (engl.: *cross validation*) [85], bei der der Datensatz mehrfach in zwei Mengen aufgeteilt wird. Eine Menge fungiert als Lerndatensatz zum Schätzen der Modellparameter, die zweite Menge dient als Testdatensatz zur Bewertung des entworfenen Modells. Die Aufteilung in Lern- und Testdatensatz wird so oft wiederholt, bis jede Menge einmal als Testdatensatz fungiert. So wird z.B. bei einer vierfachen Kreuzvalidierung der gesamte Datensatz in vier gleich große Teildatensätze unterteilt, wobei in vier Durchläufen jeweils ein anderer Teildatensatz als Testdatensatz und die restlichen drei Teildatensätze als Lerndatensatz fungieren.

Der Vergleich zwischen der Modellschätzung \hat{P}_R und der zu schätzenden Variablen P_R sollte mit aggregierten Zeitreihen durchgeführt werden, idealerweise mit mehreren 100 Haushalten je Gruppe, um haushaltsindividuelle Variationen zu reduzieren. Die Aggregation von Einzelhaushalten, die der selben Tarifgruppe T_i angehören, ist in Tabelle 2.3 gezeigt (siehe Fall GT, GT_{norm} , VT, VT_{norm} , DT und DT_{norm}). Wenn jedoch die Anzahl an Datentupeln sehr klein ist (z.B. für die Fälle GT und GT_{norm}) oder Haushalte sich nicht ausschließlich in der entsprechenden Tarifstruktur unterscheiden (z.B. durch systematische Effekte bei der Tarifauswahl aufgrund fehlender Randomisierung, die zu unterschiedlichen Lastprofilen auch für denselben Tarif führen können), sind moderate Modellierungsergebnisse hinsichtlich Genauigkeit und Generalisierungsfähigkeit zu erwarten. Der Modellentwurf und die Modellanwendung wird daher mit aggregierten Daten (Fälle VT, VT_{norm} , DT und DT_{norm}) realisiert.

Aufgrund von unterschiedlichen Lastverläufen in unterschiedlichen Jahreszeiten, an Werktagen und Wochenenden ist es zudem angebracht, weitere Subgruppen für den Fall DT einzuführen, wie z.B. der Fall $DT_{\text{summer}}^{\text{we}}$ für alle Wochenendtage im Sommer oder der Fall $DT_{\text{winter}}^{\text{wo}}$ für alle Werktage im Winter. Die Einführung weiterer Subgruppen erlaubt zudem, dass eine Subgruppe als Lerndatensatz und die anderen Subgruppen als Testdatensätze fungieren. Die Einteilung der DT- bzw. VT-Fälle in Subgruppen ist jedoch nicht für alle Zeiten festgelegt und sollte für jeden *Smart-Meter*-Datensatz individuell bestimmt werden. Mit dem höheren Aggregationsniveau der DT-Fälle und der damit verbundenen höheren Anzahl an Datentupeln sind deutlich verbesserte Modellierungsergebnisse im Vergleich zu den VT-Fällen zu erwarten. Es wird an dieser Stelle davon ausgegangen, dass die Länge der Zeitreihen der VT-Fälle im Bereich von sechs bis zehn Wochen liegt.

Wahl der Modellklasse und Festlegung der Parameterbereiche: Wie bereits zu Beginn von Abschnitt 2.6.4 beschrieben, werden für die Modellierung des preisbeeinflussten Verbrauchsverhaltens zwei Modellklassen eingesetzt. Es handelt sich dabei zum einen um ein $ARX(n_a, n_b, n_k)$ -Modell, welches bereits in Gleichung (1.4) vorgestellt wurde. Als zweite Modellklasse kommt das in Abschnitt 2.6.2 vorgestellte VS-Modell¹⁶ zum Einsatz. Im Fall des VS-Modells

¹⁶Die vier Basismodelle VSA, VSA*, VSB und VSB* sind als eine Modellklasse zusammengefasst.

müssen zudem die Wertebereiche für die Modellparameter a_{VS} , E_{VS}^{min} , E_{VS}^s und E_{VS}^{max} in Form von unteren und oberen Schranken festgelegt werden:

$$a_{VS} \in [\underline{a_{VS}}, \overline{a_{VS}}], \quad (2.78)$$

$$E_{VS}^{min} \in [\underline{E_{VS}^{min}}, \overline{E_{VS}^{min}}], \quad (2.79)$$

$$E_{VS}^s \in [\underline{E_{VS}^s}, \overline{E_{VS}^s}] \text{ und} \quad (2.80)$$

$$E_{VS}^{max} \in [\underline{E_{VS}^{max}}, \overline{E_{VS}^{max}}]. \quad (2.81)$$

Die untere Schranke ist jeweils mit einem langen Unterstrich, die obere Schranke mit einem langen Überstrich gekennzeichnet. Zudem muss je nach Stromtarif definiert werden, nach welcher Berechnungsvorschrift (siehe Spalte 3 in Tabelle 2.10) die Strompreise p_f und p_p berechnet werden sollen. Speziell für die VSB- und VSB*-Modelle gilt in diesem Schritt noch festzulegen, ob die Speicherausgleichsfunktion $a_{VS,S}$ linearen oder exponentiellen Charakter haben soll.

Modellselektion und Parameterschätzung: Für die im Rahmen der Doktorarbeit untersuchten *Smart-Meter*-Datensätze wurden die VS-Modelle manuell ausgewählt und die Modellparameter mit einem Trust-Region-Verfahren, nämlich dem *trust region reflective least-squares algorithm* der MATLAB Optimization Toolbox [282, 283], geschätzt.

Zudem werden die verschiedenen VS-Modelle mit dem in Gleichung (1.4) vorgestellten ARX-Modell hinsichtlich der Qualität der Modellschätzung und der Generalisierungsfähigkeit verglichen. Die Ordnungen n_a , n_b und n_k des ARX-Modells werden mithilfe des Informationskriteriums von Akaike (engl. *Akaike information criterion*, AIC) ausgewählt. Die Schätzung der Modellparameter des ARX-Modells erfolgt durch die Kleinste-Quadrate-Methode [284, 285].

2.6.5 Modellanwendung

Wie in Abb. 2.24 zu sehen, knüpft die Modellanwendung an den Modellentwurf an. Nach dem Modellentwurf ist nicht ersichtlich, wie gut oder schlecht das entworfene, datengetriebene Modell ist. Das Modell sollte zudem beweisen, ob der funktionale Zusammenhang zwischen der erklärenden Variablen p und der Zielvariablen P_R auch für jene Daten gilt, die zum Zeitpunkt des Modellentwurfs unbekannt waren (Generalisierungsfähigkeit des Modells). Aus diesem Grund ist es unerlässlich, das entworfene Modell an unbekanntem Daten (Testdatensatz) anzuwenden.

Zur Bewertung des Modellverhaltens und der Modellgüte eignen sich verschiedene regressionsorientierte Bewertungsmaße, die im Folgenden vorgestellt werden. Da es in der Literatur eine Vielzahl an Gütemaßen gibt, werden hier nur die gängigsten Maße beschrieben und abschließend in Tabelle 2.11 zusammengefasst.

Regressionsorientierte Bewertungsmaße: Grundsätzlich dienen die Bewertungsmaße der Beurteilung der Modellgüte. Dabei wird die Modellschätzgröße \hat{y} (\hat{P}_R) in einen mathematischen Zusammenhang zur Zielvariablen y (P_R) gesetzt. Mit dem mittleren Fehler (engl.: *Mean Error*) $Q_{C,ME}$ können systematische Modellschätzfehler erkannt werden, da die Differenz¹⁷ ($y - \hat{y}$) zwischen der Zielvariablen y und der Modellschätzung \hat{y} nicht quadriert wird. Mit den Vektoren y und \hat{y} der Länge K berechnet sich der mittlere Fehler wie folgt:

$$Q_{C,ME}(y, \hat{y}) = \frac{1}{K} \sum_{k=1}^K (y[k] - \hat{y}[k]). \quad (2.82)$$

Wird die Differenz ($y[k] - \hat{y}[k]$) in Gleichung (2.82) durch den Betrag der Differenz ersetzt, resultiert der mittlere absolute Fehler (engl.: *Mean Absolute Error*) $Q_{C,MAE}$ mit

$$Q_{C,MAE}(y, \hat{y}) = \frac{1}{K} \sum_{k=1}^K |y[k] - \hat{y}[k]|. \quad (2.83)$$

Die Summe der Fehlerquadrate (engl.: *Sum of Squares Error*) $Q_{C,SSE}$ ist durch

$$Q_{C,SSE}(y, \hat{y}) = \sum_{k=1}^K (y[k] - \hat{y}[k])^2 \quad (2.84)$$

beschrieben. $Q_{C,SSE}$ kann auf die Anzahl der betrachteten Datenpunkte K bezogen werden, um so den Mittelwert des quadratischen Schätzfehlers zu erhalten. Der mittlere quadratische Fehler (engl.: *Mean Square Error*) $Q_{C,MSE}$ reagiert sehr sensitiv auf Ausreißer, da Abweichungen zwischen Modellschätzung und Ausgangsgröße überproportional gewichtet werden. $Q_{C,MSE}$ ist folgendermaßen definiert:

$$\begin{aligned} Q_{C,MSE}(y, \hat{y}) &= \frac{1}{K} Q_{C,SSE}(y, \hat{y}) \\ &= \frac{1}{K} \sum_{k=1}^K (y[k] - \hat{y}[k])^2. \end{aligned} \quad (2.85)$$

Der Modellfehler $Q_{C,RMSE}$ ist für die Modellbewertung von großer Bedeutung und ist als Wurzel der mittleren quadratischen Abweichung (engl.: *Root Mean Square Error*) festgelegt. $Q_{C,RMSE}$ erfasst wie $Q_{C,MSE}$ die Streuung der Differenzwerte ($y[k] - \hat{y}[k]$) für $k = 1, \dots, K$.

$$\begin{aligned} Q_{C,RMSE}(y, \hat{y}) &= \sqrt{Q_{C,MSE}(y, \hat{y})} \\ &= \sqrt{\frac{1}{K} \sum_{k=1}^K (y[k] - \hat{y}[k])^2} \end{aligned} \quad (2.86)$$

¹⁷auch Residuum genannt

Ein weiteres wichtiges Bewertungsmaß ist das Bestimmtheitsmaß (engl.: *coefficient of determination*) R^2 (siehe Gleichung (2.87)), welches die Verbesserung der Modellschätzung gegenüber dem Trivialmodell angibt. In diesem Fall ist das Trivialmodell der Mittelwert \bar{y} der Werte der Zielvariablen y für $k = 1, \dots, K$. Das Bestimmtheitsmaß R^2 nimmt Werte zwischen Null (oder 0%) und Eins (oder 100%) an. Der Wert Null gibt an, dass es zwischen den Größen keinen Zusammenhang gibt. Ein eindeutiger Zusammenhang zwischen y und \hat{y} ergibt sich für $R^2 = 1$.

$$R^2(\mathbf{y}, \hat{\mathbf{y}}) = 1 - \frac{\sum_{k=1}^K (y[k] - \hat{y}[k])^2}{\sum_{k=1}^K (y[k] - \bar{y})^2} = \frac{\sum_{k=1}^K (\hat{y}[k] - \bar{y})^2}{\sum_{k=1}^K (y[k] - \bar{y})^2} \quad (2.87)$$

Die Quadratwurzel des Bestimmtheitsmaßes R^2 ergibt bei einer einfachen Regression den Korrelationskoeffizienten $r(\mathbf{y}, \hat{\mathbf{y}})$:

$$r(\mathbf{y}, \hat{\mathbf{y}}) = \pm \sqrt{R^2(\mathbf{y}, \hat{\mathbf{y}})} \quad (2.88)$$

Ein dem Bestimmtheitsmaß R^2 ähnliches Bewertungsmaß ist der *Normalized Root Mean Square Error* $Q_{C,NRMSE}$. Das Fehlermaß ist definiert als

$$Q_{C,NRMSE}(\mathbf{y}, \hat{\mathbf{y}}) = \sqrt{\frac{\sum_{k=1}^K (y[k] - \hat{y}[k])^2}{\sum_{k=1}^K (y[k] - \bar{y})^2}}, \quad (2.89)$$

mit dem Mittelwert \bar{y} der Zielvariablen y . $Q_{C,NRMSE}$ entspricht dem Quotienten aus $Q_{C,RMSE}$ und der Standardabweichung der Zielvariablen y . Die regressionsorientierten Bewertungsmaße sind abschließend in Tabelle 2.11 zusammengefasst.

2.6.6 Zusammenfassung und Diskussion

Das Ziel von Abschnitt 2.6 war es, ein neues Konzept zur Modellierung des Verbrauchsverhaltens von Haushaltsstromkunden als Reaktion auf verschiedene Preissignale zu entwickeln. Dazu wurde in Abschnitt 2.6.2 eine neue systemtechnische *Demand-Response*-Modellklasse - die Klasse der Virtuellen-Speicher-Modelle - eingeführt, um die Laständerung eines Haushaltes als Reaktion auf ein beliebiges Strompreissignal zu beschreiben. Der strukturelle Aufbau der vier VS-Basismodelle (VSA-, VSA*- , VSB- und VSB*-Modell) beruht auf den Erkenntnissen von zahlreichen DR-Feldstudien, die als *a-priori*-Wissen in die Entwicklung der neuen VS-Modelle eingeflossen sind (siehe Abschnitt 2.6.1). Tabelle 2.12 fasst zusammen, welche Modellannahmen bzw. Erkenntnisse der datengetriebenen DR-Analyse als *a-priori*-Wissen in der VS-Modellfamilie realisiert wurden. Es muss jedoch betont werden, dass das *a-priori*-Wissen nicht alle Strukturfragen der VS-Modellfamilie klärt und dass dies die Existenz verschiedener Modellvarianten

Tabelle 2.11: Zusammenfassung ausgewählter Modellbewertungsmaße (modifiziert nach [85, 286]); Abkürzungen: L - Lagemaß, S - Streuungsmaß, Z - Zusammenhangsmaß, A - Absolut, R - Relativ, P - Proportional, Ü - Überproportional, J - Ja, N - Nein.

Bezeichnung	Formelzeichen	Gleichung	Eigenschaft	Gütemaß	Fehlergewichtung	Skalenerhaltend
Mittlerer Fehler	$Q_{C,ME}$	2.82	L	A	P	J
Mittlerer absoluter Fehler	$Q_{C,MAE}$	2.83	S	A	P	J
Summe der Fehlerquadrate	$Q_{C,SSE}$	2.84	S	A	Ü	N
Mittlerer quadratischer Fehler	$Q_{C,MSE}$	2.85	S	A	Ü	N
Wurzel des mittleren quadratischen Fehlers	$Q_{C,RMSE}$	2.86	S	A	Ü	J
Bestimmtheitsmaß	R^2	2.87	Z	R	Ü	N
Korrelationskoeffizient	r	2.88	Z	R	Ü	N
Normalized Root Mean Square Error	$Q_{C,NRMSE}$	2.89	S	R	Ü	N

erfordert. Die richtige Wahl des VS-Modells muss folglich für jeden *Smart-Meter*-Datensatz geklärt werden. Um die Wirksamkeit der neuen Modellklasse bewerten und mit gängigen Modellklassen (siehe Tabelle 1.4) vergleichen zu können, wurden in Abschnitt 2.6.4 und Abschnitt 2.6.5 die Modellierungsschritte „Modellentwurf“ und „Modellanwendung“ vorgestellt. Die Einteilung des *Smart-Meter*-Datensatzes in Lern- und Testdatensatz findet dort ebenso Erwähnung wie Maße zur Bewertung der Modellgüte.

In Haushalten mit einem Energiemanagementsystem und automatischem *Demand Response* kann das preisbeeinflusste Verbrauchsverhalten durch Optimierungsmodelle (vgl. Tabelle 1.4) beschrieben werden. Das Ergebnis der Optimierung wäre dann ein kostenoptimaler Lastfahrplan, der durch automatisiertes An- und Abschalten entsprechender steuer- und programmierbarer elektrischer Geräte „abgefahren“ wird. Damit der Netzbetreiber eine genaue Kenntnis über das (zukünftige) Verbrauchsverhalten der Haushalte in einem bestimmten Gebiet hat, müsste der kostenoptimale Lastfahrplan dem Netzbetreiber gemeldet werden. Es ist in diesem Fall jedoch zu bedenken, dass eine Einhaltung des kostenoptimalen Lastfahrplans nicht garantiert werden kann, da eine Abweichung vom gemeldeten Lastfahrplan durch spontanes An- und Abschalten der steuer- und programmierbaren elektrischen Geräte möglich ist.

Reagieren dagegen die Haushaltsbewohner selbst oder zum Teil auf ein bestimmtes Preissignal, müssen andere Modellierungsansätze (z.B. VS-Modelle) gewählt werden, um das preisbeeinflusste Verbrauchsverhalten zu beschreiben. Ein Hauptvorteil der VS-Modellfamilie ist zudem die gute Interpretierbarkeit durch die direkte Integration von *a-priori*-Wissen bei der Modellentwicklung und durch die geringe Anzahl an verständlichen Modellparametern. Ein weiterer Vorteil liegt in der Erweiterbarkeit der VS-Modelle durch die Nutzung einer modularen Modellstruktur, um unterschiedliche Einschränkungen und Annahmen zu integrieren.

In Kapitel 4 werden die VS-Modelle eingesetzt, um das preisbeeinflussten Verbrauchsverhaltens anhand realer *Smart-Meter*-Daten zu modellieren.

Tabelle 2.12: Realisierung von Modellannahmen und Erkenntnisse der datengetriebenen *Demand-Response*-Analyse als *a-priori*-Wissen im Virtuellen-Speicher-Modell

Modellannahmen oder Erkenntnisse der datengetriebenen DR-Analyse	Realisierung von <i>a-priori</i> -Wissen in der VS-Modellstruktur
Quantitative Lasterhöhungen und -senkungen als Reaktion auf bekannte zukünftige Strompreise (Modellannahme a) in Abschnitt 2.6.1)	Preisberechnungsvorschriften für p_f gemäß Tabelle 2.10 und Fallunterscheidungen je VS-Basismodell
Quantitative Lasterhöhungen und -senkungen als Reaktion auf unbekannte zukünftige Strompreise (Modellannahme a) in Abschnitt 2.6.1)	Preisberechnungsvorschriften für p_p gemäß Tabelle 2.10 und Fallunterscheidungen je VS-Basismodell
Aggregierte Last bzw. Last eines Haushaltes darf nicht negativ sein (Modellannahme c) in Abschnitt 2.6.1)	Begrenzung der geschätzten Laständerung gemäß Gleichung (2.62) und Gleichung (2.63)
Haushalte zeigen eher Lastverschiebungen als reine Laständerungen als hauptsächliche Verbrauchsreaktion (Modellannahme d) in Abschnitt 2.6.1)	Lastverschiebungen werden durch die Preisberechnungsvorschriften für p_p und p_f gemäß Tabelle 2.10 und durch die Fallunterscheidungen je VS-Basismodell realisiert
Lastverlagerungen beschränken sich im Bereich von Haushaltsstromkunden auf einen Zeitraum von höchstens 24 Stunden (Modellannahme e) in Abschnitt 2.6.1)	Füllstand des virtuellen Speichers soll sich im Laufe des Tages dem Sollfüllstand E_{VS}^s annähern; Realisierung durch die lineare und exponentielle Speicherausgleichsfunktion $a_{VS,S}^{lin}$ bzw. $a_{VS,S}^{exp}$ gemäß Gleichung (2.68) und Gleichung (2.69) im VSB- und VSB*-Modell
Gradient der Laständerung	Speicherbe- bzw. -entladegeschwindigkeit a_{VS}
Vergangene Lastreduktionen bzw. -erhöhungen als Reaktion auf ein Preissignal	Speicherfüllstand des virtuellen Speichers $E_{VS}[k]$ als interne Zustandsvariable (fungiert als „Gedächtnis“)
Amplitude der täglichen Last	Größe des virtuellen Speichers $E_{VS}^{max} - E_{VS}^{min}$
Zeitraum, wann eine Lastreduktion/-erhöhung als Reaktion auf ein Preissignal beginnt oder endet	Betrachtungshorizont des Strompreises H gemäß Tabelle 2.10

3 Implementierung

Kapitel 3 befasst sich mit der Implementierung der Methoden der Pipeline aus Abb. 2.1. Zunächst wird in Abschnitt 3.1 auf die *Open-Source* MATLAB-Toolbox SciXMiner im Allgemeinen eingegangen, die die Softwarebasis für die implementierten Methoden bildet. Die softwaretechnische Umsetzung des Benchmarkgenerators als wichtiges Werkzeug zur Validierung der Methoden der Pipeline ist Abschnitt 3.2 in beschrieben. Die SciXMiner-Erweiterung *Energy* fasst die implementierten Methoden zusammen (Abschnitt 3.3). In den drei Abschnitten 3.3.1, 3.3.2 und 3.3.3 ist die Implementierung der verfahrensübergreifenden Methoden aus Abschnitt 2.4, die Implementierung der analyseorientierten Methoden aus Abschnitt 2.5 und die Implementierung der prognoseorientierten Methoden der Pipeline aus Abschnitt 2.6 beschrieben. Kapitel 3 schließt mit einer kurzen Zusammenfassung.

3.1 *Open-Source* MATLAB-Toolbox SciXMiner

Grundsätzlich kann für die datengetriebene Analyse des preisbeeinflussten Verbrauchsverhaltens jede beliebige Data-Mining-Software verwendet werden. Eine gute Übersicht hierfür bietet [287]. Im Rahmen der Doktorarbeit wurde die *Open-Source* MATLAB-Toolbox SciXMiner [288] (aktuelle Version: SciXMiner 2017a¹; alter Name: *Gait-CAD*) zur Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens eingesetzt. SciXMiner wurde am Institut für Angewandte Informatik des Karlsruher Instituts für Technologie entwickelt und zeichnet sich durch eine einfache und intuitive Handhabung bei der Auswertung und Visualisierung von Zeitreihen und Einzelmerkmalen aus. In speziellen Anwendungsfällen greift SciXMiner auf MATLAB-Standard-Toolboxen, wie *Neural Network Toolbox* oder *Signal Processing Toolbox*, zurück.

SciXMiner ist mit einer grafischen Benutzeroberfläche (Beispiel siehe Abb. 3.1) mit Menüpunkten und Bedienelementen (Auswahllisten, Checkboxen und Editierfelder) ausgestattet, was die Benutzung der Software erleichtert. Zudem bietet SciXMiner die Möglichkeit, die Betätigung von Menü- und Bedienelementen als Sequenz in Form einer Makro-Textdatei (*.makrog) aufzuzeichnen und ggfs. manuell anzupassen. Somit können relativ schnell und ohne umfassende Programmierkenntnisse automatisierte Datenauswertungen und Modellierungen vorgenommen werden. Des Weiteren ist SciXMiner so aufgebaut, dass der Entwurf und die Anwendung von gängigen Data-Mining-Verfahren, wie beispielsweise Clusterverfahren oder Regressionsverfahren, modular durch sogenannte *extension packages* erweitert werden kann. Die *extension packages* (Erweiterungen) bieten die Möglichkeit, neue Menü- und Bedienelemente zu kapseln und relativ

¹<https://sourceforge.net/projects/scixminer/>, Abrufdatum: 20.12.2017

einfach in die Basisstruktur von SciXMiner zu integrieren. Die *extension packages* können bei Bedarf an- oder ausgeschaltet werden. Die im Folgenden vorgestellten SciXMiner-Erweiterungen *Load Benchmark Generator* und *Energy* werden ständig weiterentwickelt und sind auf Anfrage verfügbar. Es ist zudem geplant, die beiden SciXMiner-Erweiterungen *Load Benchmark Generator* und *Energy* in einer *Open-Source*-Publikation zu veröffentlichen.

3.2 SciXMiner-Erweiterung *Load Benchmark Generator*

In der SciXMiner-Erweiterung *Load Benchmark Generator* sind alle Methoden zur Generierung von Benchmarkdatensätzen (siehe Benchmarkgenerator in Abschnitt 2.3.2) mit spezifischen Eigenschaften, wie z.B. die Anzahl an Haushalten, die Abtastperiodendauer oder die Gesamtstromverbräuche je Haushalt, zusammengefasst. Die Eigenschaften des Benchmarkdatensatzes werden in einer csv-Datei gespeichert, wobei pro Zeile die Parametereinstellungen eines Haushaltes hinterlegt sind. Für jeden einzustellenden Parameter ist in der csv-Datei eine Spalte vorgesehen. Die csv-Datei enthält somit alle relevanten Informationen über den zu erstellenden Benchmarkdatensatz und durch diese Vorgehensweise können synthetische *Smart-Meter*-Datensätze mit einer beliebigen Anzahl an Haushalten vorkonfiguriert werden.

Soll nun ein Benchmarkdatensatz generiert werden, liest die Funktion² `load_benchmark_generator_settlement.m` (siehe Listing B.1) die Parametereinstellungen aus der csv-Datei und erzeugt eine *Timetable*³, aus der schließlich eine SciXMiner-Projektdatei (prjz-Datei) generiert wird. Im Detail setzt sich `load_benchmark_generator_settlement.m` aus folgenden Subfunktionen zusammen:

1. `defineSettingsSettlement.m` und `setPaths.m`: Skripte, die alle relevanten Parameter in eine struct-Datei schreiben, Pfade festlegen und Ordner für die Ergebnisse erzeugen
- 2a. `runSigGen.m`: Funktion, die den Lastsignalgenerator (siehe Abschnitt 2.3.2.1) mit den entsprechenden Parametern startet und eine *Timetable* mit mehreren Spalten (u.a. Zeit, Signalkomponenten, Gesamtsignal) generiert
- 2b. `runDatImp.m`: Funktion, die alternativ einen Import realer *Smart-Meter*-Messdaten ermöglicht und daraus eine *Timetable* mit mehreren Spalten (u.a. Zeit, Lastzeitreihe) generiert
3. `runPriImp.m`: Funktion, die die in der csv-Datei gewählten Strompreisprofile (siehe Abb. 2.4) importiert und als zusätzliche Spalten in der *Timetable* hinzufügt
4. `runDRSim.m`: Funktion, die die Verbrauchsreaktion eines Haushaltes auf ein synthetisches Strompreisprofil simuliert und als zusätzliche Spalte in der *Timetable* hinzufügt
5. `runDatMan.m`: Funktion, die die (zunächst fehlerfreie) Lastzeitreihe gemäß den Parametern aus der csv-Datei mit fehlenden Werten und Ausreißern überlagert und die manipulierte Lastzeitreihe als zusätzliche Spalte in der *Timetable* hinzufügt

²siehe Definition auf Seite 205

³siehe Definition auf Seite 205

6. runSaveMAT.m und runSaveCSV.m: Skripte, die das Ergebnis des Benchmarkgenerators als separate mat- und csv-Datei pro Haushalt speichern
7. timetable2scixminer.m: Funktion, die das Ergebnis des Benchmarkgenerators als separate SciXMiner-Projektdatei pro Haushalt speichert

Die im Benchmarkgenerator einzustellenden Parameter sind in Tabelle 3.1 zusammengefasst. In der letzten Spalte von Tabelle 3.1 ist zudem angegeben, an welcher Stelle des Benchmarkgenerators die Parameter eingestellt werden.

Tabelle 3.1: Einstellende Parameter des Benchmarkgenerators zur Erzeugung von Benchmarkdatensätzen

Parameter	Symbol	Formelbezug	Parameterfestlegung
Lastsignalgenerator			
Anzahl Haushalte	N	–	csv-Datei: je Haushalt ist eine Zeile vorgesehen
Zeitreihenlänge	K	–	Parameter in Funktion benchmarkGenerator.m
Abtastperiodendauer	T_S	–	Spalte in csv-Datei
Gesamtstromverbrauch	μ_E	Gleichung (2.12)	Spalte in csv-Datei
Varianz der Gesamtstromverbräuche	σ_E^2	Gleichung (2.12)	Spalte in csv-Datei
Gewichtungsfaktor der Trendkomponente	β_T	Gleichung (1.1)	Spalte in csv-Datei
Gewichtungsfaktor der exogenen Komponente	β_E	Gleichung (1.1)	Spalte in csv-Datei
Gewichtungsfaktor der periodischen Komponente	β_P	Gleichung (1.1)	Spalte in csv-Datei
Gewichtungsfaktor der Musterkomponente	β_M	Gleichung (1.1)	Spalte in csv-Datei
Gewichtungsfaktor der stochastische Komponente	β_S	Gleichung (1.1)	Spalte in csv-Datei
DR-Simulator			
Synthetischer Stromtarif	–	Gleichung (2.13)	Spalte in csv-Datei: Angabe der Tarifgruppe je Haushalt
Festlegung: R-, SR- oder NR-Haushalt	–	Gleichung (2.15)	Spalte in csv-Datei: Festlegung, ob R-, SR- oder NR-Haushalt je Haushalt
Messfehlergenerator			
Anzahl an Bereichen mit fehlenden Werten	b_{nan}^n	–	Spalte in csv-Datei
Proz. Anteil an Bereichen mit fehlenden Werten	$\mu_{nan,g}$	Gleichung (2.16)	Spalte in csv-Datei
Varianz von $\mu_{nan,g}$	$\sigma_{nan,g}^2$	Gleichung (2.20)	Spalte in csv-Datei
Proz. Anteil an einzelnen fehlenden Werten	$\mu_{nan,s}$	Gleichung (2.17)	Spalte in csv-Datei
Varianz von $\mu_{nan,s}$	$\sigma_{nan,s}^2$	Gleichung (2.19)	Spalte in csv-Datei
Prozentualer Anteil an Ausreißern	μ_a	Gleichung (2.18)	Spalte in csv-Datei
Varianz von μ_a	σ_a^2	Gleichung (2.21)	Spalte in csv-Datei
Untere Intervallgrenze des Ausreißermultiplikators	a	Gleichung (2.22)	Spalte in csv-Datei
Obere Intervallgrenze des Ausreißermultiplikators	b	Gleichung (2.22)	Spalte in csv-Datei

3.3 SciXMiner-Erweiterung Energy

Die SciXMiner-Erweiterung *Energy* fasst die Funktionalitäten und Methoden des neuen Konzeptes zur datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens

zusammen. Im Folgenden wird die Umsetzung der verfahrensübergreifenden Methoden der Pipeline in Abschnitt 3.3.1, die Umsetzung der analyseorientierten Methoden der Pipeline in Abschnitt 3.3.2 und die Umsetzung der prognoseorientierten Methoden der Pipeline in Abschnitt 3.3.3 erläutert.

3.3.1 Automatisierung der verfahrensübergreifenden Methoden der Pipeline

Die Automatisierung der verfahrensübergreifenden Methoden der Pipeline wird in SciXMiner durch Makros realisiert, die über Menüelemente aufgerufen werden. So zeigt Abb. 3.1 einen Screenshot der relevanten Menüelemente in der SciXMiner-Erweiterung *Energy*, die zudem mit nummerierten Kreisen gekennzeichnet sind.

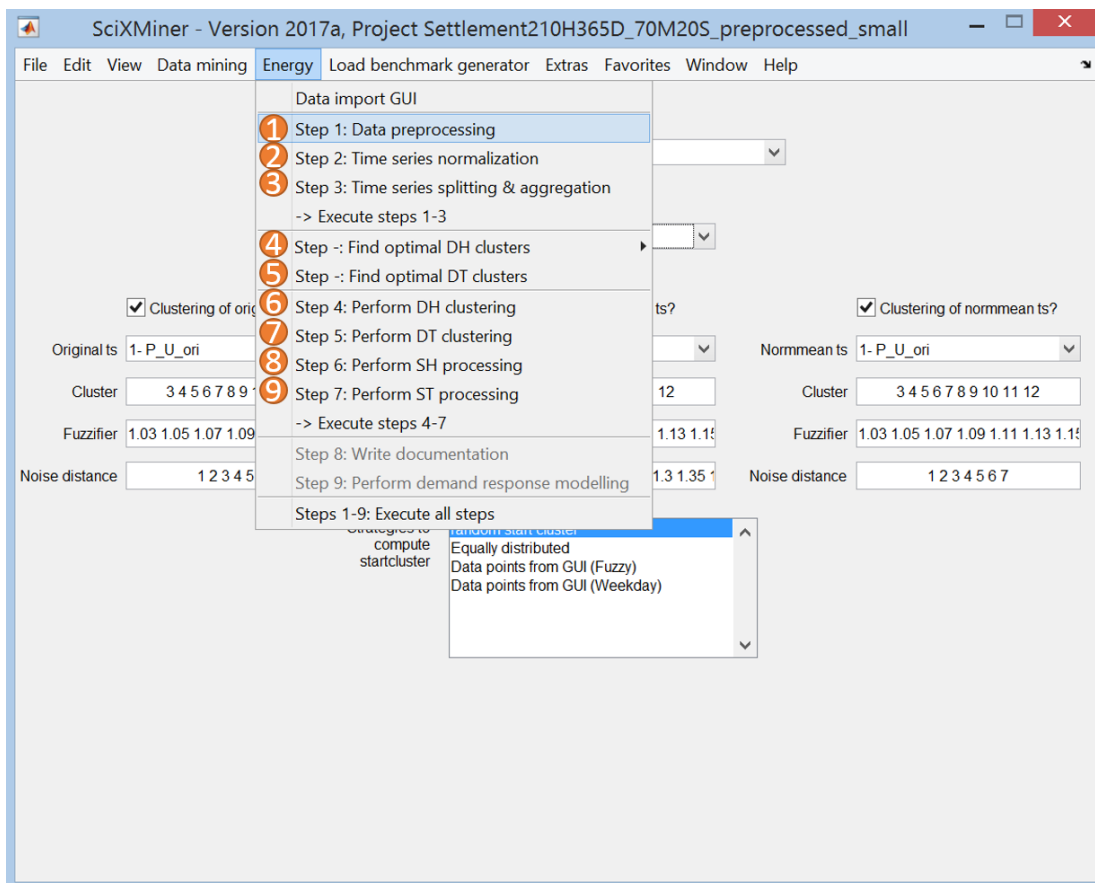


Abbildung 3.1: Exemplarischer Screenshot der Menüs und Bedienelemente in der SciXMiner-Erweiterung *Energy*: (1) Datenvorverarbeitung (Umgang mit fehlenden Messwerten und Ausreißern), (2) Zeitreihennormierung, (3) Zeitreihenzerlegung und -aggregation, (4) Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern (für Fall DH), (5) Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern (für Fall DT), (6) Zeitreihen-Clustering der Fälle DH, DH_{norm} und $DH_{normmean}$, (7) Zeitreihen-Clustering der Fälle DT, DT_{norm} und $DT_{normmean}$, (8) Analyse einzelner Abtastzeitpunkte bezüglich der Tarifeinflüsse auf die Last (für Fall SH), (9) Analyse einzelner Abtastzeitpunkte bezüglich der Tarifeinflüsse auf die Last (für Fall ST)

Schritt 1 und 2 in Abb. 3.1 sind für die Datenvorverarbeitung (siehe Abschnitt 2.4.1) bzw. Zeitreihennormierung (siehe Abschnitt 2.4.3) zuständig. Der mit der Nummer 3 gekennzeichnete Schritt vereint die Zeitreihenzerlegung (siehe Abschnitt 2.4.2) und die Zeitreihenaggregation (siehe Abschnitt 2.4.4). Um zu gewährleisten, dass die verfahrensübergreifenden Methoden der Pipeline automatisiert bei Datensätzen mit verschiedenen Bezeichnern funktioniert, sind im Skript `defineNames.m` (siehe Listing B.2) Kandidaten der möglichen Bezeichner hinterlegt. Beim Starten der Datenvorverarbeitung wird `defineNames.m` aufgerufen und die Bezeichner der Zeitreihen und der Ausgangsgrößen aus der Projektdatei mit den Kandidaten aus `defineNames.m` verglichen.

3.3.2 Automatisierung der analyseorientierten Methoden der Pipeline

Die analyseorientierten Methoden der Pipeline werden ebenfalls über die in Abb. 3.1 gezeigten Menüelemente ausgeführt. So wird beispielsweise das Clustering der Fälle DH, DH_{norm} und $DH_{normmean}$ mithilfe des Menüelements mit der Nummer 6 gestartet. Listing 3.1 zeigt das Skript `processDHClustering.m`, welches anschließend ausgeführt wird. Zuvor müssen jedoch die Einstellungen der Clusterparameter im Skript `fall_dh_set_parameter_clustering.m` in Code-Zeile 7 in Listing 3.1 festgelegt werden.

Listing 3.1: Skript `processDHClustering.m`

```

1 %% Process DH clustering
2 % Load DH project
3 loadDH;
4 % Get names and linguistic terms
5 defineNames;
6 % Set parameter for dh clustering
7 fall_dh_set_parameter_clustering;
8 % Perform DH clustering
9 fall_dh_perform_clustering_new;
10 % Save clustered dh project
11 saveClusterproject;
12 %% Plotting
13 set_standard_parameters;
14 plot_options_color; % or: plot_options_bw;
15 % fall_dh_images_optimal_clusters;
16 fall_dh_images;
17 % Cluster histograms
18 fall_dh_show_cluster_ori_norm;

```

Wie in Listing 3.1 in Code-Zeile 11 zu sehen, wird nach dem Clustering eine neue Projektdatei mit den geclusterten Zeitreihen gespeichert. Die automatisierte Generierung von Auswertefiguren, wie z.B. die Clusterkurven in Abb. 2.14, wird mithilfe der Skripte `fall_dh_images_optimal_clusters.m` und `fall_dh_images.m` (Code-Zeile 15 und 16 in Listing 3.1) realisiert. Die Bewertung der Zusammenhänge von gefundenen Clustern und Werktagen, Monaten oder Tarifen mit Verbundverteilungen und informationstheoretischen Maßen wird durch das Skript `fall_dh_show_cluster_ori_norm.m` (Code-Zeile 18 in Listing 3.1) ausgeführt. Die Analyse der Tarifeinflüsse auf die Last erfolgt durch die Analyse der Fälle SH und ST und kann in der SciXMiner-Erweiterung *Energy* über zwei weitere Menüelemente (siehe Nummer 8 und 9 in Abb. 3.1) gestartet werden.

3.3.3 Automatisierung der prognoseorientierten Methoden der Pipeline

Die datengetriebene Modellierung des preisbeeinflussten Verbrauchsverhaltens mithilfe der Virtuellen-Speicher-Modelle stellt einen wichtigen Baustein der vorliegenden Doktorarbeit dar. Die methodischen Grundlagen hierfür wurden in Abschnitt 2.6 erläutert. Die VS-Basismodelle (VSA-, VSA*-, VSB- oder VSB*-Modell) sind in der SciXMiner-Erweiterung *Energy* als jeweils eigene Modellklasse implementiert, die die SciXMiner-Standardmodellklassen, wie beispielsweise KNN oder Lineare Regressionsmodelle, um weitere Modellklassen ergänzen. Der Entwurf von Regressionsmodellen wird in SciXMiner durch die Funktion `regression_en.m` gestartet, in der alle Regressionsmodelle statisch definiert sein müssen. Die VS-Modelle sind direkt in SciXMiner implementiert, jedoch nur dann verfügbar, wenn die SciXMiner-Erweiterung *Energy* aktiviert ist. So zeigt Listing 3.2 die entsprechenden Code-Zeilen aus `regression_en.m` zum Entwurf des VSA-Modells.

Listing 3.2: Auszug aus der Funktion `regression_en.m`

```
1     case 'vsa' % Virtual Storage A
2         if exist('vsa_en.m','file') && exist('vsa_an.m','file')
3             regr_single.vsa = kp.regression.vsa;
4             regr_single     = vsa_en(d,ykont,regr_single);
5         else
6             regr_single.vsa = [];
7             myerror('Bitte erst Virtueller Speicher A Modell implementieren bzw. ...
8                 installieren!');
9         end;
```

Soll nun das preisbeeinflusste Verbrauchsverhalten mithilfe eines VS-Modells modelliert werden, muss die Auswahl der entsprechenden Basismodellklasse und die Spezifikation der VS-Basismodelle durch Bedienelemente erfolgen, was in Abb. 3.2 exemplarisch gezeigt ist.

Anschließend erfolgt der Entwurf (vgl. Abschnitt 2.6.4) und Anwendung (vgl. Abschnitt 2.6.5) des entworfenen Modells. Für jedes VS-Modell existiert jeweils eine separate Funktion für den Modellentwurf und die Modellanwendung. Beispielhaft sind in Listings B.3 und B.4 die entsprechenden Funktionen für das VSA-Modell gezeigt.

3.4 Zusammenfassung

Kapitel 3 befasste sich mit der Implementierung der Methoden und Verfahren aus Kapitel 2. Die *Open-Source* MATLAB-Toolbox SciXMiner bildet die Softwarebasis für die implementierten Methoden, wobei diese hauptsächlich in zwei SciXMiner-Erweiterungen (*Load Benchmark Generator* und *Energy*) gekapselt sind. Der Großteil der implementierten Methoden sind als Menüpunkte und Bedienelemente in den SciXMiner-Erweiterungen verwendbar, was den Vorteil bietet, dass die verfahrensübergreifenden, analyse- und prognoseorientierten Methoden relativ einfach und ohne MATLAB-Kenntnisse benutzt werden können. Daneben sind einzelne Methoden als MATLAB-Skripte realisiert und verwendbar. Das Anpassen und Verwenden dieser Skripte setzt jedoch MATLAB-Kenntnisse voraus.

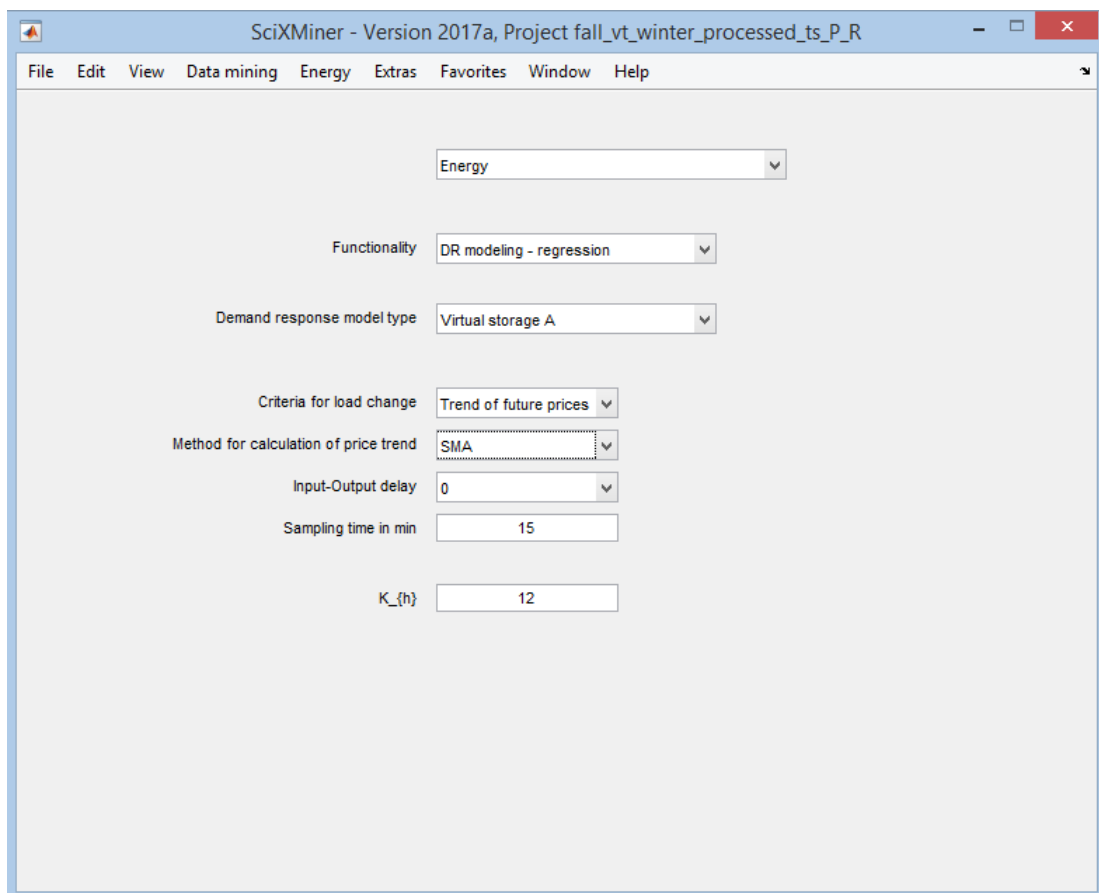


Abbildung 3.2: Exemplarischer Screenshot der Bedienelemente in der SciXMiner-Erweiterung *Energy* zur Auswahl und Spezifikation des VSA-Modells (vgl. Tabelle 2.10)

4 Anwendungen

Die Ziel von Kapitel 4 besteht darin, das in Kapitel 2 vorgestellte Konzept zur datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens an realen *Smart-Meter*-Daten zu testen. Dazu widmen sich die Abschnitte 4.1 und 4.2 den *Smart-Meter*-Datensätzen (*Olympic Peninsula Project* (OPP), RESIDENS), welche im Rahmen der Doktorarbeit untersucht werden.

In Abschnitt 2.2 wurden vier Aufgabenstellungen identifiziert, die mithilfe des neuen Konzeptes zur datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens beantwortet werden können. Die Erfüllung dieser Aufgabenstellungen hängt stark von den Eigenschaften der *Smart-Meter*-Datensätzen ab, was in den Überschriften der Abschnitte 4.1 und 4.2 erkennbar wird. So eignet sich der erste Datensatz (OPP) für die Bearbeitung der Aufgabenstellungen A1-A4 (siehe Abschnitt 4.1). Der zweite untersuchte Datensatz (RESIDENS) enthält keine Messdaten einer Kontrollgruppe, was die Bearbeitung der Aufgabenstellungen A2 und A4 erschwert bzw. nicht möglich macht. Deswegen befasst sich Abschnitt 4.2 nur mit den Aufgabenstellungen A1 und A3. Kapitel 4 schließt mit einer Zusammenfassung und Diskussion der Ergebnisse.

4.1 Olympic-Peninsula-Project-Datensatz

4.1.1 Allgemeines zum Datensatz

Der erste *Smart-Meter*-Datensatz stammt aus der *Pacific Northwest GridWise Testbed Demonstration* Feldstudie [73]. Im Teilprojekt *Olympic Peninsula Project*¹ (OPP) wurde das Verbrauchsverhalten von Haushaltsstromkunden bei variablen Stromtarifen und geeigneter Technik bzw. bidirektionaler Kommunikation untersucht. An der Feldstudie nahmen insgesamt 112 Haushalte teil, wobei die Messdatenerhebung im Zeitraum von April 2006 bis März 2007 stattfand. Die Haushalte wurden zu gleichen Teilen vier Gruppen zugeordnet:

1. Ein dreistufiger TOU-Tarif veranlasste Haushalte zur Änderung des Verbrauchsverhaltens. Der TOU-Tarif setzte sich aus drei Tarifstufen zusammen: dem *on-peak*-Tarif (Frühling, Herbst und Winter: 12,2 ¢/kWh; Sommer: 13,5 ¢/kWh), dem *off-peak*-Tarif (Frühling, Herbst und Winter: 4,1 ¢/kWh; Sommer: 5 ¢/kWh) und dem *Critical Peak Pricing* Tarif (35 ¢/kWh). Der CPP-Tarif wurde nur am 1. November 2006 für vier Stunden ausgerufen

¹Download des Datensatzes mit vorheriger Anmeldung unter <https://svn.pnl.gov/olyphen>, Abrufdatum: 12.06.2013

und diente der Netzentlastung in einer netzkritischen Situation. Der tageszeitliche Verlauf des TOU-Tarifs war den Haushaltsstromkunden im Voraus bekannt.

2. Den Haushalten der RTP-Gruppe wurde ein zeitvariabler Stromtarif übermittelt, der in Echtzeit an einem virtuellen Strommarkt gemäß Angebot und Nachfrage (*transactive control*) gebildet wurde. Aufgrund von fehlenden Messdaten im April 2006 ist die aufgezeichnete Datenmenge der RTP-Gruppe geringer als die der anderen Tarifgruppen.
3. Die Haushalte der FIXED-Gruppe erhielten einen konstanten Stromtarif von 8,1 ¢/kWh.
4. In Ergänzung zu den drei Tarifgruppen wurden 24 Haushalte der CTRL-Gruppe zugewiesen. Da die Messdaten der CTRL-Gruppe eine sehr schlechte Datenqualität besitzen, werden diese im Folgenden vernachlässigt und stattdessen die FIXED-Gruppe als Kontrollgruppe behandelt.

Der Verlauf des Strompreises der einzelnen Tarifgruppen ist in Abb. 4.1 zu sehen.

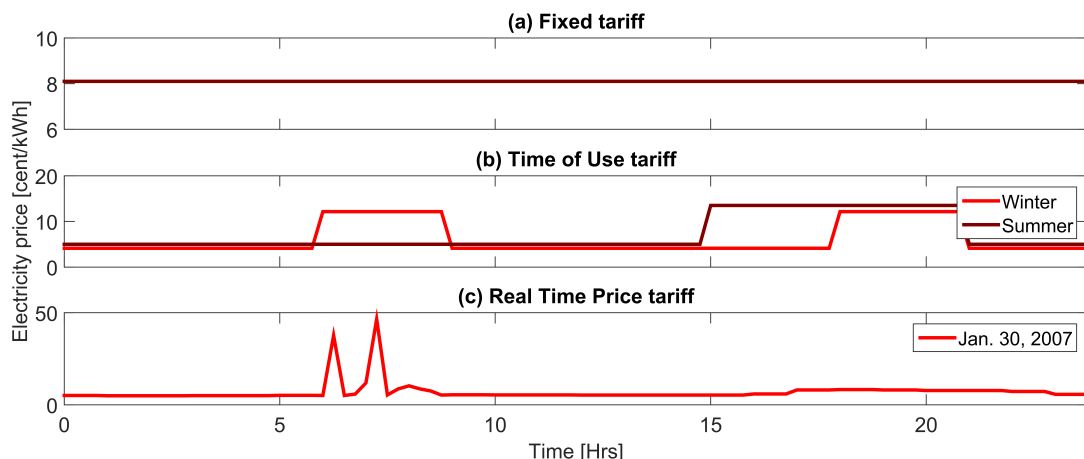


Abbildung 4.1: Strompreise der Tarifgruppen des *Olympic Peninsula Project*: (a) Strompreis der FIXED-Tarifgruppe, (b) saisonal unterschiedliches Strompreisprofil der TOU-Tarifgruppe, (c) Echtzeit-Strompreisprofil der RTP-Tarifgruppe für einen beispielhaften Tag.

Die Haushalte, die an der OPP-Feldstudie teilnahmen, wurden mit Geräten zur elektrischen Warmwasserbereitung und Luftheizungsanlagen (Luft/Wasser-Wärmepumpe oder elektrische Warmluftheizung) ausgestattet [73]. Zudem konnten die Haushalte den Warmwasserbereiter und den Raumtemperaturregler entsprechend ihrer Komfortbedürfnisse so konfigurieren, dass die elektrischen Geräte automatisiert auf Preisanreize reagieren (Automatisches *Demand Response*, siehe Abschnitt 1.2.5). Die Haushaltslast setzt sich nicht nur aus der Last der Geräte zur elektrischen Warmwasserbereitung und Luftheizungsanlagen zusammen, sondern aus der Last aller im Haushalt befindlichen Geräte (z.B. Kühlschrank, Waschmaschine, Küchengeräte oder TV). Des Weiteren war an bestimmten Haushaltsgeräten - wie der Waschmaschine - eine Anzeige des aktuellen Strompreises angebracht. Die beeinflusste Verbraucherlast der preisbeeinflussten Haushalte besteht demnach aus drei Anteilen:

- dem unbeeinflussten Anteil der Verbraucherlast,

- dem beeinflussten Anteil der Verbraucherlast, der durch automatisches *Demand Response* hervorgerufen wurde und
- dem beeinflussten Anteil der Verbraucherlast, der durch das manuelle Ein- bzw. Ausschalten bestimmter Haushaltsgeräte, wie Wäschetrockner oder Waschmaschine, hervorgerufen wurde.

Die Gesamtlast aller steuerbaren Geräte belief sich im Durchschnitt auf 1,5 kW je Haushalt [73]. Die Struktur des OPP-Datensatzes ist in Tabelle B.12 zusammengefasst. Vor der Datenvorverarbeitung enthält der *Smart-Meter*-Datensatz Zeitreihen von 112 Haushalten. Da allerdings Zeitreihen mit mehr als 3% fehlenden Werten komplett gelöscht werden und die Haushalte der CTRL-Gruppe ausgeschlossen werden, enthält der GH-Fall des OPP-Datensatzes nur 84 Datentupel.

4.1.2 Optimale Clusterparameter

Zur Bestimmung von optimalen Clusterparametern wird die in Abschnitt 2.5.4.1 vorgestellte Sensitivitätsanalyse an den DH- und DT-Fällen des OPP-Datensatzes durchgeführt. Die Festlegung, welche Parameterkombination aus C , q und δ zu einer optimalen Clustereinteilung führt, geschieht auf Basis von Heatmaps, die die Sensitivität des Trennungsgrades, die Sensitivität des prozentualen Anteils der Datentupel im Rauschcluster und die Sensitivität der Werte der relativen Transinformation bezüglich der Eingangsparameter C , q und δ darstellen. In Tabelle 4.1 ist das Ergebnis der Bestimmung von optimalen Clusterparametern für den OPP-Datensatz eingetragen. Jede der dort gewählten Parameterkombination stellt eine „manuell ausgewählte Kompromisslösung“ bezüglich der drei oben genannten Ziele dar. Wie bereits in Abb. 2.18

Tabelle 4.1: Ergebnis der Bestimmung von optimalen Clusterparametern für den OPP-Datensatz

Fall	Tage	opt. Clusteranzahl	opt. Fuzzifizierungsgrad	opt. Rauschabstand	Verweis
DH	alle Wochentage	$C_{opt} = 5$	$q_{opt} = 1.05$	$\delta_{opt} = 10$	Abb. 4.2
DH _{norm}	alle Wochentage	$C_{opt} = 4$	$q_{opt} = 1.25$	$\delta_{opt} = 1.5$	Abb. B.12
DH _{normmean}	alle Wochentage	$C_{opt} = 5$	$q_{opt} = 1.70$	$\delta_{opt} = 30$	Abb. B.13
DH ^{wo}	nur Werktage	$C_{opt} = 5$	$q_{opt} = 1.10$	$\delta_{opt} = 6$	Abb. B.14
DH _{norm} ^{wo}	nur Werktage	$C_{opt} = 4$	$q_{opt} = 1.25$	$\delta_{opt} = 1.5$	Abb. B.15
DH _{normmean} ^{wo}	nur Werktage	$C_{opt} = 5$	$q_{opt} = 1.70$	$\delta_{opt} = 30$	Abb. B.16
DT	alle Wochentage	$C_{opt} = 4$	$q_{opt} = 1.40$	$\delta_{opt} = 2.75$	Abb. B.17
DT _{norm}	alle Wochentage	$C_{opt} = 9$	$q_{opt} = 1.05$	$\delta_{opt} = 1.3$	Abb. B.18
DT _{normmean}	alle Wochentage	$C_{opt} = 5$	$q_{opt} = 1.20$	$\delta_{opt} = 3$	Abb. B.19
DT ^{wo}	nur Werktage	$C_{opt} = 4$	$q_{opt} = 1.55$	$\delta_{opt} = 3.25$	Abb. B.20
DT _{norm} ^{wo}	nur Werktage	$C_{opt} = 4$	$q_{opt} = 1.05$	$\delta_{opt} = 1.8$	Abb. B.21
DT _{normmean} ^{wo}	nur Werktage	$C_{opt} = 8$	$q_{opt} = 1.10$	$\delta_{opt} = 3.0$	Abb. B.22

und 2.19 gezeigt, eignen sich 3D-Oberflächengrafiken und Heatmaps gut, um die Sensitivität des Trennungsgrades, die Sensitivität des prozentualen Anteils der Datentupel im Rauschcluster

und die Sensitivität der Werte der relativen Transinformation bezüglich der Eingangsparameter C , q und δ beurteilen zu können. Jedoch können an dieser Stelle aus Platzgründen nicht für alle zwölf Fälle aus Tabelle 4.1 die Grafiken dargestellt werden. Das Clustering des Falls DH mit den in Tabelle 4.1 gezeigten Clusterparametern führt zu den in Abb. 4.2 dargestellten Ergebnissen. Die entsprechenden Abbildungen der restlichen elf Fälle sind in Anhang B.4.1 zu finden.

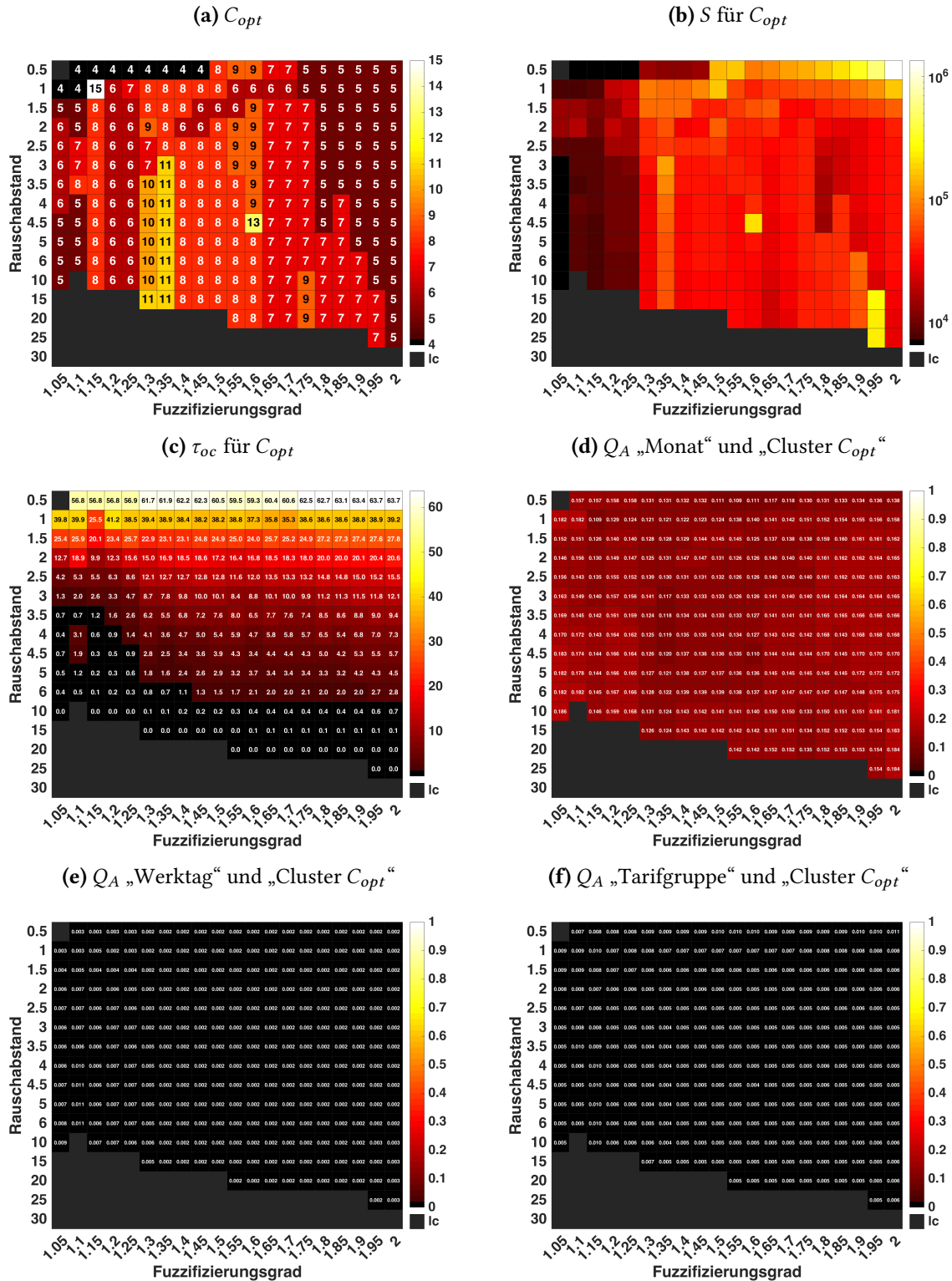


Abbildung 4.2: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH des OPP-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(f) relative Transinformation pro Ausgangsentropie). Für Fall DH_{norm} siehe Abb. B.12, für Fall $DH_{normmean}$ siehe Abb. B.13. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

4.1.3 Aufgabenstellung A1: Identifikation von saisonalen und wochentagsbedingten Unterschieden für eine Demand-Response-Maßnahme

Es sei nochmals an die Aufgabenstellung A1 aus Abschnitt 2.2 erinnert: „Es sollen saisonale und wochentagsbedingte Schwankungen für eine Demand-Response-Maßnahme, z.B. für Fälle mit ungewöhnlichen Preissignalen, identifiziert werden. Das neue Konzept soll beispielsweise Unterschiede in der Verbrauchsreaktion auf ein Preissignal zwischen Winter- und Sommermonaten aufdecken.“ In Abschnitt 4.1.3 werden nun nacheinander die preisbeeinflussten Typtage der DH -, DH_{norm} -, $DH_{normmean}$ -, DH^{wo} -, DH_{norm}^{wo} -, $DH_{normmean}^{wo}$ -, DT -, DT_{norm} -, $DT_{normmean}$ -, DT^{wo} -, DT_{norm}^{wo} - und $DT_{normmean}^{wo}$ -Fälle vorgestellt.

4.1.3.1 DH -, DH_{norm} - und $DH_{normmean}$ -Cluster als preisbeeinflusste Typtage (Alle Wochentage)

Wie bereits in Tabelle 2.9 zusammengefasst, können saisonale und wochentagsbedingte Schwankungen durch die visuelle Inspektion der gefundenen Clusterkurven aufgedeckt werden. Mögliche Zusammenhänge zwischen den gefundenen Clusterkurven und den Ausgangsgrößen „Monat“ oder „Werktag (ja/nein)“ werden anschließend mit den Methoden aus Abschnitt 2.5.3 bewertet. Die gefundenen Clusterkurven, die auch als preisbeeinflusste Typtage bzw. typische preisbeeinflusste Lastprofile bezeichnet werden, sollen im Folgenden für die Fälle DH , DH_{norm} und $DH_{normmean}$ vorgestellt werden, wobei zunächst keine Wochentage ausgeschlossen werden. Wie in Tabelle 4.1 zu sehen, ergibt sich für den Fall DH eine optimale Clusteranzahl von $C_{opt} = 5$. Das Clustering der varianznormierten Tageskurven führt zu $C_{opt} = 4$, das Clustering der mittelwertnormierten Tageskurven führt zu $C_{opt} = 5$, wobei in allen drei Fällen eines der Cluster ein Rauschcluster ist.

Nachdem nun die optimale Anzahl an Clusterzentren ermittelt wurde, werden im Folgenden die Ergebnisse des Clusterings - die preisbeeinflussten Typtage in Abb. 4.5a - präsentiert und diskutiert. Bei der visuellen Inspektion der haushaltsindividuellen Tageskurven sind gut interpretierbare Cluster auszumachen. Die Tagesverläufe der einzelnen DH -Cluster in Abb. 4.5a(i) variieren hauptsächlich in der Gesamtstrommenge, wohingegen die DH_{norm} -Cluster in Abb. 4.5a(ii) und die $DH_{normmean}$ -Cluster in Abb. 4.5a(iii) sehr deutlich tageszeitliche Lastverschiebungen zeigen.

Das Balkendiagramm in Abb. 4.3a verdeutlicht, dass die Tagesverläufe der einzelnen DH -Cluster in Abb. 4.5a(i) das saisonal geprägte Verbrauchsverhalten von Haushaltsstromkunden abbilden. Während die Cluster C2, C3 und C4 das Verbrauchsverhalten zur kalten Jahreszeit repräsentieren, enthält das Cluster C1 hauptsächlich Tageskurven aus der warmen Jahreszeit. Die Werte der relativen Transinformation der Ausgangsgröße „Monat“ in Tabelle 4.2 zeigen zudem, dass durch die DH - und DH^{wo} -Cluster der Einfluss des Monats (saisonale Einflüsse) am besten bewertet werden kann. Die Fälle DH_{norm} , DH_{norm}^{wo} , $DH_{normmean}$ und $DH_{normmean}^{wo}$ eignen sich hingegen nicht, um saisonale Unterschiede zu erklären.

Wochentagsbedingte Unterschiede im preisbeeinflussten Verbrauchsverhalten lassen sich erklären, wenn die Werte für Q_A „Werktag“ für die Fälle DH , DH_{norm} und $DH_{normmean}$ in Tabelle 4.2

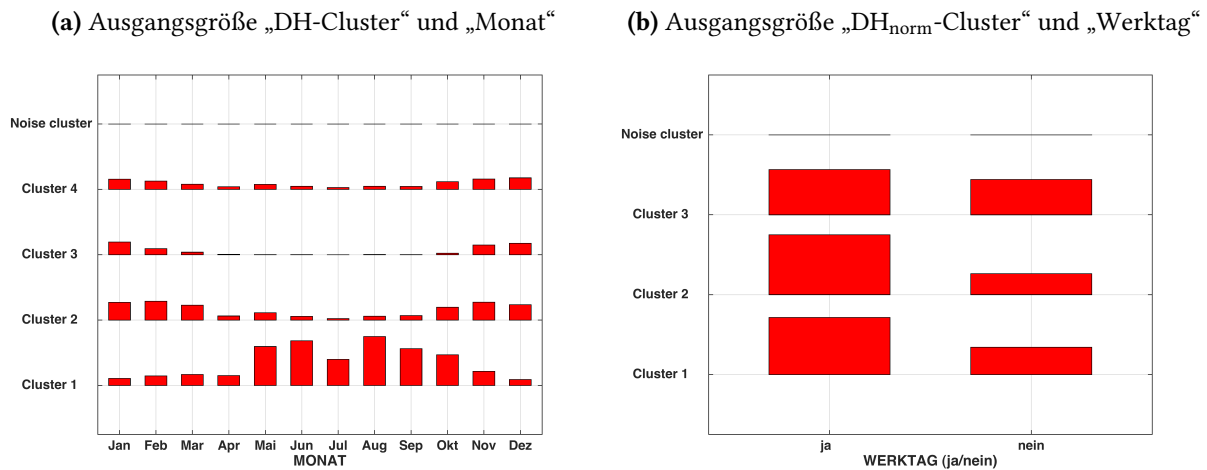


Abbildung 4.3: Verbundverteilungen zweier Ausgangsgrößen des OPP-Datensatzes als Balkendiagramme: (a) Ausgangsgröße „DH-Cluster“ und „Monat“, (b) Ausgangsgröße „DH_{norm}-Cluster“ und „Werktag“

entsprechend hoch sind. Da allerdings der Wert für Q_A in allen drei Fällen maximal 0,01 beträgt, liefert das Clustern dieser Fälle keine Erkenntnisse über den Zusammenhang der Klassen der Ausgangsgrößen „Werktag (ja/nein)“ und den entsprechenden Clustern (siehe Abb. 4.3b). Die Werte der relativen Transinformation pro Ausgangsentropie für die Ausgangsgröße „Tarifgruppe“ sind bereits in Tabelle 4.2 gezeigt, werden aber erst in den Abschnitten 4.1.4 und 4.1.5 bewertet und diskutiert.

Tabelle 4.2: Untersuchte Fälle des OPP-Datensatzes: Auswertung der Clusteranalysen mit informationstheoretischen Maßen; C_{opt} - optimale Clusteranzahl, N_D - Anzahl an Datentupeln, Q_A - relative Transinformation pro Ausgangsentropie.

Fall	Tage	C_{opt}	N_D	Q_A Monat	Q_A Werktag (ja/nein)	Q_A Tarifgruppe
DH	alle Wochentage	5	21318	0.19	0.01	0.01
DH _{norm}	alle Wochentage	4	21318	0.03	0.01	0.01
DH _{norm} mean	alle Wochentage	5	21318	0.02	0.00	0.00
DH ^{wo}	nur Werktage	5	14076	0.20	–	0.01
DH _{norm} ^{wo}	nur Werktage	4	14076	0.03	–	0.01
DH _{norm} mean ^{wo}	nur Werktage	5	14076	0.03	–	0.01
DT	alle Wochentage	4	931	0.62	0.01	0.01
DT _{norm}	alle Wochentage	9	931	0.25	0.09	0.07
DT _{norm} mean	alle Wochentage	5	931	0.21	0.13	0.03
DT ^{wo}	nur Werktage	4	613	0.63	–	0.01
DT _{norm} ^{wo}	nur Werktage	4	613	0.35	–	0.03
DT _{norm} mean ^{wo}	nur Werktage	8	613	0.31	–	0.12

4.1.3.2 DH^{wo} -, DH_{norm}^{wo} - und $DH_{normmean}^{wo}$ -Cluster als preisbeeinflusste Typtage (Nur Werktage)

Um den Einfluss von unterschiedlichen Tagesroutinen an Werktagen und Wochenendtagen zu unterdrücken, beschränkt sich die folgende Clusteranalyse auf das Clustering der Fälle DH^{wo} , DH_{norm}^{wo} und $DH_{normmean}^{wo}$, die nur Werktage enthalten. Der hochgestellte Index „wo“ steht für *working days* (englisch für Werktage). Die **optimale Clusteranzahl** wurde bereits in Abschnitt 4.1.2 ermittelt. Für den Fall DH^{wo} ergibt sich eine optimale Clusteranzahl von $C_{opt} = 5$. Das Clustering der varianznormierten Tageskurven führt hier zu $C_{opt} = 4$, das Clustering der mittelwertnormierten Tageskurven führt hier zu $C_{opt} = 5$, wobei in allen drei Fällen eines der Cluster ein Rauschcluster ist.

Die **preisbeeinflussten Typtage** der drei Fälle DH^{wo} , DH_{norm}^{wo} und $DH_{normmean}^{wo}$ sind in Abb. 4.5b(i)-(iii) zu sehen. Die Tagesverläufe der DH^{wo} -Cluster in Abb. 4.5b(i) ähneln stark denen der DH-Cluster in Abb. 4.5a(i). Der Zusammenhang zwischen der Ausgangsgröße „Monat“ und den gefundenen Clustern ist für den Fall DH^{wo} am stärksten ($Q_A = 0.20$). Saisonale Unterschiede lassen sich somit am ehesten mit den Clusterkurven der nicht-normierten Tageskurven beschreiben. Das Clustering der normierten Fälle DH_{norm}^{wo} und $DH_{normmean}^{wo}$ liefert keine Erkenntnisse über den Zusammenhang der gefundenen Clusterkurven und der Ausgangsgröße „Monat“.

Wochentagsbedingte Unterschiede im preisbeeinflussten Verbrauchsverhalten lassen sich nicht erklären, da alle Samstage und Sonntage ausgeschlossen und alle Montage bis Freitage zur Klasse „Werktag“ zusammengefasst wurden. Eine weitere Aufteilung der Werktage in die Klassen {„Montag“, . . . , „Freitag“} ist möglich, wurde im Rahmen der Doktorarbeit nicht gemacht.

4.1.3.3 DT -, DT_{norm} - und $DT_{normmean}$ -Cluster als preisbeeinflusste Typtage (Alle Wochentage)

Durch die Aggregation von Tageszeitreihen einzelner Haushalte (Fälle DH , DH_{norm} , $DH_{normmean}$ in Tabelle 2.3) zu Tageszeitreihen aller Haushalte einer Tarifgruppe (Fälle DT , DT_{norm} , $DT_{normmean}$ in Tabelle 2.3) verringert sich die Varianz der zu clusternden Zeitreihen. Auch hier wurde die **optimale Clusteranzahl** in einer Sensitivitätsanalyse ermittelt, deren Ergebnisse in Tabelle 4.1 zu sehen sind. Für den Fall DT ergibt sich eine optimale Clusteranzahl von $C_{opt} = 4$. Das Clustering der varianznormierten Tageskurven führt hier zu $C_{opt} = 9$, das Clustering der mittelwertnormierten Tageskurven führt hier zu $C_{opt} = 5$, wobei in allen drei Fällen eines der Cluster ein Rauschcluster ist.

Die **preisbeeinflussten Typtage** der drei Fälle DT , DT_{norm} und $DT_{normmean}$ sind in Abb. 4.6a(i)-(iii) dargestellt. Analog zum Clustering der DH- und DH^{wo} -Fälle zeigen auch hier die Clusterkurven des nicht-normierten Falls DT am besten saisonale Verbrauchsunterschiede ($Q_A = 0.62$ für Ausgangsgröße „Monat“ in Tabelle 4.2). Cluster C1 ist ein deutliches „Sommercluster“, wohingegen das Cluster C3 als „Wintercluster“ betrachtet werden kann. Zudem ist in Abb. 4.4a zu erkennen, dass das Cluster C2 aus Tageszeitreihen bestehen, die in den Monaten Februar, März, Oktober und November gemessen wurden.

In Abb. 4.4b ist die Verbundverteilung zwischen der Ausgangsgröße „DT_{normmean}-Cluster“ und „Werktag“ dargestellt. Der Wert für Q_A mit $Q_A = 0.13$ belegt, dass wochentagsbedingte Unterschiede durch die Clusterkurven der DT_{normmean}-Falls teilweise erklärbar sind. So bilden die Cluster C3 und C4 das preisbeeinflusste Verbrauchsverhalten an Werktagen ab, wohingegen das Cluster C1 hauptsächlich aus Wochenendtageskurven besteht. Für das Cluster C2 kann keine eindeutige Zuordnung diesbezüglich getroffen werden, da sowohl Werktage als auch Wochenendtage im Cluster enthalten sind.

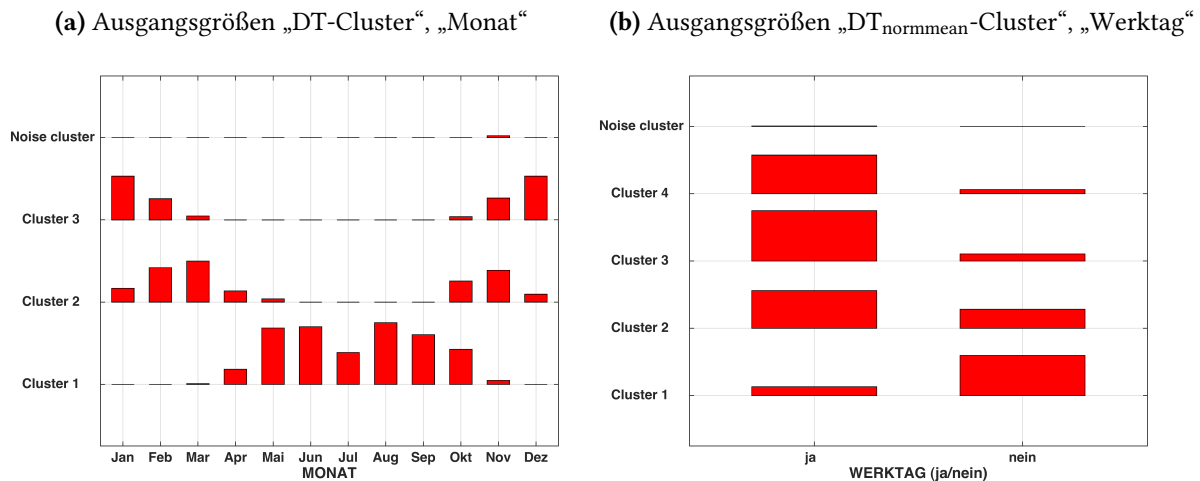


Abbildung 4.4: Verbundverteilungen zweier Ausgangsgrößen des OPP-Datensatzes als Balkendiagramme: (a) Ausgangsgröße „DT-Cluster“ und „Monat“, (b) Ausgangsgröße „DT_{normmean}-Cluster“ und „Werktag“

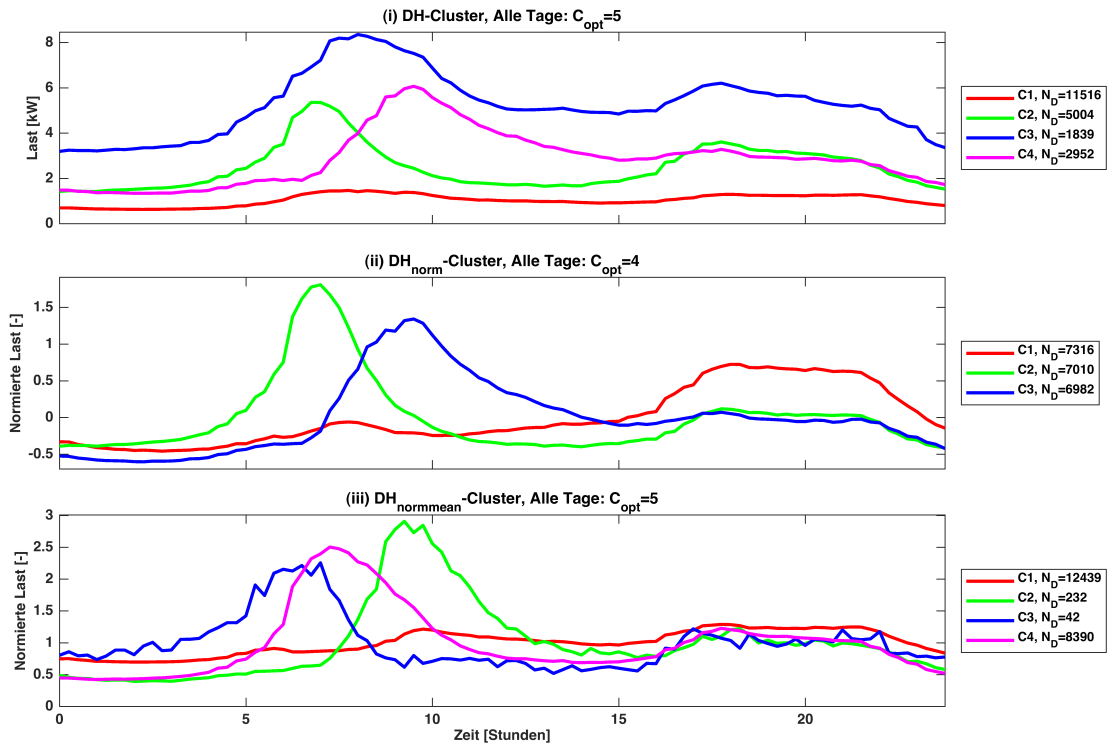
4.1.3.4 DT^{wo}-, DT_{norm}^{wo}- und DT_{normmean}^{wo}-Cluster als preisbeeinflusste Typtage (Nur Werktage)

In den Fällen DT^{wo}, DT_{norm}^{wo} und DT_{normmean}^{wo} sind Wochenendtage ausgeschlossen, was dazu führt, dass der in Abb. 4.6a und Abb. 4.4b gezeigte Einfluss der Werk- bzw. Wochenendtage für die nun folgende Clusteranalyse keine Rolle spielt. Das Ergebnis der Bestimmung von optimalen Clusterparametern steht in Tabelle 4.1. Für den Fall DT^{wo} ergibt sich eine optimale Clusteranzahl von $C_{opt} = 4$. Das Clustering der varianznormierten Tageskurven führt hier zu $C_{opt} = 4$, das Clustering der mittelwertnormierten Tageskurven führt hier zu $C_{opt} = 8$, wobei in allen drei Fällen eines der Cluster ein Rauschcluster ist.

Die **preisbeeinflussten Typtage** der Fälle DT^{wo}, DT_{norm}^{wo} und DT_{normmean}^{wo} sind in Abb. 4.6b(i)-(iii) dargestellt. Die Clusterkurven in Abb. 4.6b unterscheiden sich von den entsprechenden Clusterkurven in Abb. 4.6a insofern, als die Fälle DT_{norm} und DT_{normmean} eine andere optimale Clusteranzahl haben als die Fälle DT_{norm}^{wo} und DT_{normmean}^{wo}.

Auch hier gilt, dass sich monatliche Verbrauchsunterschiede sehr gut durch die Clusterkurven des DT^{wo}-Falls erklären lassen ($Q_A = 0.63$). Das Clustering der DT_{norm}^{wo}- und DT_{normmean}^{wo}-Fälle führt ebenfalls zu Clusterkurven, die den saisonalen Einfluss auf das preisbeeinflusste Verbrauchsverhalten gut erklären ($Q_A = 0.35$ bzw. $Q_A = 0.31$).

(a) Preisbeeinflusste Typtage der (i) DH-Cluster, (ii) DH_{norm} -Cluster und (iii) $DH_{normmean}$ -Cluster



(b) Preisbeeinflusste Typtage der (i) DH^{wo} -Cluster, (ii) DH_{norm}^{wo} -Cluster und (iii) $DH_{normmean}^{wo}$ -Cluster

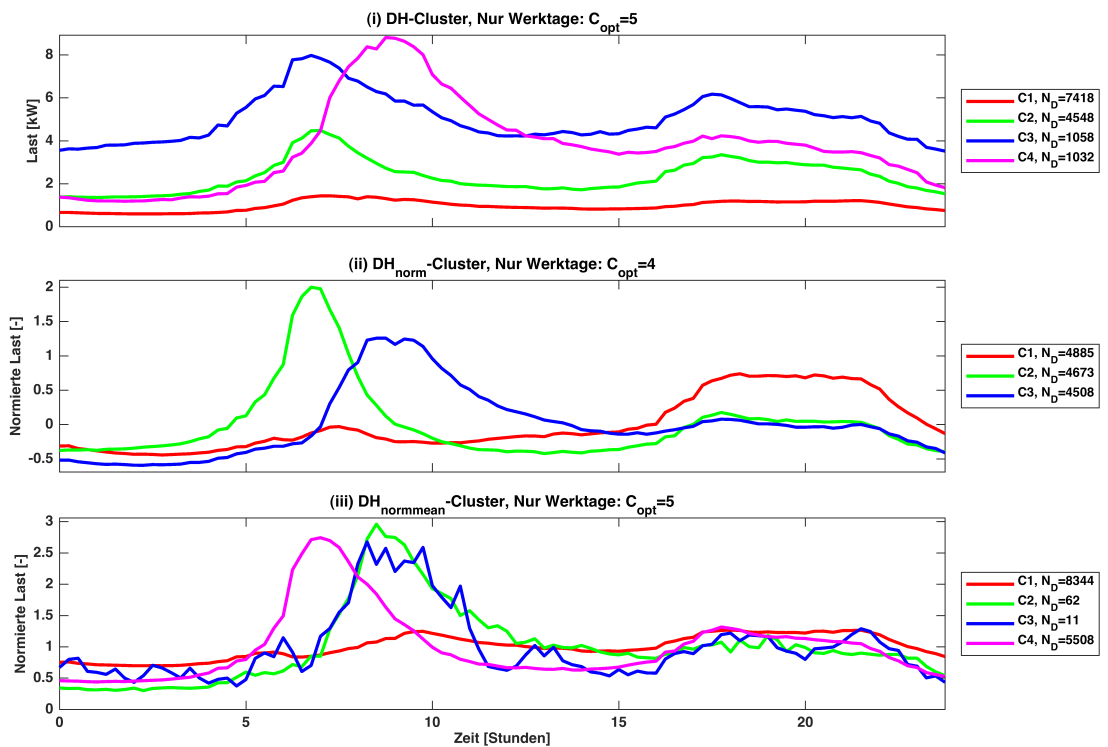
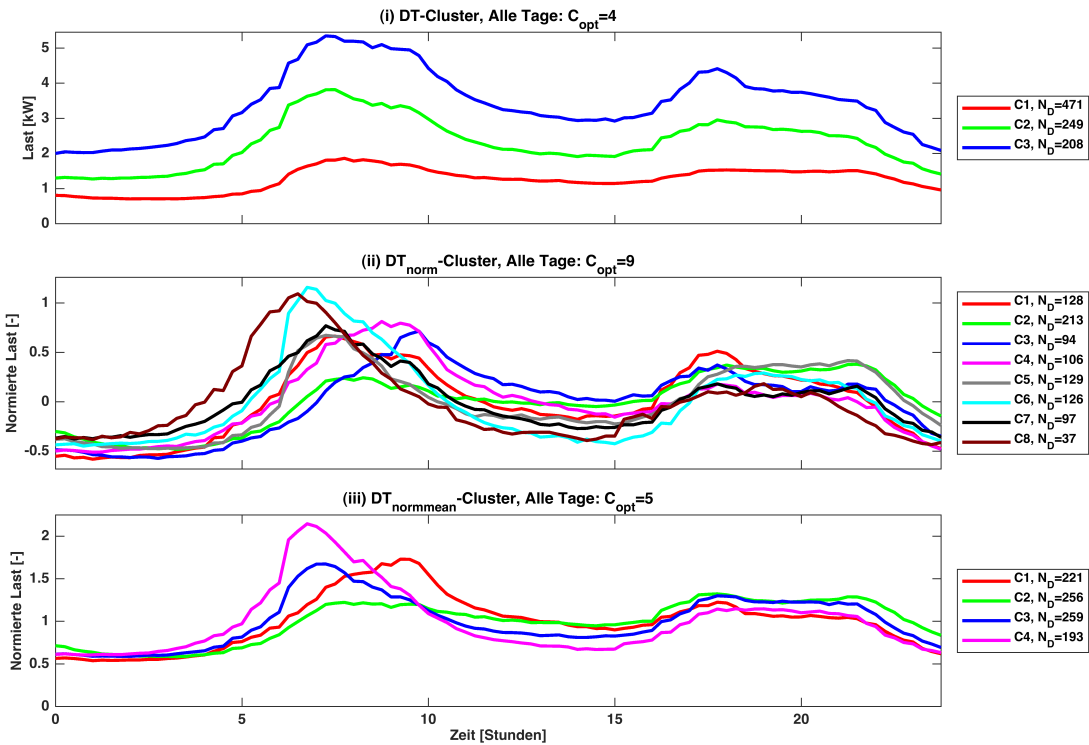
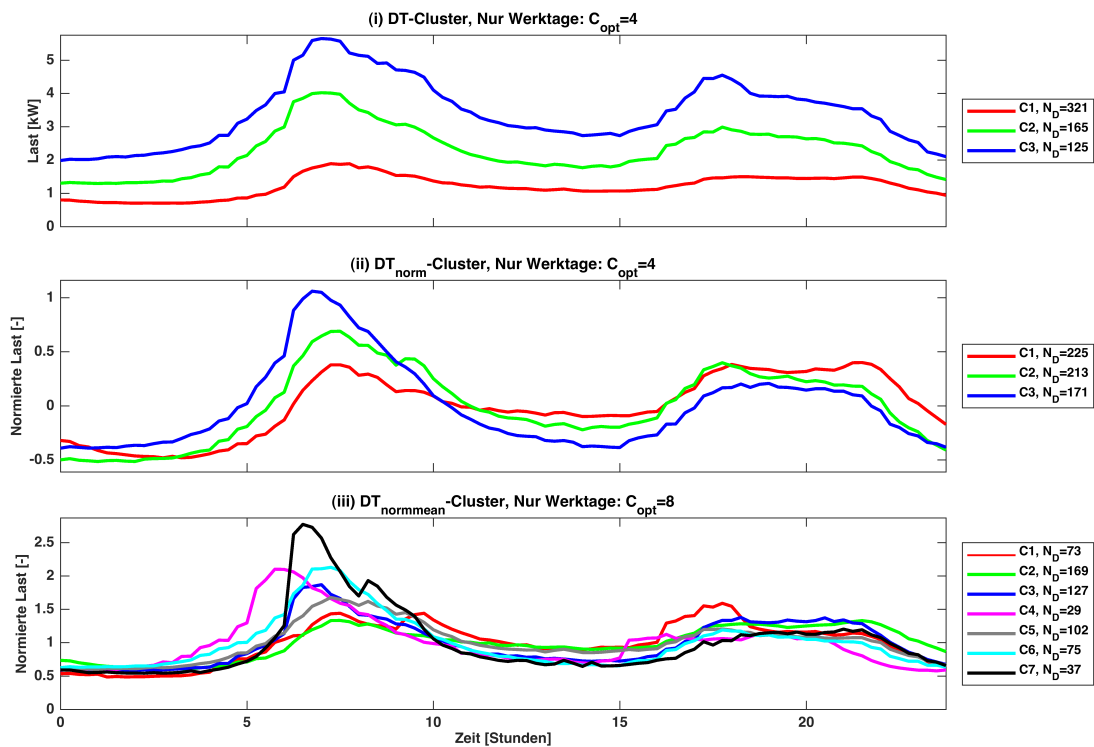


Abbildung 4.5: Gefundene DH-Cluster des OPP-Datensatzes: (a) alle Tage, (b) nur Werktage. N_D : Anzahl an Datentupeln je Cluster.

(a) Preisbeeinflusste Typtage der (i) DT-Cluster, (ii) DT_{norm} -Cluster und (iii) $DT_{normmean}$ -Cluster(b) Preisbeeinflusste Typtage der (i) DT^{wo} -Cluster, (ii) DT_{norm}^{wo} -Cluster und (iii) $DT_{normmean}^{wo}$ -ClusterAbbildung 4.6: Gefundene DT-Cluster des OPP-Datensatzes: (a) alle Tage, (b) nur Werktage. N_D : Anzahl an Datentupeln je Cluster.

4.1.4 Aufgabenstellung A2: Identifikation von tarifbedingten Unterschieden für eine Demand-Response-Maßnahme

Es sei nochmals an die Aufgabenstellung A2 aus Abschnitt 2.2 erinnert: „Ein Echtzeitstrompreis (vgl. Abb. 1.5c) bewirkt eine andere Verbrauchsreaktion als ein Zweistufentarif (vgl. Abb. 1.5a), da die Preisstruktur des Zweistufentarifs i.d.R. im Vorfeld bekannt ist und der Haushaltsstromkunde im Sinne einer Tagesoptimierung auf das Preissignal reagiert. Das neue Konzept soll den unterschiedlichen Stromtarifen Rechnung tragen und Unterschiede in der Verbrauchsreaktion, hervorgerufen durch die verschiedenen Stromtarifstrukturen, erklären.“

Im vorigen Abschnitt wurden die zwölf Fälle des OPP-Datensatzes mit dem in Abschnitt 2.5.1 vorgestellten Clusteranalyse-Konzept untersucht. In Abschnitt 4.1.4 findet nun kein erneutes Clustering der zwölf Fälle statt. Vielmehr werden die Clusterkurven hinsichtlich der neuen Aufgabenstellung bewertet, wofür die Verfahren aus Abschnitt 2.5.3 zum Einsatz kommen. Den größten Zusammenhang zwischen den gefundenen Clusterkurven und den Klassen der Ausgangsgröße „Tarifgruppe“ liefert das Clustering des Falls $DT_{\text{normmean}}^{\text{wo}}$ ($Q_A = 0.12$). Die Werte der relativen Transinformation pro Ausgangsentropie für die Ausgangsgröße „Tarifgruppe“ in Tabelle 4.2 belegen allerdings, dass die Clusterkurven der nicht-aggregierten Fälle tarifbedingte Verbrauchsunterschiede nicht erklären. Die aggregierten Clusterkurven sind bezüglich des Einflusses des Tarifmodells etwas aussagekräftiger als die nicht-aggregierten Clusterkurven. Dennoch sind die Werte für Q_A für die Ausgangsgröße „Tarifgruppe“ sehr gering.

Die Verteilung der einzelnen Tageszeitreihen der gefundenen Clusterkurven aus Abb. 4.6b(iii) ist in Abb. 4.7 als Heatmap gezeigt.

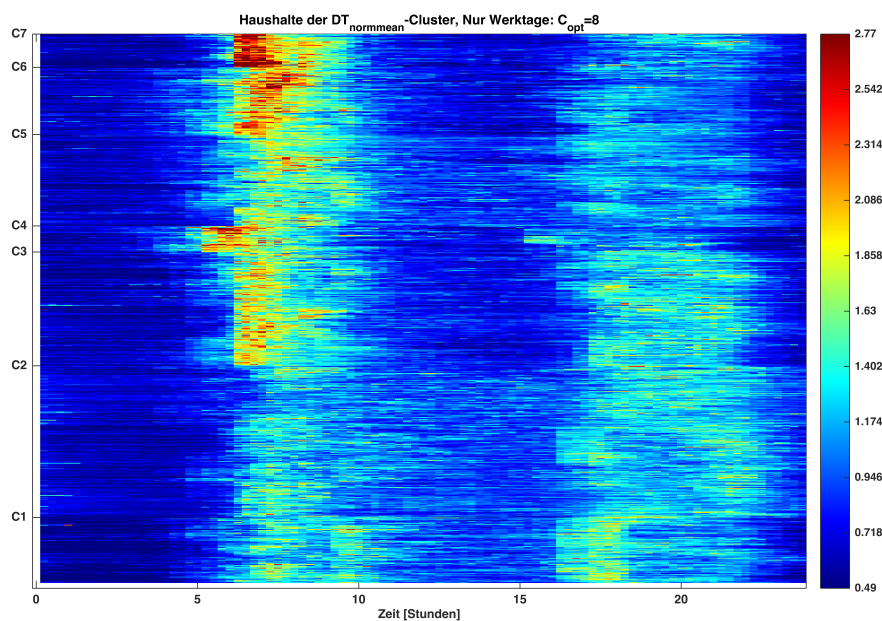


Abbildung 4.7: Verteilung der Typtage für Fall $DT_{\text{normmean}}^{\text{wo}}$ des OPP-Datensatzes mit farbkodiertem normierten Verbrauch

Beim Betrachten der Clusterkurven in Abb. 4.6b(iii) und der Tageskurven in Abb. 4.7 zeigt sich, dass das Cluster C1 ein dem TOU-Tarif entsprechendes Verbrauchsverhalten aufweist. Cluster C1 zeigt eine deutliche Lasterhöhung nach der Hochpreiszeit am Morgen; eine nach vorne verlagerte Abendlastspitze und eine reduzierte Last zu Hochpreiszeiten am Abend. Diese Beobachtungen und die daraus abgeleiteten Hypothesen über das preisbeeinflusste Verbrauchsverhalten werden durch Abb. 4.8 bestätigt. Cluster C1 setzt sich hauptsächlich aus TOU-Haushalten zusammen.

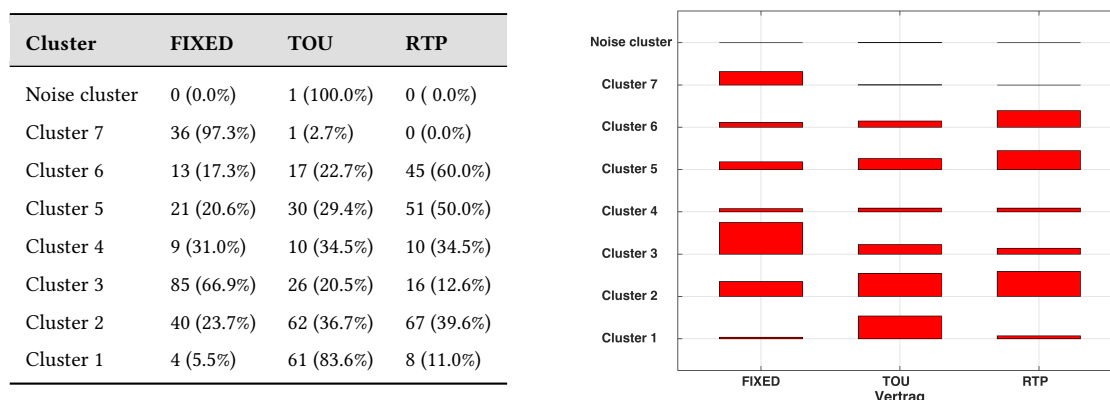


Abbildung 4.8: Verbundverteilung der Ausgangsgrößen „DT_{normmean}^{wo}-Cluster“ und „Tarifgruppe“ des OPP-Datensatzes: (links) als Tabelle mit der Anzahl an Datentupeln und der prozentualen Verteilung der Datentupel je Zeile, (rechts) als Balkendiagramme

Cluster C3 und C7 haben eine ausgeprägte Lastspitze zu Hochpreiszeiten am Morgen und keine reduzierte Last zu Hochpreiszeiten am Abend, was vermuten lässt, dass diese Cluster hauptsächlich aus Haushalten der FIXED-Tarifgruppe bestehen. Die entsprechenden Werte in Abb. 4.8 bestätigen diese Annahme.

Wie in Abschnitt 2.5.5 beschrieben, lassen sich zeitliche Lastverschiebungen durch die Analyse der Fälle SH_{normmean} beurteilen. In Tabelle 4.3 ist die Auswertung der Preisklassen für den TOU-Tarif bzw. RTP-Tarif gegenüber der Kontrollgruppe (FIXED) dargestellt. Hierfür wurde die statistische Signifikanz von Mittelwertunterschieden mithilfe des Zweistichproben-Welch-Tests bewertet. Die Höhe der Signifikanz der Ergebnisse ist in beiden Tabellen mit Sternen gekennzeichnet: signifikante Ergebnisse ($\alpha < 0.05$) sind mit *, hochsignifikante Ergebnisse ($\alpha < 0.01$) mit ** und höchstsignifikante Ergebnisse ($\alpha < 0.001$) mit *** gekennzeichnet.

Tabelle 4.3: Auswertung der Preisklassen (PK in ct/kWh) (links) für den TOU-Tarif gegenüber der Kontrollgruppe und (rechts) für den RTP-Tarif gegenüber der Kontrollgruppe für Fall SH_{normmean} des OPP-Datensatzes; N_D : Anzahl der Datentupel pro Gruppe; Angabe der mittleren normierten Leistung, Standardabweichung und Datentupelanzahl je Preisklasse und Gruppe; Zweistichproben-Welch-Test, zweiseitig: FIXED vs. TOU und FIXED vs. RTP

PK	N_D	FIXED	TOU	PK	N_D	FIXED	RTP
4.12	405552	0.870+-0.799	0.923+-0.909***	ca. 2.07	212491	0.954+-1.027	0.954+-1.020
5	102384	0.965+-1.070	0.937+-0.961***	ca. 5.48	438780	1.009+-0.965	1.006+-0.935
12.15	135312	1.390+-1.309	1.232+-1.190***	ca. 11.01	12661	1.199+-1.013	1.203+-1.050
13.5	34128	1.104+-0.890	1.189+-1.148***	ca. 20.39	7576	1.120+-0.910	1.150+-0.976*
35	384	0.982+-0.788	0.604+-0.583***	ca. 36.94	6252	1.381+-1.229	1.310+-1.304**

Die Ergebnisse in Tabelle 4.3 belegen, dass ein sehr hoher TOU-Tarif von 35 ct/kWh zu einer starken Lastabsenkung von ca. 38% führt. Bei der Betrachtung der Preisklassen des TOU-Wintertarifs (4,12 ct/kWh und 12,15 ct/kWh, siehe Abb. 4.1) fällt auf, dass die Lastverschiebung ca. 15% des Tagesdurchschnittes beträgt. Ein gänzlich anderes Verhalten zeigen die Haushalte im Sommer. Der TOU-Sommertarif (5,0 ct/kWh und 13,5 ct/kWh, siehe Abb. 4.1) führt dazu, dass die durchschnittliche Leistung der TOU-Haushalte in den Hochtarifzeiten um ca. 8% höher ist als die durchschnittliche Leistung der FIXED-Haushalte. Es gilt zu beachten, dass die Datenmenge für den TOU-Sommertarif deutlich geringer ist als bei der entsprechenden Hochtarifstufe im Winter ($N_D = 34128$ zu $N_D = 135312$). Eine mögliche Erklärung für das untypische Verhalten („Mehrverbrauch in der Hochtarifzeit“) der TOU-Haushalte im Sommer ist das Fehlen einer hohen Tarifstufe am Morgen. Ein TOU-Tarif, der Haushaltsstromkunden morgens und abends die Möglichkeit zur Lastverschiebung bietet, scheint geeigneter zu sein als ein TOU-Tarif, der lediglich abends Anreize zur Lastverlagerung schafft. Eine weitere Erklärung für das untypische Verhalten zielt auf den Beginn der Hochpreisphase und die Verdopplung der Hochpreisphase von drei auf sechs Stunden in den Abendstunden (15:00-21:00) ab. Ein TOU-Tarif mit einer Hochpreisphase von sechs Stunden am Abend führt scheinbar zu falschen Anreizen zur Lastverschiebung.

Die Auswertung der Preisklassen für den RTP-Tarif² in Tabelle 4.3 zeigt, dass nur geringe Unterschiede in den Mittelwerten der Preisklassen zu finden sind. Einzig die sehr hohe Preisklasse von 36,94 ct/kWh bewirkt eine (höchst-) signifikante Lastabsenkung von ca. 7% im Vergleich zu den FIXED-Haushalten. Jedoch liegt der Stromverbrauch der RTP-Haushalte in der zweithöchsten Preisklasse (ca. 20,39 ct/kWh) 3% höher als der Stromverbrauch der FIXED-Haushalte zu gleichen Zeiten. Das untypische Verhalten der RTP-Haushalte („Mehrverbrauch in der zweithöchsten Preisklasse“) scheint in dem schnell veränderlichen Strompreis und der Geräteausstattung begründet zu sein. Ein Echtzeitstrompreis erfordert, dass zu jeder Zeit ein möglichst großer Anteil der Haushaltslast verschiebbar ist. Allerdings limitieren elektrische Geräte, deren Leistungsaufnahme nicht unterbrech- oder verschiebbar ist (z.B. gestartetes Waschprogramm einer Waschmaschine), den Anteil der verschiebbaren Last, was im Extremfall zu dem oben genannten untypischen Verbrauchsverhalten führen kann.

4.1.5 Aufgabenstellung A3: Identifikation von (Non-) Responder-Haushalten

Es sei nochmals an die Aufgabenstellung A3 aus Abschnitt 2.2 erinnert: „Erfolgt die Verbrauchsreaktion auf ein Preissignal durch manuelles Ein- und Ausschalten bestimmter Haushaltsgeräte, spielt die Anwesenheit des Menschen und die individuelle Bereitschaft zur Laständerung eine entscheidende Rolle. Mithilfe des neuen Konzeptes sollen haushaltsindividuelle Unterschiede in der Verbrauchsreaktion analysiert und sogenannte „Responder-Haushalte“ (Abk.: R-Haushalt; Haushalt reagiert stark auf Preissignal), „Non-Responder-Haushalte“ (Abk.: NR-Haushalt; Haus-

²Die Preisklassen des RTP-Tarifs werden durch Fuzzifizierung der Preise $p[k]$ des RTP-Tarifs gebildet. Unter Fuzzifizierung ist die graduelle Zuordnung dieser Werte zu einem von fünf linguistischen Termen, wie PVS (*Positive Very Small*), PS (*Positive Small*), PM (*Positive Mean*), PB (*Positive Big*) oder PVB (*Positive Very Big*), gemeint (vgl. Abb. 2.11).

halt reagiert nicht bzw. schwach auf Preissignal) und „Semi-Responder-Haushalte“ (Abk.: SR-Haushalt; Haushalt reagiert mäßig auf Preissignal) identifiziert werden.“

Der unterschiedliche Verlauf des TOU-Tarifs im Winter (Standardtarif) und im Sommer rechtfertigt eine weitere Unterteilung der $DH_{\text{TOU,Standard}}^{\text{wo}}$ -, $DH_{\text{TOU,Standard,norm}}^{\text{wo}}$ - und $DH_{\text{TOU,Standard,normmean}}^{\text{wo}}$ -Fälle. Es sind nur diejenigen Tage ausgewählt, die die gleiche **TOU-Standardtarifstruktur** besitzen. Die Anzahl der Datentupel reduziert sich so auf $N_D = 3751$. Die optimalen Clusterparameter (siehe Tabelle B.13) wurden ebenfalls mit einer Sensitivitätsanalyse bestimmt. Die entsprechenden Heatmaps sind in Abb. B.23, B.24 und B.25 zu finden. Die **preisbeeinflussten Typtage** der drei Fälle sind in Abb. 4.9(i)-(iii) zu sehen.

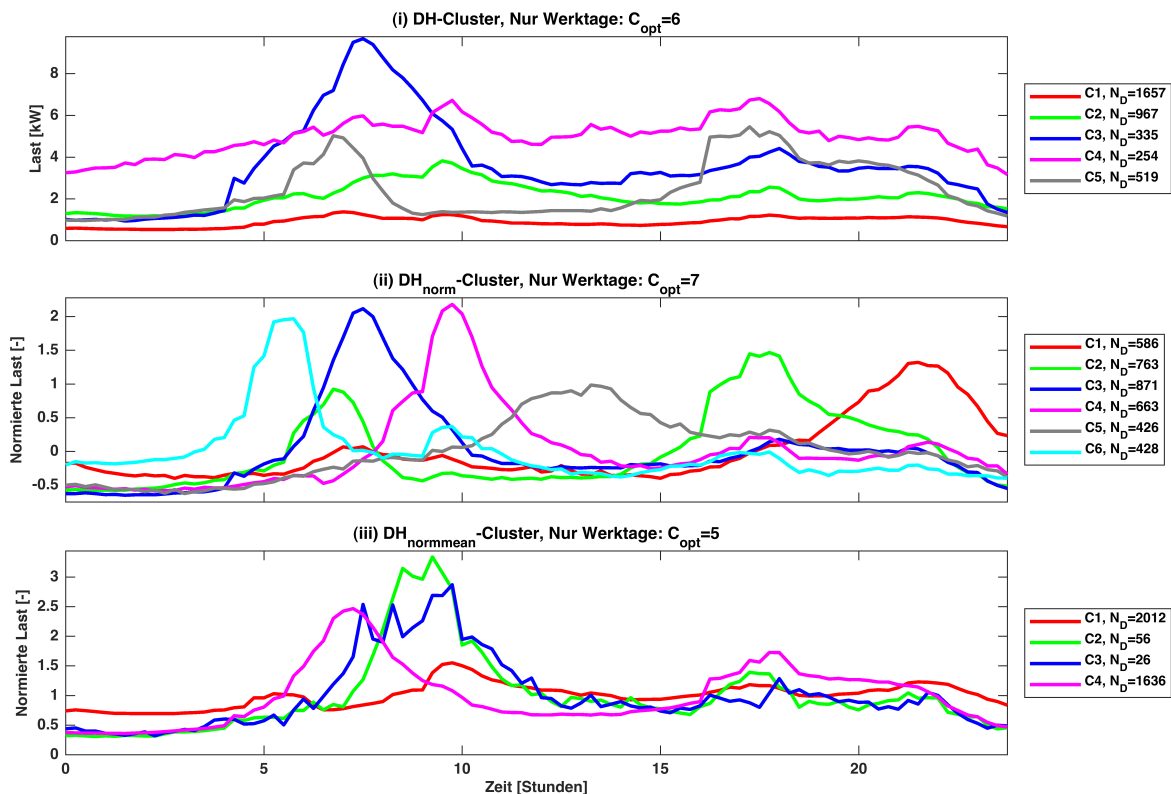


Abbildung 4.9: Gefundene $DH_{\text{TOU,Standard}}^{\text{wo}}$ -Cluster des OPP-Datensatzes: Preisbeeinflusste Typtage der (i) $DH_{\text{TOU,Standard}}^{\text{wo}}$ -Cluster, (ii) $DH_{\text{TOU,Standard,norm}}^{\text{wo}}$ -Cluster und (iii) $DH_{\text{TOU,Standard,normmean}}^{\text{wo}}$ -Cluster N_D : Anzahl an Datentupeln je Cluster. Es sind nur diejenigen Datentupel ($N_D = 3751$) ausgewählt, die die gleiche TOU-Standardtarifstruktur besitzen.

Die Identifikation von (Non-) Responder-Haushalten erfolgt durch visuelle Inspektion der gefundenen Clusterkurven in Abb. 4.9. Das Ergebnis der Identifikation von (Non-) Responder-Haushalten ist in Tabelle 4.4 zusammengefasst. Auf Basis der Clusterkurven wurde eine subjektive Einstufung der Cluster in Responder-, Semi-Responder- und Non-Responder-Haushalte vorgenommen. Cluster, bei denen keine klare Einstufung vorgenommen werden konnte, wurden in Tabelle 4.4 mit mehreren Kreuzen gekennzeichnet. Im Folgenden wird die Identifikation von (Non-) Responder-Haushalten am Beispiel der Clusterkurven in Abb. 4.9(ii) diskutiert.

Obwohl alle Haushalte das gleiche Preissignal (TOU-Standardtarifstruktur im Winter) bekommen und der Einfluss unterschiedlicher Tagesroutinen an Werk- und Wochenendtagen ausgeschlossen ist, unterscheiden sich die gefundenen Clusterkurven stark voneinander. Besonders deutlich wird dies bei den Clusterkurven in Abb. 4.9(ii) sichtbar. Die Cluster C4 und C6 zeigen eine klare Lastreduktion zu Zeiten des Hochtarifs am Morgen/Abend und einer Lasterhöhung kurz vor oder kurz nach dem Hochpreistarif (Responder-Haushalte). Die Cluster C1 und C2 zeigen ein ähnliches Verhalten, wobei die Spitzenlast nicht am Morgen, sondern am Abend auftritt. Cluster C2 hat zudem eine ausgeprägte Lastspitze zu Zeiten des Hochtarifs am Morgen. Das differenzierte Verhalten von Cluster C2 (keine Lastverschiebung am Morgen, Lastverschiebung am Abend) führt dazu, dass dieses Cluster nicht eindeutig den Responder- oder den Semi-Responder-Haushalten zugeordnet werden kann. Cluster C5 hat eine ausgeprägte Mittagslastspitze und eine geringe Last zu Zeiten des Hochtarifs am Morgen/Abend (Responder-Haushalte). Für die Haushalte des Clusters C3 scheint der hohe Strompreis am Morgen kein passendes Preissignal zu sein, um die Last zu dieser Zeit zu senken (Non-Responder-Haushalte). Alles in allem zeigen ca. 77% der Datentupel³ in Abb. 4.9(ii) eine erwartbare Reaktion auf das TOU-Strompreissignal. Die Einstufung der $DH_{TOU,Standard}^{wo}$ - und $DH_{TOU,Standard,normmean}^{wo}$ -Cluster in Responder-, Semi-Responder- und Non-Responder-Haushalte ist in Tabelle 4.4(a) bzw. Tabelle 4.4(c) gezeigt, wird aber nicht ausführlich diskutiert.

Tabelle 4.4: Einstufung der $DH_{TOU,Standard}^{wo}$ -Cluster in Responder-, Semi-Responder- und Non-Responder-Haushalte. Die Einstufung erfolgt durch visuelle Inspektion der Clusterkurven in Abb. 4.9. Cluster, bei denen keine klare Einstufung vorgenommen werden konnte, werden mit mehreren Kreuzen gekennzeichnet. Abkürzungen: C - Cluster, RC - Rauschcluster, RH - Responder-Haushalt, SRH - Semi-Responder-Haushalt, NRH - Non-Responder-Haushalt.

(a) $DH_{TOU,Standard}^{wo}$ -Cluster					(b) $DH_{TOU,Standard,norm}^{wo}$ -Cluster					(c) $DH_{TOU,Standard,normmean}^{wo}$ -Cluster				
C	N_D	RH	SRH	NRH	C	N_D	RH	SRH	NRH	C	N_D	RH	SRH	NRH
C1	1657	x	x	x	C1	586	x			C1	2012	x		
C2	967	x			C2	763	x	x		C2	56		x	
C3	335			x	C3	871			x	C3	26		x	
C4	254	x			C4	663	x			C4	1636		x	x
C5	519	x	x		C5	426	x			RC	21			
RC	19				C6	428	x							
					RC	14								

4.1.6 Aufgabenstellung A4: Prognose des preisbeeinflussten Verbrauchsverhaltens

Es sei nochmals an die Aufgabenstellung A4 aus Abschnitt 2.2 erinnert: „Die Ergebnisse aus A1 bis A3 haben analytischen Charakter und eignen sich daher nur begrenzt für den Einsatz

³Die Angabe von 77% bezieht sich hier auf die Anzahl der betrachteten Tage und nicht auf die Anzahl der betrachteten Haushalte. Es wird dazu geraten, weitergehende Analysen durchzuführen, um Aussagen bezüglich der Anzahl an (Non-) Responder-Haushalten treffen zu können.

als Prognosemodell zur Vorhersage des preisbeeinflussten Verbrauchsverhaltens von Haushaltsstromkunden. Aus diesem Grund soll, in Ergänzung zu den Aufgabenstellungen A1 bis A3, das preisbeeinflusste Verbrauchsverhaltens mit einem systemtheoretischen *Demand-Response*-Modell beschrieben werden. Das zu entwickelnde Prognosemodell (vgl. Gleichung (2.4)) soll die haushaltsindividuelle Verbrauchsreaktion auf ein bestimmtes Preissignal beschreiben. Das Prognosemodell soll in der Lage sein, die analytischen Ergebnisse der Aufgabenstellung A1 bis A3 bei der Modellentwicklung zu berücksichtigen. Des Weiteren soll das Prognosemodell in komplexen Simulationsszenarien als Simulationsmodell eingesetzt werden.“

Die in Abschnitt 2.6.2 entwickelten Virtuelle-Speicher-Modelle sollen nun verwendet werden, um das preisbeeinflusste Verbrauchsverhalten der Haushaltsstromkunden zu modellieren. Der Ablauf der Modellierung wurde bereits in Abb. 2.24 beschrieben, sodass an dieser Stelle auf die Ergebnisse eines Modellvergleichs zwischen den neu entwickelten VS-Modellen und etablierten dynamischen Regressionsmodellen (ARX-Modell) eingegangen wird. Wie in Abb. 2.24 gezeigt, kann die DR-Modellierung anhand verschiedener Datensätze und Aggregationsniveaus durchgeführt werden, wie beispielsweise VT- oder DT-Fälle (vgl. Tabelle 2.3). Aus diesem Grund sind die folgenden Ergebnisse in zwei Abschnitte gegliedert: erstens, die Ergebnisse der Modellierung auf Basis der VT-Fälle (siehe Abschnitt 4.1.6.1) und zweitens, die Ergebnisse der Modellierung auf Basis der DT-Fälle des OPP-Datensatzes (siehe Abschnitt 4.1.6.2). In Abschnitt 4.1.6.3 wird eine qualitative Bewertung der Ergebnisse vorgenommen.

Zur Bewertung des Modellverhaltens und der Modellgüte werden aus der Vielzahl an regressionsorientierten Bewertungsmaßen (siehe Abschnitt 2.6.5) der mittlere absolute Fehler $Q_{C,MAE}(P_R, \hat{P}_R)$ (siehe Gleichung (2.83)) und der Korrelationskoeffizient $r(P_R, \hat{P}_R)$ (siehe Gleichung (2.88)) zwischen der Zielvariablen P_R und der Modellschätzgröße \hat{P}_R ausgewählt.

4.1.6.1 Ergebnis der Modellierung der VT-Fälle

Die Ergebnisse der hier gezeigten Modellierung basieren auf den VT-Fällen des OPP-Datensatzes, die in Tabelle 4.5 gezeigt sind. Die VT-Fälle stellen die größtmöglichen Zeiträume ohne fehlende Werte dar. Der Modellentwurf und die Modellanwendung für die VT-Fälle wurde auf Basis des in Abb. 2.24 gezeigten Prozesses durchgeführt. Für den Modellvergleich wurden alle vier VS-Modelle (VSA, VSA*, VSB, VSB*) und ein ARX-Modell eingesetzt. Die Modelle wurden mit einem der vier VT^{TOU}-Fälle (z.B. Fall VT^{TOU}_{Sommer}) entworfen und die restlichen drei Fälle (z.B. Fälle VT^{TOU}_{Herbst}, VT^{TOU}_{Winter} und VT^{TOU}_{Frühling}) dienten jeweils als Testdatensatz zum Testen der entworfenen Modelle. Der beschriebene Modellierungsvorgang wurde insgesamt dreimal wiederholt, bis jeder der vier VT^{TOU}-Fälle einmal als Lerndatensatz fungierte. Diese vierstufige Kreuzvalidierung wurde in analoger Weise für die VT^{RTP}-Fälle durchgeführt.

Bei der Einteilung des OPP-Datensatzes in die vier VT-Fälle muss angemerkt werden, dass sich das Strompreisprofil des Falls VT^{TOU}_{Sommer} von dem der Fälle VT^{TOU}_{Herbst}, VT^{TOU}_{Winter} und VT^{TOU}_{Frühling} unterscheidet, was bereits in Abb. 4.1 gezeigt wurde.

Tabelle 4.5: Strukturierung und Einteilung des OPP-Datensatzes in VT-Fälle für die Aufgabenstellung A4; D - Anzahl an Tagen, K - Zeitreihenlänge, N_D - Anzahl an Datentupeln

Saison	FIXED-Haushalte	TOU-Haushalte	RTP-Haushalte
Frühling	Fall $VT_{\text{Frühling}}^{\text{FIXED}}$ ($N_D = 3, D = 60, K = 5760$)	Fall $VT_{\text{Frühling}}^{\text{TOU}}$ ($N_D = 3, D = 60, K = 5760$)	Fall $VT_{\text{Frühling}}^{\text{RTP}}$ ($N_D = 3, D = 60, K = 5760$)
Sommer	Fall $VT_{\text{Sommer}}^{\text{FIXED}}$ ($N_D = 3, D = 65, K = 6240$)	Fall $VT_{\text{Sommer}}^{\text{TOU}}$ ($N_D = 3, D = 65, K = 6240$)	Fall $VT_{\text{Sommer}}^{\text{RTP}}$ ($N_D = 3, D = 65, K = 6240$)
Herbst	Fall $VT_{\text{Herbst}}^{\text{FIXED}}$ ($N_D = 3, D = 44, K = 4224$)	Fall $VT_{\text{Herbst}}^{\text{TOU}}$ ($N_D = 3, D = 44, K = 4224$)	Fall $VT_{\text{Herbst}}^{\text{RTP}}$ ($N_D = 3, D = 44, K = 4224$)
Winter	Fall $VT_{\text{Winter}}^{\text{FIXED}}$ ($N_D = 3, D = 64, K = 6144$)	Fall $VT_{\text{Winter}}^{\text{TOU}}$ ($N_D = 3, D = 64, K = 6144$)	Fall $VT_{\text{Winter}}^{\text{RTP}}$ ($N_D = 3, D = 64, K = 6144$)

Die Modellierungsergebnisse auf Basis der VT^{TOU} -Fälle sind in Abb. 4.10a als Boxplots der Werte des Korrelationskoeffizienten r und der Werte des mittleren absoluten Fehlers $Q_{C,MAE}$ dargestellt und lassen sich wie folgt zusammenfassen:

- Ein Korrelationskoeffizient von $r < 0.5$ bedeutet, dass eine schwache Korrelation zwischen der Zielvariablen P_R und der Modellschätzgröße \hat{P}_R besteht. Dies gilt für jeden der fünf DR-Modelltypen und kann durch die geringe Aggregationsstufe der VT-Fälle (d.h. VT-Fälle bestehen aus wenigen aber langen Zeitreihen) und durch die hohen Abweichungen in den aggregierten Lastkurven erklärt werden.
- Die VSA- und VSA*-Modelle beschreiben das preisbeeinflusste Verbrauchsverhalten der TOU-Tarifgruppe besser als die VSB- und VSB*-Modelle, da VSB- und VSB*-Modelle nicht für bekannte zukünftige Preise konzipiert wurden (vgl. Gleichungen (2.67) und (2.71)).
- VSA- und VSA*-Modelle erklären den Zusammenhang zwischen Preissignal und Verbraucherlast besser als aus der Literatur bekannte ARX-Modelle (vgl. Gleichung (1.4)).

Die Modellierungsergebnisse auf Basis der VT^{RTP} -Fälle sind in Abb. 4.10b dargestellt und lassen sich zusammenfassend folgendermaßen beschreiben:

- Die Werte des Korrelationskoeffizienten und des mittleren absoluten Fehlers sind für die VT^{RTP} -Fälle schlechter als für die VT^{TOU} -Fälle. Das Verbrauchsverhalten als Reaktion auf Echtzeitpreise lässt sich schlechter prognostizieren als das Verbrauchsverhalten als Reaktion auf den statischen TOU-Tarif.
- Unter allen getesteten Modelltypen eignen sich VSB-Modelle am besten, um den Zusammenhang zwischen Preissignal und Verbraucherlast zu modellieren. Allerdings werden lediglich 19% ($r = 0.19$ im Fall des besten VSB-Modells) der Gesamtvariation in \hat{P}_R durch den Zusammenhang zwischen Preissignal und Verbraucherlast erklärt. Die restlichen 81% der Gesamtvariation werden durch das Modell nicht abgebildet.
- Die Werte des Korrelationskoeffizienten und des mittleren absoluten Fehlers machen deutlich, dass der Modellvergleich anhand von Daten mit einem höheren Aggregationsni-

veaus (z.B. DT-Fälle) durchgeführt werden muss. Aus diesem Grund werden im Folgenden die Ergebnisse der Modellierung der DT-Fälle vorgestellt.

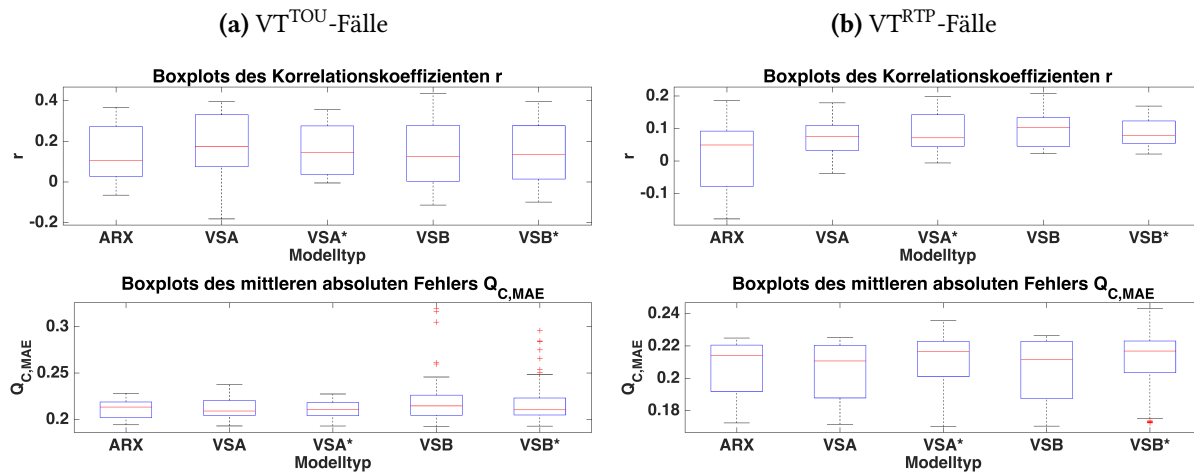


Abbildung 4.10: Modellvergleich verschiedener VS-Modelle mit ARX-Modellen auf Basis der VT-Fälle des OPP-Datensatzes (siehe Tabelle 4.5); Boxplots der Werte des Korrelationskoeffizienten r und Boxplots der Werte des mittleren absoluten Fehlers $Q_{C,MAE}$: (a) VT^{TOU} -Fälle und (b) VT^{RTP} -Fälle

4.1.6.2 Ergebnis der Modellierung der DT-Fälle

Die mäßigen Ergebnisse für die RTP-Tarifgruppe im vorigen Abschnitt rechtfertigen, dass die RTP-Tarifgruppe für den hier durchgeführten Modellvergleich ausgeschlossen wurde. Die Strukturierung und Einteilung des OPP-Datensatzes für den nun folgenden Modellvergleich ist in Tabelle 4.6 gezeigt. Aufgrund des unterschiedlichen Verlaufs des TOU-Strompreisprofils (TOU-Sommertarif, TOU-Standardtarif und CPP-Tarif; vgl. Abb. 4.1) werden die DT-Fälle aus Tabelle 2.3 in die DT-Subgruppen aus Tabelle 4.6 unterteilt. Der Entwurf und die Validierung der

Tabelle 4.6: Strukturierung und Einteilung des OPP-Datensatzes in DT-Fälle für die Aufgabenstellung A4; N_D - Anzahl an Datentupeln, K - Zeitreihenlänge

Tage	Alle Tage mit Sommer-TOU-Tarif	Alle Tage mit Standard-TOU-Tarif	Alle Tage mit CPP-Tarif
Werktage	Fall DT_{Sommer}^{wo} ($N_D = 3001, K = 96$)	Fall $DT_{Standard}^{wo}$ ($N_D = 10970, K = 96$)	Fall DT_{CPP}^{wo} ($N_D = 74, K = 96$)
Wochenendtage	Fall DT_{Sommer}^{we} ($N_D = 1427, K = 96$)	Fall $DT_{Standard}^{we}$ ($N_D = 5681, K = 96$)	–

DR-Modelle wurde, analog zum vorigen Abschnitt, nach dem in Abb. 2.24 dargestellten Prozess durchgeführt. Die Modelle wurden mit einem der fünf DT-Fälle aus Tabelle 4.6 entworfen und getestet. Für jeden der fünf Modelltypen ergeben sich so 25 separate Modelle, deren Bewertung

mithilfe der Bewertungsmaße r und $Q_{C,MAE}$ in Tabelle B.14 dargestellt ist. In der rechten Spalte ist für jeden Lerndatensatz der Mittelwert der Korrelationskoeffizienten und der mittleren absoluten Fehler angegeben. Bei der Mittelwertbildung wurden für jede Zeile diejenigen Werte ausgeschlossen, bei denen der Lern- und Testdatensatz identisch ist. In den diagonalen Zellen von Tabelle B.14 stehen die Bewertungsmaße jener Modelle, die mit dem selben Datensatz trainiert und getestet wurden.

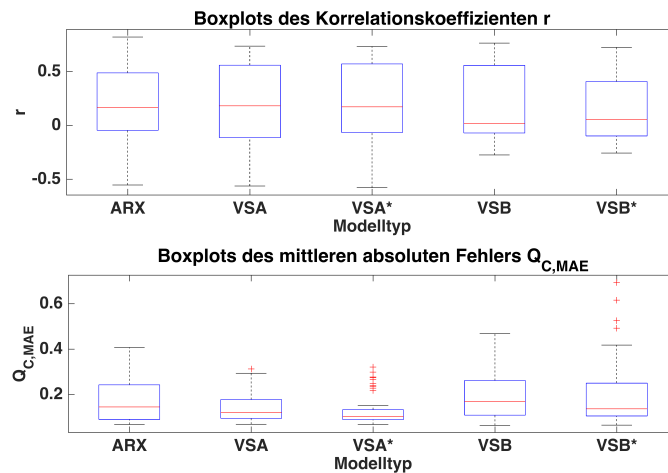


Abbildung 4.11: Modellvergleich verschiedener VS-Modelle mit ARX-Modellen auf Basis der DT-Fälle (siehe Tabelle 4.6) des OPP-Datensatzes: Boxplots der Werte des Korrelationskoeffizienten r und Boxplots der Werte des mittleren absoluten Fehlers $Q_{C,MAE}$.

Die Modellierungsergebnisse auf Basis der DT-Fälle sind in Abb. 4.11 als Boxplots der Werte des Korrelationskoeffizienten r und der Werte des mittleren absoluten Fehlers $Q_{C,MAE}$ dargestellt und lassen sich wie folgt zusammenfassen:

- Insgesamt weisen die DR-Modelle der DT-Fälle eine höhere Genauigkeit und Generalisierungsfähigkeit auf als die DR-Modelle der VT-Fälle. Dies lässt sich durch die höhere Aggregationsstufe und den damit einhergehenden geringeren Varianzen in den Lastkurven erklären.
- Für jeden Modelltyp entsteht das beste Modell aus dem Training mit einem der folgenden drei Fälle: $DT_{\text{Standard}}^{\text{wo}}$, $DT_{\text{Standard}}^{\text{we}}$ und $DT_{\text{CPP}}^{\text{wo}}$. Das Lernen der Modelle mit den Fällen $DT_{\text{Sommer}}^{\text{wo}}$ und $DT_{\text{Sommer}}^{\text{we}}$ führt zu eher schlechten Ergebnissen. Dies ist vor allem auf die höhere Anzahl an Datenpunkten zurückzuführen. Darüber hinaus ist durch das Fehlen der Hochtarifstufe am Morgen im Fall des TOU-Sommertarifs die Möglichkeit der Lastverlagerung reduziert.
- Auch hier gilt, dass sich die VSB-/VSB*-Modelle im Durchschnitt am wenigsten eignen, um den Zusammenhang zwischen Preissignal und Verbraucherlast zu schätzen, da VSB-/VSB*-Modelle nicht für bekannte zukünftige Preise konzipiert wurden (vgl. Gleichungen (2.67) und (2.71)).
- VSA- und VSA*-Modelle erklären den Zusammenhang zwischen Preissignal und Verbraucherlast im Durchschnitt besser als ARX-Modelle. Etwa 54% ($r = 0.54$ im Fall des besten

VSA-Modells) der Gesamtvariation in \hat{P}_R werden durch den Zusammenhang von P_R und \hat{P}_R erklärt. Jedoch wird die insgesamt höchste Modellgenauigkeit durch die Verwendung eines ARX-Modells erreicht.

4.1.6.3 Qualitative Bewertung der Ergebnisse des Modellvergleichs

Eine abschließende Bewertung aller in Abschnitt 4.1.6 durchgeführten Modellvergleiche liefert Tabelle 4.7. Die qualitative Bewertung der Kategorien „Genauigkeit“ und „Generalisierungsfähigkeit“ ergibt sich direkt aus den Ergebnissen des Modellvergleichs in Tabelle B.14 und Abb. 4.10a, 4.10b und 4.11. Eine gute Interpretierbarkeit der VS-Modelle im Vergleich zu den ARX-Modellen ergibt sich aus der geringen Anzahl von Modellparametern. Darüber hinaus sind die VS-Modelle *Gray-Box-Modelle* mit einer verständlichen inneren Modellstruktur, wie z.B. der Speicherfüllstand des virtuellen Speichers E_{VS} oder die Speicherbe- bzw. -entladegeschwindigkeit a . Im

Tabelle 4.7: Bewertung verschiedener *Demand-Response*-Modelltypen: ++: sehr geeignet, sehr nützlich; +: geeignet, nützlich; 0: wenig geeignet, wenig nützlich; -: nicht geeignet, nicht nützlich. Kategorie „Erweiterbarkeit“: Erweiterbarkeit wird durch die Integration von *a priori* Wissen erzielt.

Modelltyp	Genauigkeit (VT ^{TOU})		Genauigkeit (VT ^{RTP})		Genauigkeit (DT)		Generalisierungsfähigkeit	Interpretierbarkeit	Eignung für Einsatz in Simulationen	Erweiterbarkeit
	max.	im Mittel	max.	im Mittel	max.	im Mittel				
ARX	0	-	+	0	++	+	0	0	0	0
VSA	+	+	0	-	+	+	0	++	+	++
VSA*	0	+	+	0	+	++	+	++	+	++
VSB	++	0	++	+	+	+	0	++	+	++
VSB*	+	0	-	0	0	0	0	++	+	++

Rahmen einer Simulation ist es entscheidend, parametrierbare DR-Modelle zu verwenden, um die Modelle an bestimmte Rahmenbedingungen, wie z.B. Haushaltsgröße oder technische Geräteausstattung, anzupassen. Im Hinblick auf dieses Kriterium eignen sich VS-Modelle für die Simulation des preisbeeinflussten Verbrauchsverhaltens (siehe Abb. 2.2). Ein wesentlicher Vorteil aller VS-Modelle ist deren Erweiterbarkeit auf Basis der modularen Struktur. Die VS-Modelle lassen sich durch die Integration unterschiedlicher Einschränkungen und Modellannahmen leicht erweitern.

4.2 RESIDENS-Datensatz

4.2.1 Allgemeines zum Datensatz

Der zweite *Smart-Meter*-Datensatz, der mit dem neuen Konzept untersucht wurde, ist der RESIDENS⁴ Datensatz [77, 78]. Die *Demand-Response*-Feldstudie fand im Zeitraum November 2010 bis November 2011 in Thüringen statt. Die 189 teilnehmenden Haushalte erhielten alle den gleichen *Time-of-Use*-Tarif, wobei die Tarifstruktur von Montag bis Donnerstag, an Freitagen und von Samstag bis Sonntag konstant war (siehe Abb. 4.12).

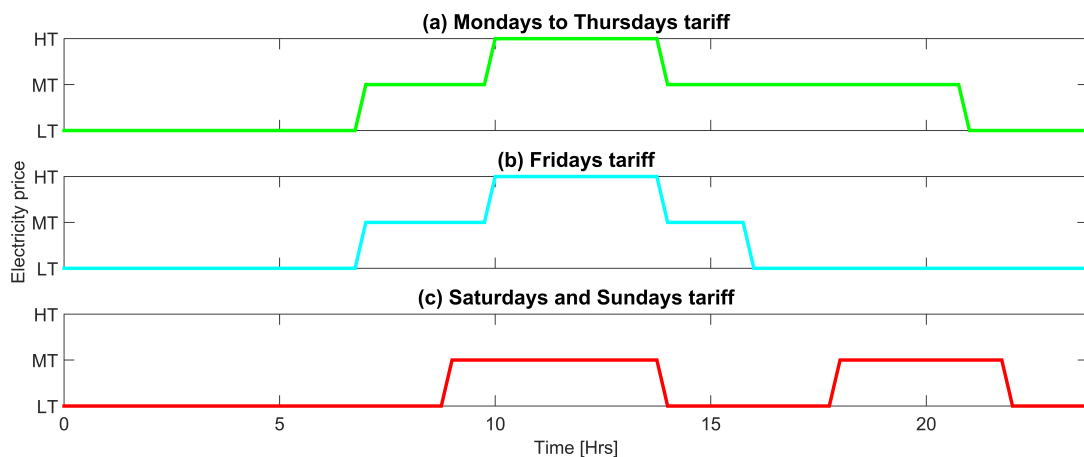


Abbildung 4.12: RESIDENS: Strompreisstufen (LT=low tariff, MT=medium tariff, HT=high tariff): (a) dreistufiger TOU-Tarif von Montag bis Donnerstag, (b) dreistufiger TOU-Tarif am Freitag, (c) zweistufiger TOU-Tarif von Samstag bis Sonntag.

Wie in Eigenschaft E10 in Tabelle 2.1 beschrieben, hat das Fehlen einer Kontrollgruppe für die Aufgabenstellung A1 einen geringen Einfluss, da auch ohne eine Kontrollgruppe Unterschiede identifiziert werden können, die auf wochentagsbedingte und saisonale Einflüsse zurückzuführen sind. Hinsichtlich der Aufgabenstellung A3 (vgl. Tabelle 2.2) gilt ähnliches. Eine Kontrollgruppe hilft zwar bei der Identifizierung von (Non-) Responder-Haushalten. Allerdings können unterschiedlich starke Verbrauchsreaktionen auch ohne eine Kontrollgruppe ausfindig gemacht werden. Somit kann der RESIDENS-Datensatz zur Bearbeitung der in den Abschnitten 4.2.3 und 4.2.4 gezeigten Aufgabenstellungen eingesetzt werden.

Im Vergleich zur OPP-Feldstudie war im RESIDENS-Projekt kein automatisches *Demand Response* möglich. Die Struktur des RESIDENS-Datensatzes ist in Tabelle B.15 zusammengefasst.

⁴Abkürzung für EffizienteRe_Energienutzung durch Systemtechnische Integration Des privaten ENdabnehmers

4.2.2 Optimale Clusterparameter

Analog zu Abschnitt 4.1.2 werden für den RESIDENS-Datensatz die optimalen Clusterparameter in einer umfangreichen Sensitivitätsanalyse bestimmt. Das Ergebnis der Ermittlung von optimalen Clusterparametern für die sechs Fälle des RESIDENS-Datensatzes ist kompakt in Tabelle 4.8 zusammengefasst.

Tabelle 4.8: Ergebnis der Bestimmung von optimalen Clusterparametern für den RESIDENS-Datensatz

Fall	Tage	opt. Clusteranzahl	opt. Fuzzifizierungsgrad	opt. Rauschabstand	Verweis
DH	alle Wochentage	$C_{opt} = 4$	$q_{opt} = 1.05$	$\delta_{opt} = 25$	Abb. B.26
DH _{norm}	alle Wochentage	$C_{opt} = 5$	$q_{opt} = 1.40$	$\delta_{opt} = 2.2$	Abb. B.27*
DH _{normmean}	alle Wochentage	$C_{opt} = 7$	$q_{opt} = 1.45$	$\delta_{opt} = 15$	Abb. B.28*
DH ^{wo}	nur Werktage	$C_{opt} = 4$	$q_{opt} = 1.05$	$\delta_{opt} = 25$	Abb. B.29
DH _{norm} ^{wo}	nur Werktage	$C_{opt} = 4$	$q_{opt} = 1.05$	$\delta_{opt} = 1.8$	Abb. B.30
DH _{normmean} ^{wo}	nur Werktage	$C_{opt} = 4$	$q_{opt} = 1.45$	$\delta_{opt} = 10$	Abb. B.31

Die beiden Fälle DH_{norm} und DH_{normmean} sind in Tabelle 4.8 in der letzten Spalte mit einem Stern gekennzeichnet. Die optimalen Clusterparameter für diese beiden Fälle stehen nicht in Abb. B.27 bzw. Abb. B.28, da nach der Sensitivitätsanalyse festgestellt wurde, dass die obere Grenze des Rauschabstandes ($\bar{\delta}$) zu niedrig gewählt wurde und so zu viele Datentupel im Rauschcluster landen. Es wurden deshalb weitere Clusterdurchläufe mit größeren Rauschabständen getestet und schlussendlich die in Tabelle 4.8 gezeigten optimalen Clusterparameter gewählt.

4.2.3 Aufgabenstellung A1: Identifikation von saisonalen und wochentagsbedingten Unterschieden für eine Demand-Response-Maßnahme

In Abschnitt 4.2.3 werden analog zu Abschnitt 4.1.3 die preisbeeinflussten Typtage der DH-, DH_{norm}⁻, DH_{normmean}⁻, DH^{wo}-, DH_{norm}^{wo}- und DH_{normmean}^{wo}-Fälle vorgestellt. Da der RESIDENS-Datensatz nur eine Tarifgruppe hat, existieren keine DT-Fälle.

4.2.3.1 DH-, DH_{norm}⁻ und DH_{normmean}⁻-Cluster als preisbeeinflusste Typtage (Alle Wochentage)

Analog zu den Untersuchungen in Abschnitt 4.1.3 wurden die optimalen Clusterparameter in einer umfassenden Sensitivitätsanalyse (vgl. Abschnitt 2.5.4.1) ermittelt. Die Ergebnisse der Analyse stehen in Tabelle 4.8. Für den Fall DH ergibt sich eine optimale Clusteranzahl von $C_{opt} = 4$. Das Clustering der varianznormierten Tageskurven führt zu $C_{opt} = 5$, das Clustering der mittelwertnormierten Tageskurven führt zu $C_{opt} = 7$, wobei eines der Cluster ein Rauschcluster ist.

Die **preisbeeinflussten Typtage** der drei Fälle DH -, DH_{norm} - und $DH_{normmean}$ sind in Abb. 4.14a(i)-(iii) dargestellt. Da alle Haushalte der gleichen Tarifgruppe angehörten und kein automatisches DR möglich war, geben die gefundenen Clusterkurven der Fälle DH -, DH_{norm} - und $DH_{normmean}$ die manuelle Verbrauchsreaktion der Haushaltsstromkunden auf ein Preissignal wider.

Wochentagsbedingte Unterschiede im preisbeeinflussten Verbrauchsverhalten lassen sich dann gut erklären, wenn die Werte der relativen Transinformation pro Ausgangsentropie für die Ausgangsgröße „Werktag“ (hier: „Typtag“) in Tabelle 4.9 entsprechend hoch ($\rightarrow 1$) sind. Da der Wert für Q_A mit $Q_A = 0.05$ sehr niedrig ist, erklären die DH -, DH_{norm} - und $DH_{normmean}$ -Cluster nur bedingt das wochentagsbedingte Verbrauchsverhalten. Die Verbundverteilung in Abb. 4.13 zeigt, dass ca. 42% aller Datentupel des Clusters C2 nicht an Werktagen aufgezeichnet wurden. Das Cluster C2 bildet somit am ehesten das preisbeeinflusste Verbrauchsverhalten an Wochenendtagen ab. Cluster C4 kann dagegen als „Werktagscluster“ bezeichnet werden.

Cluster	Werktage	Samstage	Sonntage
Noise cluster	32 (66.7%)	9 (18.8%)	7 (14.6%)
Cluster 4	29366 (81.9%)	2931 (8.2%)	3555 (9.9%)
Cluster 3	165 (74.3%)	28 (12.6%)	29 (13.1%)
Cluster 2	18953 (57.5%)	6869 (20.8%)	7168 (21.7%)
Cluster 1	41 (71.9%)	6 (10.5%)	10 (17.5%)

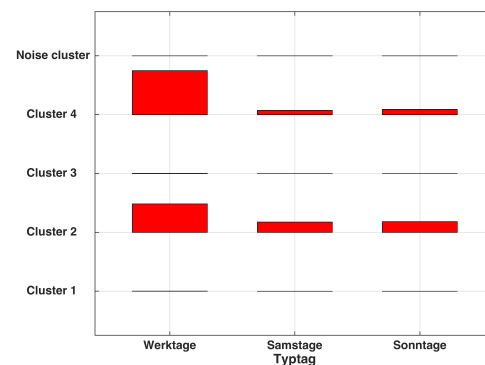


Abbildung 4.13: Verbundverteilung der Ausgangsgrößen „ DH_{norm} -Cluster“ und „Werktag“ (hier: „Typtag“) des RESIDENS-Datensatzes: (links) als Tabelle mit der Anzahl an Datentupeln und der prozentualen Verteilung der Datentupel je Zeile, (rechts) als Balkendiagramme

Die Werte der relativen Transinformation pro Ausgangsentropie für die Ausgangsgröße „Monat“ in Tabelle 4.9 belegen, dass saisonale Unterschiede für eine DR-Maßnahme am ehesten mit dem Fall DH ($Q_A = 0.04$) aufgedeckt werden können. Die Verbundverteilung in Tabelle B.16 belegt, dass das Cluster C1 aus Tageszeitreihen besteht, die vermehrt in den Frühlings- und Sommermonaten aufgezeichnet wurden (Ausnahme: November). Cluster C2 setzt sich aus Tageszeitreihen zusammen, die in den Herbst- und Wintermonaten aufgezeichnet wurden. Die Zeitreihen des Clusters C3 wurden ausschließlich in den Monaten „Dezember“ und „Januar“ aufgezeichnet. Eine nachgelagerte Analyse hat ergeben, dass alle $N_D = 33$ Datentupel des Clusters C3 zu einem Haushalt gehören. Es ist davon auszugehen, dass dieser Haushalt in den Monaten „Dezember“ und „Januar“ einen zusätzlichen elektrischen Verbraucher (z.B. elektrischer Heizlüfter) in Betrieb hatte, der dann zu dem stark erhöhten Strombezug führte. Diese Beobachtung macht deutlich, wie sorgsam bei der Auswahl von Haushalten für eine Feldstudie und der Interpretation der Ergebnisse dieser Feldstudie vorgegangen werden muss.

Tabelle 4.9: Untersuchte Fälle des RESIDENS-Datensatzes: Auswertung der Clusteranalysen mit informationstheoretischen Maßen; C_{opt} - optimale Clusteranzahl, N_D - Anzahl an Datentupeln, Q_A - relative Transinformation pro Ausgangsentropie.

Fall	Tage	C_{opt}	N_D	Q_A Monat	Q_A Werktag	Q_A Tarifgruppe
DH	alle Wochentage	4	69169	0.04	0.01	–
DH_{norm}	alle Wochentage	5	69169	0.01	0.05	–
$DH_{normmean}$	alle Wochentage	7	69169	0.01	0.03	–
DH^{wo}	Mo-Do	4	38918	0.04	–	–
DH_{norm}^{wo}	Mo-Do	4	38918	0.03	–	–
$DH_{normmean}^{wo}$	Mo-Do	4	38918	0.01	–	–

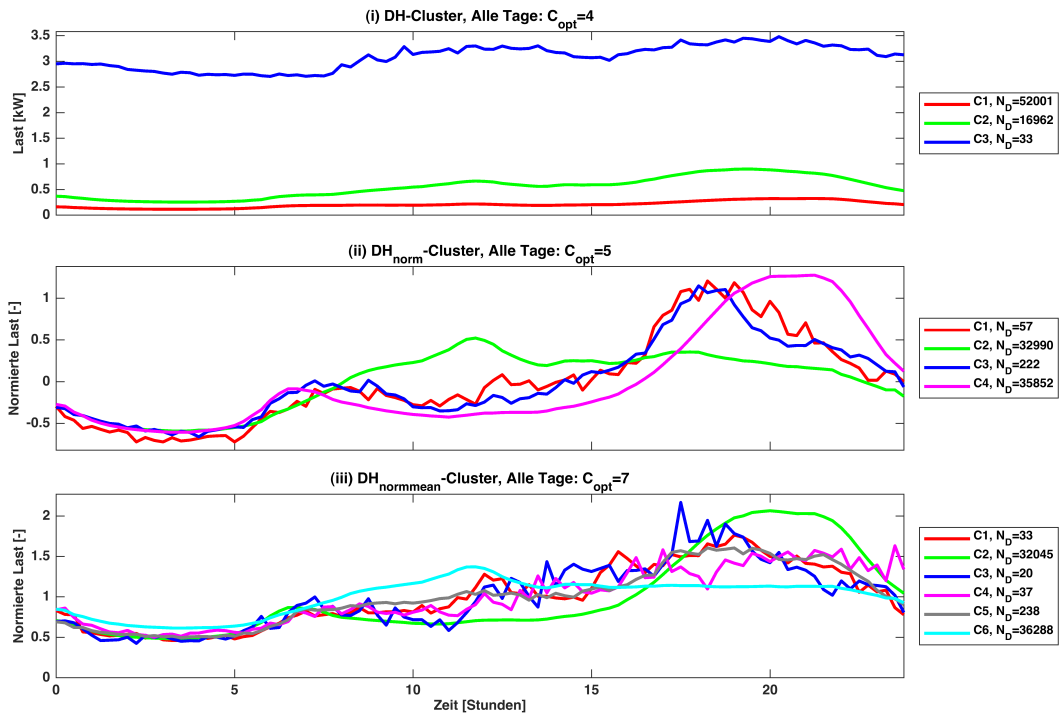
4.2.3.2 DH^{wo} -, DH_{norm}^{wo} - und $DH_{normmean}^{wo}$ -Cluster als preisbeeinflusste Typtage (Montag-Donnerstag)

Wie in Abb. 4.12 zu erkennen, unterscheidet sich im RESIDENS-Projekt die Tarifstruktur am Freitag von der Tarifstruktur an allen anderen Werktagen. Um den Einfluss der veränderten Tarifstruktur auf das Verbrauchsverhalten zu unterdrücken, wurden alle Freitage, Samstage und Sonntage ausgeschlossen. Die optimalen Clusterparameter wurden mit einer Sensitivitätsanalyse bestimmt und stehen in Tabelle 4.8. Für den Fall DH^{wo} ergibt sich eine optimale Clusteranzahl von $C_{opt} = 4$. Das Clustering des Falls DH_{norm}^{wo} führt zu $C_{opt} = 4$, das Clustering des Falls $DH_{normmean}^{wo}$ führt zu $C_{opt} = 4$, wobei in allen drei Fällen eines der Cluster ein Rauschcluster ist.

Die **preisbeeinflussten Typtage** der drei Fälle DH^{wo} , DH_{norm}^{wo} und $DH_{normmean}^{wo}$ sind in Abb. 4.14b(i)-(iii) zu sehen. Analog zu den DH-Clustern in Abb. 4.14a(i) unterscheidet sich das Cluster C3 in Abb. 4.14b(i) deutlich von den Clustern C1 und C2. Die Erklärung hierfür wurde bereits im vorigen Abschnitt gegeben. Die mittleren DH_{norm}^{wo} -Clusterkurven in Abb. 4.14b(ii) unterscheiden sich nur marginal von den entsprechenden Clusterkurven in Abb. 4.14a(ii). So lässt sich Cluster C2 in Abb. 4.14b(ii) als Zusammenschluss der Cluster C1 und C3 in Abb. 4.14a(ii) begreifen. Cluster C3 in Abb. 4.14b(iii) und der Zusammenschluss der Cluster C1, C3, C4 und C5 in Abb. 4.14a(iii) weisen einen ähnlichen Verlauf auf.

Die Werte der relativen Transinformation pro Ausgangsentropie für die Ausgangsgröße „Monat“ ist für den DH^{wo} -Fall am größten. Jedoch zeigt der sehr geringe Wert von $Q_A = 0.04$, dass von den Monaten sehr schlecht eingeschränkt auf die gefundenen Clusterkurven geschlossen werden kann. Die Identifikation von saisonalen Unterschieden für eine *Demand-Response*-Maßnahme gelingt somit nicht.

(a) Preisbeeinflusste Typtage der (i) DH-Cluster, (ii) DH_{norm} -Cluster und (iii) $DH_{normmean}$ -Cluster



(b) Preisbeeinflusste Typtage der (i) DH^{wo} -Cluster, (ii) DH_{norm}^{wo} -Cluster und (iii) $DH_{normmean}^{wo}$ -Cluster

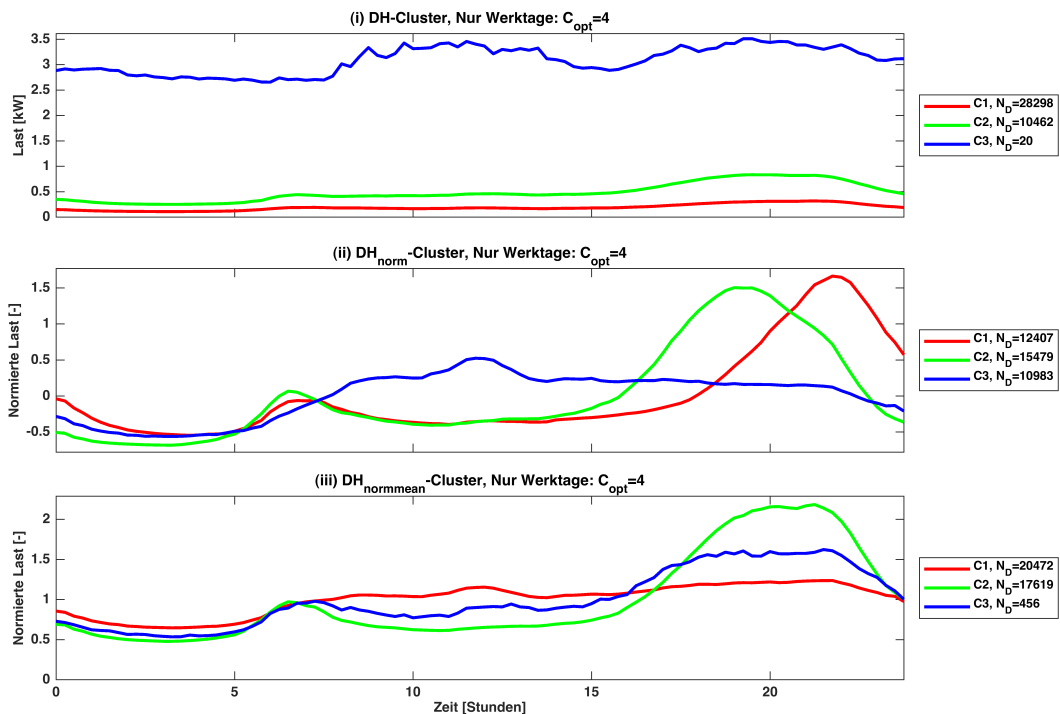


Abbildung 4.14: Gefundene DH-Cluster des RESIDENS-Datensatzes: (a) alle Tage, (b) Mo-Do. N_D : Anzahl an Datentupeln je Cluster.

4.2.4 Aufgabenstellung A3: Identifikation von (Non-) Responder-Haushalten

Analog zu der Identifikation von (Non-) Responder-Haushalten in Abschnitt 4.1.5 findet auch für diese Aufgabenstellung kein neues Clustering statt. Die visuelle Inspektion der Clusterkurven des $DH_{\text{norm}}^{\text{wo}}$ -Falls in Abb. 4.14b(ii) und des $DH_{\text{normmean}}^{\text{wo}}$ -Falls in Abb. 4.14b(iii) kann helfen, (Non-) Responder-Haushalte zu identifizieren. Um den Einfluss unterschiedlicher Tagesroutinen an Werk- und Wochenendtagen zu unterdrücken, wurden für diese Aufgabenstellung alle Freitage, Samstage und Sonntage ausgeschlossen.

In Abb. 4.14b(ii) zeigen die Cluster C1 und C2 eine Lastreduktion zu Zeiten des Hochtarifs (vgl. Abb. 4.12: Hochtarif werktags von 10-14 Uhr). Es kann allerdings nicht abschließend geklärt werden, ob die Lastsenkung eine direkte Reaktion auf das Hochpreissignal ist oder der Tatsache geschuldet ist, dass die Haushaltsbewohner zur Mittagszeit nicht zuhause waren. Des Weiteren ist in Abb. 4.14b(ii) zu erkennen, dass die Haushalte des Clusters C2 den höchsten Strombezug zu Zeiten des Mitteltarifs (ca. 19 Uhr) haben und potenzielle Kosteneinsparungen durch eine Lastverschiebung hin zu Zeiten des Niedertarifs ab 21 Uhr nicht nutzen. Einzig die Haushalte des Clusters C1 reagieren mit einer Lastverschiebung hin zu Zeiten des Niedertarifs ab 21 Uhr (Responder-Haushalte). Das Cluster C3 hat eine relativ flache mittlere Tageslastkurve, wobei der höchste Strombezug um ca. 12 Uhr zu Zeiten des Hochtarifs ist (Non-Responder-Haushalte).

Die visuelle Inspektion der mittelwertnormierten Clusterkurven in Abb. 4.14b(iii) lässt erkennen, dass auch hier ein Cluster (Cluster C1, Non-Responder-Haushalte) auszumachen ist, dessen Haushalte einen (lokalen) maximalen Strombezug zu Zeiten des Hochtarifs um ca. 12 Uhr haben. Des Weiteren findet sich ein Cluster (Cluster C2, Responder-Haushalte) unter den drei gezeigten Clustern, dessen Haushalte ein erwartbares Verhalten zeigen, indem ein Großteil des Strombezugs hin zu Zeiten des Niedertarifs verschoben wird.

4.3 Zusammenfassung und Diskussion

In Abschnitt 4.3 werden die Ergebnisse der datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens auf Basis des *Olympic-Peninsula-Project*- (Abschnitt 4.1) und RESIDENS-Datensatzes (Abschnitt 4.2) zusammengefasst. Das Clustering der nicht-aggregierten Fälle (alle DH-Fälle) des OPP- und RESIDENS-Datensatzes lässt aufgrund der großen Varianz innerhalb und zwischen den haushaltsindividuellen Lastzeitreihen nur geringe Erkenntnisse über den Zusammenhang zwischen den Klassen der Ausgangsgröße „Monat“ und den gefundenen Clusterkurven zu. Am ehesten erklären die Clusterkurven der nicht-normierten Fälle den saisonalen Einfluss auf das preisbeeinflusste Verbrauchsverhalten. Das Clustering der aggregierten Fälle (alle DT-Fälle) des OPP-Datensatzes, insbesondere der Fälle DT und DT^{wo}, liefert preisbeeinflusste Typtage, die sehr gut saisonale Unterschiede für eine DR-Maßnahme beschreiben. Wochentagsbedingte Verbrauchsunterschiede lassen sich mit den Clusterkurven der nicht-aggregierten Fälle nicht erklären. Die Aggregation und Normierung der Tageslastzeitreihen (Fälle DT_{norm} und DT_{normmean}) führt dazu, dass die Zuordnung von den Klassen der Ausgangsgröße „Werktag“ zu den gefundenen Clusterkurven etwas besser gelingt, wobei eine eindeutige Zuordnung weiter nicht gegeben ist.

Eine valide Aussage über saisonale und wochentagsbedingte Unterschiede für eine DR-Maßnahme ist nur dann zulässig, wenn sie sich auf Analyseergebnisse stützt, die aus mehreren unterschiedlichen DR-Feldstudien gewonnen wurden. Darum lassen sich die preisbeeinflussten Typtage des OPP-Datensatzes nicht losgelöst von den preisbeeinflussten Typtagen des RESIDENS-Datensatzes betrachten. Die preisbeeinflussten Typtage der DH-Cluster in Abb. 4.5 (OPP) und Abb. 4.14 (RESIDENS) zeigen signifikante Unterschiede im Tagesverlauf und Tagesstromverbrauch, welche vermutlich von den Unterschieden in der Haushaltsgröße, der Geräteausstattung und den Tarifmodellen herrührt. Des Weiteren differiert der Verlauf der OPP-Cluster (Lastspitzen am Morgen) wesentlich vom Verlauf der RESIDENS-Cluster (Lastspitzen am Abend). Insbesondere bei der Gestaltung des Preissignals (siehe Planung des Preissignals (Block A) in Abb. 2.2) müssen die Unterschiede im Tagesverlauf und Tagesstromverbrauch berücksichtigt werden.

Es muss an dieser Stelle betont werden, dass eine datensatzspezifische Analyse und, daran anknüpfend, eine datensatzspezifische Bewertung der Analyseergebnisse eher sinnvoll ist als eine datensatzübergreifende Analyse. Die Unterschiede in den Eigenschaften der getesteten *Smart-Meter*-Datensätze (z.B. Tarifmodell, Preisspreizung zwischen Hoch- und Niedertarif (vgl. Abb. 1.6), Geräteausstattung oder Vorhandensein einer Kontrollgruppe) verhindern, dass generelle Aussagen über das preisbeeinflusste Verbrauchsverhalten getroffen werden können. Eine Verallgemeinerung von Analyseergebnissen über *Smart-Meter*-Datensätze hinweg kann nur dann erfolgen, wenn vergleichbare Datensätze untersucht werden. Da im Rahmen der Doktorarbeit entsprechende Datensätze nicht verfügbar waren, müssen entweder Datensätze aus vergangenen Feldstudien verfügbar gemacht werden oder neue Datensätze in weiteren, gut geplanten Feldstudien gewonnen werden.

Um tarifbedingte Unterschiede für eine DR-Maßnahme zu erklären, muss der Zusammenhang zwischen den gefundenen Clusterkurven und den Tarifgruppen mit informationstheoretischen

Maßen bewertet werden. Das Clustering der nicht-aggregierten Fälle (alle DH-Fälle) des OPP-Datensatzes liefert preisbeeinflusste Typtage, die den Einfluss des Tarifmodells nicht erklären können. Die Clusterkurven der aggregierten Fälle (alle DT-Fälle) des OPP-Datensatzes sind diesbezüglich aussagekräftiger, was durch die geringere Varianz in den Clustern zu erklären ist. Als weiteres Ergebnis gilt festzuhalten, dass sich die normierten Fälle grundsätzlich besser eignen als die nicht-normierten Fälle, um tarifbedingte Unterschiede für eine DR-Maßnahme zu identifizieren.

Mithilfe der Analyse einzelner Abtastzeitpunkte bezüglich der Tarifeinflüsse auf die Last konnten zeitliche Lastverschiebungen für die OPP-Haushalte aufgedeckt werden (siehe Tabelle 4.3). Extrem hohe Strompreise sorgten sowohl in der TOU- als auch in der RTP-Tarifgruppe für eine starke Lastabsenkung. Allerdings zeigten die Haushalte der TOU- und RTP-Tarifgruppe z.T. untypische Reaktionen. So blieb beispielsweise eine Lastabsenkung zu Zeiten des TOU-Hochpreistarifs im Sommer aus. Eine mögliche Erklärung könnte die Verdopplung des Hochtarifzeitraums von drei (TOU-Wintertarif) auf sechs (TOU-Sommertarif) Stunden sein. Die Reaktion der RTP-Haushalte des OPP-Projektes zeigte, dass der Echtzeitstrompreis kein geeignetes Preissignal zur Beeinflussung der Haushalte darstellt. Möglicherweise führt die schnelle Änderung des Strompreises zu einer Überforderung der RTP-Haushalte.

In den Abschnitten 4.1.5 und 4.2.4 konnte gezeigt werden, dass die Identifikation von (Non-) Responder-Haushalten durch die visuelle Inspektion von Clusterkurven gelingt. Jedoch müssen vor der Inspektion der Clusterkurven diejenigen Tageslastzeitreihen im Datensatz ausgeschlossen werden, die unerwünschte Einflüsse auf das preisbeeinflusste Verbrauchsverhalten haben. So sollten nur noch Tageslastzeitreihen mit den gleichen Tarifstrukturen (z.B. nur TOU) und gleichen Wochentagen (z.B. nur Werktag) im Datensatz übrig sein.

Grundsätzlich ließen sich der Einfluss der Monate, der Wochentage und der Tarife auf das preisbeeinflusste Verbrauchsverhalten auch durch eine Mittelwertbildung und -analyse der Tageslastzeitreihen bezüglich der Ausgangsgrößen „Monat“, „Werktag (ja/nein)“ und „Tarifgruppe“ erklären. Eine reine Mittelwertbildung/-analyse setzt jedoch voraus, dass es keine Vermischung der drei Einflüsse (Monat, Wochentag, Tarifgruppe) gibt. Die untersuchten *Smart-Meter*-Daten müssten demnach Lastzeitreihen enthalten, deren Verlauf und Amplitude streng nach Monat, Wochentag und Tarifzugehörigkeit⁵ unterscheidbar sind. Bei realen *Smart-Meter*-Datensätzen ist eine derartige Mittelwertbildung/-analyse insofern nicht sinnvoll, als es beim preisbeeinflussten Verbrauchsprozess (siehe Abschnitt 2.1) zur Überlagerung der drei Einflüsse und sonstiger Effekte (z.B. unterschiedliche Tagesroutinen oder Responder-/Non-Responder-Haushalte) kommt. Eine reine Mittelwertbildung/-analyse würde die Vermischung der Einflüsse vernachlässigen. Dagegen hat das Clustering zur Ermittlung von preisbeeinflussten Typtagen und eine anschließende Clusterzuordnung zu den Ausgangsgrößen „Monat“, „Werktag (ja/nein)“ und „Tarifgruppe“ den Vorteil, dass das reale preisbeeinflusste Verbrauchsverhalten inklusive der Vermischung der Einflüsse aufgedeckt und bewertet werden kann. Außerdem können durch das Clustering unterschiedliche Reaktionen auf ein Preissignal aufgedeckt und somit Responder- von Non-Responder-Haushalten unterschieden werden.

⁵Alle Haushalte einer Tarifgruppe müssten ähnlich auf das Preissignal reagieren. Somit wären alle Haushalte Responder-Haushalte.

In Abschnitt 2.6.2 wurde ein neuer Modelltyp („Virtueller Speicher“) zur Modellierung des preisbeeinflussten Verbrauchsverhaltens eingeführt. Diese VS-Modelle bestehen aus einem System von Differenzgleichungen mit internen Parametern (z.B. Speicherfüllstand als interne Zustandsvariable des Modells), das die Änderung der Verbraucherlast als Reaktion auf ein beliebiges Preissignal beschreibt. Um das reale preisbeeinflusste Verbrauchsverhalten zu modellieren, wurde ein Modellvergleich zwischen VS-Modellen und etablierten dynamischen Regressionsmodellen unter Verwendung von experimentellen Daten aus einer DR-Feldstudie durchgeführt. Sowohl die ARX-Modelle als auch die VS-Modelle sind in der Lage, das grundlegende Verbrauchsverhalten zu erklären. Allerdings lassen die verbleibende Gesamtvariation, auch in den aggregierten DT-Fällen, und die strukturellen Unterschiede des DR-Verhaltens für verschiedene Jahreszeiten erkennen, dass die neu entwickelten DR-Modelle weiter optimiert werden müssen. So können die saisonalen Aspekte künftig mit zeitvarianten Modellparametern modelliert werden. Die hierfür erforderliche Parametereauswertung bedarf größere Datensätze, die im Rahmen der Doktorarbeit nicht verfügbar waren. Gleichwohl zeigen die Ergebnisse der Modellierung und der Modellvergleiche, dass die gut interpretierbaren VS-Modelle in gleichem Maße wie weniger interpretierbare und komplexere ARX-Modelle geeignet sind, um das preisbeeinflusste Verbrauchsverhalten von Haushaltsstromkunden zu prognostizieren.

Abschließend gilt festzuhalten, dass erst das Zusammenspiel aller Methoden der Pipeline ein umfassende und schlüssige Erklärung des preisbeeinflussten Verbrauchsverhaltens von Haushaltsstromkunden liefert. Als Beispiel sei auf das Zeitreihen-Clustering in Abschnitt 2.5.1 und dessen Anwendung in den Abschnitten 4.1.2 und 4.2.2 verwiesen. So liefert das Zeitreihen-Clustering zunächst nur preisbeeinflusste Typtage, die erst durch die nachgelagerte Bewertung der Clusterlösung und nachgelagerte Clusterzuordnung an Aussagekraft gewinnen. Darüber hinaus kann eine Aussage über prozentuale Energieeinsparungen erst dann getroffen werden, wenn der Tarifeinfluss auf die Last mit den in Abschnitt 2.5.5 vorgestellten Methoden analysiert wurde.

5 Zusammenfassung und Ausblick

Der verstärkte Ausbau von Erneuerbare-Energien-Anlagen in Deutschland führt dazu, dass der Anteil der erneuerbaren Energien am gesamten Bruttostromverbrauch im Jahr 2016 bei ca. 32% lag. Die Erzeugungscharakteristik von Photovoltaik- und Windkraftanlagen macht jedoch nachfrage- und erzeugungsseitige Anpassungen im deutschen Elektrizitätsmarkt notwendig. Im Energiesystem der Zukunft (*Smart Grid*) wird die Bereitstellung von lastseitigen Flexibilitäten eine bedeutende Rolle einnehmen. Die Veränderung der Verbraucherlast durch Preisanreize steht dabei ausdrücklich nicht in Konkurrenz zum verstärkten Ausbau von Energiespeichern, Energieübertragungs- und -verteilnetzen. Die Verbraucherbeeinflussung mithilfe von Steuer- (*Demand Side Management, DSM*) und Preissignalen (*Demand Response, DR*) setzt allerdings voraus, dass die Auswirkungen von Preissignalen für Energieverbraucher auf deren Verbrauchsverhalten untersucht, mathematisch modelliert und validiert werden. Die vorliegende Arbeit beschäftigte sich aus diesem Grund mit der datengetriebenen Analyse und Modellierung des Verbrauchsverhaltens als Reaktion auf variable Stromtarife.

Dazu wurde in Kapitel 2 ein Gesamtkonzept vorgestellt, welches den Kern der vorliegenden Doktorarbeit bildet. Das Gesamtkonzept besteht aus verschiedenen Vorverarbeitungs-, Analyse- und Modellierungsmethoden, die in einer Methodenpipeline kombiniert werden. In Abschnitt 2.1 fand die systemtechnische Betrachtung des preisbeeinflussten Verbrauchsprozesses und dessen Einordnung in das Energiesystem statt. Darauf aufbauend wurde in einer Anforderungsanalyse in Abschnitt 2.2 geklärt, welche wissenschaftlichen Aufgabenstellungen mit dem neu entwickelten Konzept untersucht werden können. Hinzu kam eine umfassende Analyse der Beschaffenheit der Messdaten, die maßgeblich den Erfolg der Aufgabenstellungen bestimmen. Die Struktur der *Smart-Meter*-Daten und die Generierung von Benchmarkdatensätzen zur Validierung der Methoden des Konzeptes wurden in Abschnitt 2.3 behandelt. Die Methodenpipeline des Konzeptes besteht aus verfahrensübergreifenden, analyseorientierten und prognoseorientierten Methoden, die separat in den Abschnitten 2.4, 2.5 und 2.6 beschrieben wurden. Kapitel 3 hatte die Implementierung der Methoden der Pipeline sowie die softwaretechnische Umsetzung des Benchmarkgenerators als wichtiges Werkzeug zur Validierung der Methoden zum Thema. Im Rahmen der Doktorarbeit wurde das Konzept zur datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens an zwei realen *Smart-Meter*-Datensätzen angewendet. Die Ergebnisse hierfür wurden in Kapitel 4 behandelt.

Die wesentlichen Beiträge der Doktorarbeit werden im Folgenden zusammengefasst:

1. Es wurde ein neues Gesamtkonzept zur systematischen datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens bei Haushaltsstromkunden entwickelt. Die Besonderheit des Konzeptes besteht darin, dass es verfahrensübergreifende, analyseorientierte und prognoseorientierte Methoden so miteinander kombiniert,

dass von einer Methodenpipeline gesprochen werden kann. Das Konzept behandelt die gesamte Prozesskette von der Erfassung der *Smart-Meter*-Rohdaten bis hin zur Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens.

2. Eine ganzheitliche Beschreibung des preisbeeinflussten Verbrauchsprozesses (siehe Abschnitt 2.1) wurde im Rahmen der Doktorarbeit vorgenommen, die es ermöglicht, das Zusammenwirken zwischen preisbeeinflussten Haushalten, Netzbetreibern und Energieversorgungsunternehmen aus systemtechnischen Gesichtspunkten zu charakterisieren.
3. Auf Basis der systemtechnischen Beschreibung des preisbeeinflussten Verbrauchsprozesses und der *Smart-Meter*-Messdaten wurde eine umfassende Anforderungsanalyse (siehe Abschnitt 2.2) durchgeführt. Diese Anforderungsanalyse macht deutlich, welche Eigenschaften ein *Smart-Meter*-Datensatz enthalten muss, um bestimmte wissenschaftliche Fragestellungen beantworten zu können. Die dabei entwickelten Checklisten (siehe Tabellen 2.1 und 2.2) können in zukünftigen Projekten, die sich mit der datengetriebenen Analyse von Verbrauchs- bzw. Lastzeitreihen beschäftigen, zum Einsatz kommen.
4. Zur Validierung der entwickelten Methoden der Pipeline wurden Benchmarkdatensätze generiert (siehe Abschnitt 2.3). Der hierfür entwickelte Benchmarkgenerator ist so konfigurierbar, dass synthetische *Smart-Meter*-Datensätze mit spezifischen Eigenschaften, wie beispielsweise die Anzahl an Haushalten, die Signalcharakteristik der unbeeinflussten Verbraucherlast oder die Verbrauchsreaktion auf ein Preissignal, erzeugt werden können.
5. Das Gesamtkonzept besteht aus drei aufeinander aufbauenden Methodenkettens (siehe Abb. 2.1), wobei die erste Methodenkette aus den verfahrensübergreifenden Methoden der Pipeline besteht. Die verfahrensübergreifenden Methoden umfassen die Datenvorverarbeitung, die Zeitreihenzerlegung, die Zeitreihennormierung und die Zeitreihenaggregation. Damit wurde erstmalig ein ganzheitliches Konzept zur Aufbereitung und Strukturierung von *Smart-Meter*-Rohdaten vorgestellt. In zwei Testverfahren in Abschnitt 2.4.1 konnte gezeigt werden, dass die gewählten Verfahren der Datenvorverarbeitung sinnvolle Ersatzwerte für fehlende Werte und Ausreißer in Lastzeitreihen liefern.
6. In der zweiten Methodenkette (siehe Abschnitt 2.5) wurde erstmalig die systematische Analyse von *Smart-Meter*-Datensätzen mit dem Ziel der Ermittlung von preisbeeinflussten Typtagen beschrieben. Der neue Ansatz erlaubt zudem, Aussagen über saisonale, wochentagsbedingte, tarifbedingte Unterschiede für eine *Demand-Response*-Maßnahme treffen zu können. Des Weiteren können mit dem neuen Konzept haushaltsindividuelle Unterschiede in der Verbrauchsreaktion (Responder-, Semi-Responder-, Non-Responder-Haushalte) identifiziert werden.
7. Das Zeitreihen-Clustering bildete den Kern der analyseorientierten Methoden der Pipeline bzw. der dritten Methodenkette. In drei Sensitivitätsanalysen (siehe Abschnitt 2.5.4) konnte der Einfluss der Clusterparameter, der Einfluss der Signalcharakteristik der Lastzeitreihen und der Einfluss der Startcluster auf das Ergebnis des Zeitreihen-Clusterings auf Basis der Benchmarkdatensätze ermittelt werden.

-
8. In Abschnitt 2.6 wurde eine neue *Demand-Response*-Modellklasse (Virtuelle-Speicher-Modelle) entwickelt, deren Modelle aus einem System von Differenzgleichungen bestehen und das Verbrauchsverhalten von Haushaltsstromkunden als Reaktion auf verschiedene Preissignale beschreiben. Die Anwendung der VS-Modelle an realen *Smart-Meter*-Daten und der Vergleich mit herkömmlichen dynamischen Regressionsmodellen zeigte, dass sich die VS-Modelle hinsichtlich „Modellgenauigkeit“ und „Modellgeneralisierungsfähigkeit“ zur Modellierung des preisbeeinflussten Verbrauchsverhaltens von Haushaltsstromkunden eignen.
 9. Das neue Konzept zur DR-Analyse und -Modellierung wurde in Kapitel 4 an zwei realen *Smart-Meter*-Datensätzen getestet: *Olympic-Peninsula-Project*- und *RESIDENS*-Datensatz. Es konnte anhand des *OPP*-Datensatzes gezeigt werden, dass der saisonale Einfluss auf das preisbeeinflusste Verbrauchsverhalten durch Clusterkurven erklärbar ist. Des Weiteren gelang mithilfe der Analyse einzelner Abtastzeitpunkte bezüglich der Tarifeinflüsse auf die Last, dass zeitliche Lastverschiebungen aufgedeckt werden konnten. Durch eine geschickte Datentupelselektion und anschließendem Clustering konnte zudem demonstriert werden, wie unterschiedliche Reaktionen auf ein Preissignal entdeckt und somit Responder- von Non-Responder-Haushalten unterschieden werden können.
 10. Die Implementierung der Methoden und Verfahren aus Kapitel 2 wurde in Kapitel 3 beschrieben. Die Softwarebasis für die implementierten Methoden bildet die *Open-Source* MATLAB-Toolbox *SciXMiner*, wobei die Methoden hauptsächlich in zwei neu entwickelten *SciXMiner*-Erweiterungen (*Load Benchmark Generator* und *Energy*) gekapselt sind. Die Realisierung der Methoden als Menüpunkte und Bedienelemente in den *SciXMiner*-Erweiterungen bietet den großen Vorteil, dass die verfahrensübergreifenden, analyse- und prognoseorientierten Methoden relativ einfach und ohne MATLAB-Kenntnisse benutzt werden können.

Die kritische Reflektion des neu entwickelten Konzeptes zur datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens lässt Stärken und Schwächen erkennen. So ist der Benchmarkgenerator aus Abschnitt 2.3.2 in der Lage, zahlreiche Eigenschaften eines *Smart-Meter*-Datensatzes zu simulieren und somit reproduzierbare synthetische Benchmarkdatensätze zu erzeugen. Jedoch zeigen die Ergebnisse der Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern (siehe Abschnitt 2.5.4.1), dass die Simulation der Verbrauchsreaktion zu ungenügenden Ergebnissen bei der Identifikation von (Non-) Responder-Haushalten führt. Es wird vermutet, dass bei der Erzeugung der Benchmarkdatensätze, speziell bei der Simulation der Verbrauchsreaktion im DR-Simulator, noch zu wenig zwischen Responder, Non-Responder und Semi-Responder differenziert wird. Aus diesem Grund sollte der *Demand-Response*-Simulator dahingehend modifiziert werden, dass eine Identifikation von (Non-) Responder-Haushalten besser gelingt.

Ein wichtiges Ziel von Abschnitt 2.6 war es, sowohl das preisbeeinflusste Verbrauchsverhalten einzelner Haushalte als auch das aggregierte preisbeeinflusste Verbrauchsverhalten zu modellieren. Die Ergebnisse aus Abschnitt 4.1.6 belegen, dass sich die neu entwickelten Virtuelle-Speicher-Modelle eher eignen, um das aggregierte preisbeeinflusste Verbrauchsverhalten zu schätzen als das preisbeeinflusste Verbrauchsverhalten einzelner Haushalte. Dies lässt sich dadurch erklären, dass der stochastische Anteil durch die Zeitreihenaggregation ab- und der

Musteranteil im Lastsignal zunimmt. Um das preisbeeinflusste Verbrauchsverhalten von Einzelhaushalten zukünftig besser prognostizieren zu können, sollte die VS-Modellklasse um weitere spezifische VS-Modelle erweitert werden.

In zukünftigen Arbeiten sollte das Konzept an weiteren *Smart-Meter*-Datensätzen getestet werden; speziell an Datensätzen, die Eigenschaften besitzen, die bisher nicht vertreten waren, wie z.B. Wetterinformationen. Mithilfe von Wetterinformationen kann dann der Einfluss der Außentemperatur auf die gefundenen Cluster bewertet werden. Des Weiteren konnte im Rahmen der Doktorarbeit nur ein Datensatz mit Echtzeitstrompreisen analysiert werden. Um zu allgemeingültigen Aussagen über das beeinflusste Verbrauchsverhalten bei Echtzeitstrompreisen zu kommen, ist es notwendig, weitere *Smart-Meter*-Datensätze mit möglichst unterschiedlichen Tarifstrukturen mit dem vorgestellten Konzept zu analysieren.

Die in Abb. 2.24 dargestellte Modellierung des preisbeeinflussten Verbrauchsverhaltens berücksichtigt bisher zwei unterschiedliche Modellklassen: dynamische Regressionsmodelle und VS-Modelle. In zukünftigen Untersuchungen sollten weitere Modellklassen (z.B. ökonomische Preiselastizitätsmodelle, siehe Tabelle 1.4) berücksichtigt und der Modellvergleich in Abschnitt 4.1.6.3 erweitert werden.

In Abschnitt 2.1 wurde die systemtechnische Beschreibung und Einordnung des preisbeeinflussten Verbrauchsprozesses vorgenommen. Die Beeinflussung von Haushaltsstromkunden durch variable Stromtarife führt dazu, dass ein zuvor unbeeinflusstes und bekanntes¹ Verbrauchsverhalten verändert wird und sich ein beeinflusstes und unbekanntes² Verbrauchsverhalten einstellt. In einem realen Energiesystem kann der beeinflusste Verbrauchsprozess nicht losgelöst von der Planung des Preissignals und den Auswirkungen des veränderten Verbrauchsverhaltens auf die Versorgungssituation betrachtet werden (siehe „Preisbeeinflusster Verbrauchsprozess mit Systemrückkopplung“ in Abschnitt 2.1). Daher ist es zwingend notwendig, das Systemverhalten in komplexen Simulationsszenarien unter Verwendung der entwickelten DR-Modelle zu untersuchen.

¹Ein nicht durch Strompreise beeinflusstes Verbrauchsverhalten kann deswegen als bekannt angesehen werden, da es bewährte Verfahren (siehe Abschnitt 1.2.7) gibt, die dieses Verhalten gut schätzen.

²Sofern es kein Konzept zur Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens gibt, kann ein durch Strompreise beeinflusstes Verbrauchsverhalten als unbekannt bezeichnet werden.

A Verzeichnis der verwendeten Abkürzungen und Formelzeichen

Allgemeine Regeln beim Erstellen des Verzeichnisses:

- Das Abkürzungsverzeichnis setzt sich aus Abkürzungen deutscher und englischer Begriffe (*Kursivschreibweise*) zusammen.
- Es muss zwischen zeitkontinuierlichen ($x(t)$) und zeitdiskreten ($x[k]$) Signalen unterschieden werden.
- Der erstellte Bezeichnerapparat orientiert sich an [85].
- Das Symbol \hat{x} zeigt immer an, dass es sich um eine Schätzung für x handelt.
- Ein \tilde{x} kennzeichnet den Median/Zentralwert für das Symbol x .
- Ein \bar{x} bedeutet in der Statistik einen (zeitlichen) Mittelwert für das Symbol x , in der Logik eine Negation.
- Ein \dot{x} kennzeichnet eine Ableitung des Merkmals x nach der Zeit t .
- Sofern nicht anderweitig vermerkt, kennzeichnen fett gedruckte Großbuchstaben Matrizen (z.B. **A**), fett gedruckte Kleinbuchstaben Vektoren (z.B. **y**) und normal gesetzte Zeichen in Symbolen skalare Werte.
- Optimale Lösungen für ein Symbol x werden durch x_{opt} bezeichnet.
- Mengen werden durch das Aufzählen von Elementen mit beispielsweise $y \in \{0, 1\}$ durch die Angabe von offenen Intervallen $y \in (0, 1)$ (alle Werte zwischen Null und Eins gehören zur Menge) oder geschlossenen Intervallen $y \in [0, 1]$ (Werte Null, Eins und alle Werte dazwischen gehören zur Menge) angegeben.
- Je nach untersuchtem Smart Meter Datensatz wird die Währungseinheit Dollar \$ oder Euro € verwendet. In nachfolgender Tabelle werden die unterschiedlichen Währungseinheiten mit GE für Geldeinheit abgekürzt.
- Energiegrößen, wie beispielsweise die verbrauchte Menge an elektrischer Energie, werden in der Literatur mit den Einheiten J, Ws, Wh, kWh oder MWh angegeben. In der nachfolgenden Tabelle werden die unterschiedlichen Energieeinheiten zusammengefasst und mit E für Energie abgekürzt.

- Leistungsgrößen, wie beispielsweise die elektrische Last eines Haushaltes, werden oftmals mit den Einheiten J/s, W, kW oder MW angegeben. In der nachfolgenden Tabelle werden die unterschiedlichen Leistungseinheiten zusammengefasst und mit LE für Leistungseinheit abgekürzt.

Abkürzungen

A/S	Strommarkt für Systemdienstleistung (<i>Ancillary Service Market</i>)
ADR	Automatisches <i>Demand Response</i> (<i>Automated Demand Response</i>)
AR	<i>AutoRegressiv model</i>
ARIMA	<i>AutoRegressiv Integrated Moving Average model</i>
ARIMAX	<i>AutoRegressiv Integrated Moving Average model with eXogeneous input</i>
ARMA	<i>AutoRegressiv Moving Average model</i>
ARMAX	<i>AutoRegressiv Moving Average model with eXogeneous input</i>
ARX	<i>AutoRegressiv model with eXogeneous input</i>
BDEW	Bundesverband der Energie- und Wasserwirtschaft
CAES	<i>Consumer Automated Energy Management System</i>
CAP	<i>Capacity Market Program</i>
CBL	<i>Customer Baseline Load</i>
CPP	Zeitvariabler Stromtarif mit Event-Preisstufen (<i>Critical Peak Pricing</i>)
CPP-F	<i>Critical Peak Pricing-Fixed</i>
CPP-V	<i>Critical Peak Pricing-Variable</i>
DB	Nachfrage-Bieterverfahren (<i>Demand Bidding</i>)
DCT	Diskrete Kosinustransformation
DLC	Direkte Laststeuerung (<i>Direct Load Control</i>)
DR	<i>Demand Response</i>
DSM	Lastmanagement (<i>Demand Side Management</i>)
EDR	Notfallprogramm (<i>Emergency Demand Response</i>)
EE	Erneuerbare Energien
EMS	<i>Energy Management System</i>
ES	<i>Exponential Smoothing</i>
EU	Europäische Union
EVU	Energieversorgungsunternehmen
Fall DH	Fall eines <i>Smart-Meter</i> -Datensatzes mit Tageszeitreihen (<i>day, D</i>) einzelner Haushalte (<i>household, H</i>); Optional: Index „norm“ bei varianznormierten Tageszeitreihen, Index „normmean“ bei mittelwertnormierten Tageszeitreihen

Fall DT	Fall eines <i>Smart-Meter</i> -Datensatzes mit aggregierten Tageszeitreihen (<i>day</i> , D) aller Haushalte einer Tarifgruppe (T); Optional: Index „norm“ bei aggregierten, varianznormierten Tageszeitreihen, Index „normmean“ bei aggregierten, mittelwertnormierten Tageszeitreihen
Fall GH	Fall eines <i>Smart-Meter</i> -Datensatzes mit Gesamtzeitreihen (G) einzelner Haushalte (<i>household</i> , H); Optional: Index „norm“ bei varianznormierten Gesamtzeitreihen, Index „normmean“ bei mittelwertnormierten Gesamtzeitreihen
Fall GT	Fall eines <i>Smart-Meter</i> -Datensatzes mit aggregierten Gesamtzeitreihen (G) aller Haushalte einer Tarifgruppe (T); Optional: Index „norm“ bei aggregierten, varianznormierten Gesamtzeitreihen, Index „normmean“ bei aggregierten, mittelwertnormierten Gesamtzeitreihen
Fall SH	Fall eines <i>Smart-Meter</i> -Datensatzes mit Abtastzeitpunkten (<i>sample points</i> , S) der Zeitreihen einzelner Haushalte (<i>household</i> , H); Optional: Index „norm“ bei Abtastzeitpunkten der varianznormierten Zeitreihen, Index „normmean“ bei Abtastzeitpunkten der mittelwertnormierten Zeitreihen
Fall ST	Fall eines <i>Smart-Meter</i> -Datensatzes mit Abtastzeitpunkten (<i>sample points</i> , S) der aggregierten Zeitreihen aller Haushalte einer Tarifgruppe (T); Optional: Index „norm“ bei Abtastzeitpunkten der aggregierten, varianznormierten Zeitreihen, Index „normmean“ bei Abtastzeitpunkten der aggregierten, mittelwertnormierten Zeitreihen
Fall VH	Fall eines <i>Smart-Meter</i> -Datensatzes mit Zeitreihen variabler Länge (V) einzelner Haushalte (<i>household</i> , H); Optional: Index „norm“ bei varianznormierten Zeitreihen variabler Länge, Index „normmean“ bei mittelwertnormierten Zeitreihen variabler Länge
Fall VT	Fall eines <i>Smart-Meter</i> -Datensatzes mit aggregierten Zeitreihen variabler Länge (V) aller Haushalte einer Tarifgruppe (T); Optional: Index „norm“ bei aggregierten, varianznormierten Zeitreihen variabler Länge, Index „normmean“ bei aggregierten, mittelwertnormierten Zeitreihen variabler Länge
FIR	<i>Finite Impulse Response</i>
HVAC	<i>Heating, Ventilation and Air Conditioning</i>
I/C	Unterbrechbare Lasten (<i>Interruptible Load</i>), Lastbegrenzung (<i>Curtable Load</i>)
IBP	Anreizorientiertes Programm (<i>Incentive-Based Program</i>)
IKT	Informations- und Kommunikationstechnik
KNN	Künstliches Neuronales Netz (<i>Artificial Neural Network</i> , ANN)
LOCF	<i>Last-Observation-Carried-Forward</i>
LP	Lineare Programmierung
MA	<i>Moving Average</i>
NaN	Kennzeichnung eines fehlerhaften Messwertes (<i>Not a Number</i>)
NEOI	<i>NEarest-Observation-Imputation</i>

NFIR	<i>Nonlinear Finite Impulse Response</i>
NIALM	<i>Non-Intrusive Appliance Load Monitoring</i>
NOCB	<i>Next-Observation-Carried-Backward</i>
NR-Haushalt	Non-Responder-Haushalt
OPP	<i>Olympic Peninsula Project</i>
OTC	<i>Over The Counter</i>
PBP	Preisorientiertes Programm (<i>Price-Based Program</i>)
PK	Preisklasse
PLC	<i>Powerline Communication</i>
PLS	<i>Penalized Least Square</i>
PTR	Zeitvariabler Tarif mit expliziten Spitzenlastperioden (<i>Peak-Time Rebate</i>)
PV	Photovoltaik
R-Haushalt	Responder-Haushalt
RESIDENS	Forschungsprojekt EffizienteRe Energienutzung durch Systemtechnische Integration Des privaten ENdabnehmerS
RLS	Rekursives Verfahren der kleinsten Fehlerquadrate (<i>Recursive Least Squares</i>)
RTP	Dynamischer Echtzeitstromtarif (<i>Real-Time Pricing</i>)
SCI	Helmholtz-Energieprogramm „Speicher und vernetzte Infrastrukturen“ (<i>Storage and Cross-linked Infrastructures</i>)
SLP	Standardlastprofil
SMD	<i>Smart-Meter-Datensatz</i>
SPP	<i>Statewide Pricing Pilot</i>
SR-Haushalt	Semi-Responder-Haushalt
SVM	<i>Support Vector Machines</i>
SVR	<i>Support Vector Regression</i>
TOU	Zeitvariabler Tarif mit statischen Preisstufen (<i>Time-of-Use</i>)
USA	<i>United States of America</i>
V2G	<i>Vehicle-to-Grid</i>

Griechische Formelzeichen

α_{EMA}	Glättungsfaktor	[–]
α	Signifikanzniveau	[–]
β_E	Gewichtungsfaktor der exogenen Komponente im additiven Komponentenmodell	[–]
β_M	Gewichtungsfaktor der Musterkomponente im additiven Komponentenmodell	[–]

β_P	Gewichtungsfaktor der periodischen Komponente im additiven Komponentenmodell	[–]
β_S	Gewichtungsfaktor der stochastischen Komponente im additiven Komponentenmodell	[–]
β_T	Gewichtungsfaktor der Trendkomponente im additiven Komponentenmodell	[–]
β_{VS}	Parameter der exponentiellen Speicherausgleichsfunktion $a_{VS,S}^{\text{exp}}$	[–]
$\Delta\delta$	Stufenbreite des Rauschabstandes	[–]
ΔC	Stufenbreite der Clusteranzahl	[–]
$\Delta P[k]$	Laständerung zum Zeitpunkt k	[LE]
$\Delta p[k]$	Strompreisänderung zum Zeitpunkt k	[GE/E]
Δq	Stufenbreite des Fuzzifizierungsgrades	[–]
δ	Rauschabstand	[–]
γ_{VS}	Parameter der exponentiellen Speicherausgleichsfunktion $a_{VS,S}^{\text{exp}}$	[–]
μ_a	Erwartungswert der prozentualen Anteile an Ausreißern	[–]
μ_E	Erwartungswert der Gesamtstromverbräuche	[E]
$\mu_{n_D,c}$	Zugehörigkeitsfunktion des n_D -ten Datentupels zum c -ten Cluster	[–]
$\mu_{nan,g}$	Erwartungswert der prozentualen Anteile an zusammenhängenden fehlenden Werten	[–]
$\mu_{nan,s}$	Erwartungswert der prozentualen Anteile an einzelnen fehlenden Werten	[–]
μ_c	Clusterzugehörigkeitsfunktion des c -ten Clusters	[–]
σ	Standardabweichung	[–]
σ^2	Varianz	[–]
σ_a^2	mittlere quadratische Abweichung der prozentualen Anteile an Ausreißern um den Erwartungswert μ_a	[–]
σ_E^2	mittlere quadratische Abweichung der Gesamtstromverbräuche E_n um den Erwartungswert μ_E	[E ²]
$\sigma_{nan,g}^2$	mittlere quadratische Abweichung der prozentualen Anteile an zusammenhängenden fehlenden Werten um den Erwartungswert $\mu_{nan,g}$	[–]
$\sigma_{nan,s}^2$	mittlere quadratische Abweichung der prozentualen Anteile an einzelnen fehlenden Werten um den Erwartungswert $\mu_{nan,s}$	[–]
σ_P	Standardabweichung einer Lastzeitreihe	[LE]
τ_{oc}	prozentualer Anteil der Datentupel im Rauschcluster bezogen auf die Gesamtanzahl der Datentupel	[–]
$\underline{\delta}, \bar{\delta}$	untere und obere Schranke des Rauschabstandes	[–]
ε	Preiselastizität	[–]

$\varepsilon[j, k]$	Kreuzpreiselastizität (Abtastzeitpunkte j und k)	[–]
$\varepsilon[k, k]$	Eigenpreiselastizität (Zeitpunkt k)	[–]

Lateinische Formelzeichen

$\bar{P}^{T_i}[k]$	mittlere Kurve der Verbraucherlast aller N_{T_i} Haushalte der Tarifgruppe T_i zum Zeitpunkt k	[LE]
$\bar{P}_R^{T_i}[k]$	mittlere Kurve des preisbeeinflussten Anteils der Verbraucherlast aller N_{T_i} Haushalte der Tarifgruppe T_i zum Zeitpunkt k	[LE]
$\bar{P}_U^{T_i}[k]$	mittlere Kurve des unbeeinflussten Anteils der Verbraucherlast aller N_{T_i} Haushalte der Tarifgruppe T_i zum Zeitpunkt k	[LE]
$\bar{P}_d^{T_i}[k]$	mittlere Tageslastkurve aller N_{T_i} Haushalte der Tarifgruppe T_i am d -ten Tag zum Zeitpunkt k	[LE]
$\bar{P}_{d,n}$	zeitliche Mittelung der Last für den n -ten Haushalt am d -ten Tag	[LE]
$\bar{x}_b[k]$	k -ter Abtastzeitpunkt der Mittelwertzeitreihe aus allen $N_{D,b}$ Datentupeln in der b -ten Klasse der Ausgangsgröße y_i	[–]
$\bar{P}[k]$	mittlere Kurve der Verbraucherlast für alle Haushalte zum Zeitpunkt k	[LE]
$\bar{P}^{T_i}[k]$	mittlere Kurve der Verbraucherlast für alle N_{T_i} Haushalte einer Tarifgruppe T_i zum Zeitpunkt k	[LE]
$\bar{P}_R^{T_i}[k]$	mittlere Kurve der preisbeeinflussten Anteile der Verbraucherlast für alle N_{T_i} Haushalte einer Tarifgruppe T_i zum Zeitpunkt k	[LE]
$\bar{P}_R^{T_i}[k]$	mittlere Kurve der preisbeeinflussten Anteile der Verbraucherlast für alle N_{T_i} Haushalte einer Tarifgruppe T_i zum Zeitpunkt k	[LE]
$\bar{P}_U^{T_i}[k]$	mittlere Kurve der unbeeinflussten Anteile der Verbraucherlast für alle N_{T_i} Haushalte einer Tarifgruppe T_i zum Zeitpunkt k	[LE]
$\bar{P}_R[k]$	mittlere Kurve der preisbeeinflussten Anteile der Verbraucherlast für alle Haushalte zum Zeitpunkt k	[LE]
$\bar{P}_R[k]$	mittlere Kurve der preisbeeinflussten Anteile der Verbraucherlast für alle Haushalte zum Zeitpunkt k	[LE]
$\bar{P}_U[k]$	mittlere Kurve der unbeeinflussten Anteile der Verbraucherlast für alle Haushalte zum Zeitpunkt k	[LE]
\mathbf{v}_i	Clusterzentrum des i -ten Clusters	[–]
$\mathbf{V}_{K \times C}$	Matrix der Clusterzentren für alle K Abtastzeitpunkte und C Clusterzentren	[–]
$\mathbf{X}_{K \times N_D}$	Matrix der Werte der Verbraucherlast für alle K Abtastzeitpunkte und N_D Datentupel	[–]
\mathcal{N}	normalverteilte Zufallsvariable	[–]
\mathcal{U}	gleichverteilte Zufallsvariable	[–]

$\underline{a}_{VS}, \overline{a}_{VS}$	untere und obere Schranke der Speicherbe-/entladegeschwindigkeit	[1/h]
$\underline{C}, \overline{C}$	untere und obere Schranke der Clusteranzahl	[-]
$\underline{E}_{VS,n}^{max}, \overline{E}_{VS,n}^{max}$	untere und obere Schranke des maximalen Füllstands des virtuellen Speichers (n -ter Haushalt)	[E]
$\underline{E}_{VS,n}^{min}, \overline{E}_{VS,n}^{min}$	untere und obere Schranke des minimalen Füllstands des virtuellen Speichers (n -ter Haushalt)	[E]
$\underline{E}_{VS,n}^s, \overline{E}_{VS,n}^s$	untere und obere Schranke des Sollwerts des Füllstandes des virtuellen Speichers (n -ter Haushalt)	[E]
$\underline{q}, \overline{q}$	untere und obere Schranke des Fuzzifizierungsgrades	[-]
a, b	untere und obere Intervallgrenze der gleichverteilten Zufallsvariablen $X_{ha}^n \sim \mathcal{U}(a, b)$ zur Berechnung der Werte der Ausreißer	[-]
$a_{VS,S}^{exp}$	exponentielle Speicherausgleichsfunktion	[-]
$a_{VS,S}^{lin}$	lineare Speicherausgleichsfunktion	[-]
a_i, b_j	Regressionsparameter (z.B. des ARX-Modells)	[-]
$a_{VS,n}$	Speicherbe- und -entladegeschwindigkeit (n -ter Haushalt)	[1/h]
$a_{VS,S}[K]$	Wert der Speicherausgleichsfunktion zum Zeitpunkt k	[-]
b	Indexvariable der Klassen einer Ausgangsgröße	[-]
b_{nan}^n	Anzahl an Bereichen mit fehlenden Werten	[-]
C	Clusteranzahl	[-]
c	Index des Clusters	[-]
$C_{h,\alpha}^{Ham}$	kritischer Wert nach Hampel	[-]
C_{opt}	optimale Clusteranzahl	[-]
D	Anzahl der Tage	[-]
d	Index des Tages	[-]
$d_{Euk,oc}$	euklidischer Abstand zwischen Rauschcluster \mathbf{v}_{oc} und allen Datentupeln \mathbf{x}_{nD}	[-]
d_{Euk}	euklidischer Abstand	[-]
$e[k]$	gaußverteilttes weißes Rauschen zum Zeitpunkt k	[-]
$E_{VS,n}^{max}$	maximaler Füllstand des virtuellen Speichers (n -ter Haushalt)	[E]
$E_{VS,n}^{min}$	minimaler Füllstand des virtuellen Speichers (n -ter Haushalt)	[E]
$E_{VS,n}^s$	Sollwert des Füllstandes des virtuellen Speichers (n -ter Haushalt)	[E]
E_n	Gesamtenergieinhalt der Lastzeitreihe (n -ter Haushalt)	[E]
$E_{VS,n}[k]$	Füllstand des virtuellen Speichers zum Zeitpunkt k (n -ter Haushalt)	[E]
f_s	Abtastfrequenz	[1/h]

H	Entropie	[–]
H	Vorhersagehorizont	[–]
h	Anzahl der Abtastzeitpunkte der Hampel-Filterbreite	[–]
$H(y)$	Entropie einer Ausgangsgröße y	[–]
$H(y_1)$	Eingangsentropie der Ausgangsgröße y_1	[–]
$H(y_1, y_2)$	Verbundentropie (Ausgangsgrößen y_1 und y_2)	[–]
$H(y_2)$	Ausgangsentropie der Ausgangsgröße y_2	[–]
$I(y_1; y_2)$	Transinformation (Ausgangsgrößen y_1 und y_2)	[–]
i, j, l	Indexvariablen, die je nach Kontext eine andere Bedeutung haben	[–]
J_q	Zielfunktion des Fuzzy-C-Means-Clusterverfahrens	[–]
K	Anzahl der Abtastzeitpunkte im Betrachtungsintervall	[–]
k	Abtastzeitpunkt (wertediskret)	[–]
K_a	Anzahl an Ausreißern	[–]
k_a	Indexposition eines Ausreißers	[–]
$K_{nan,g}$	Anzahl an zusammenhängenden fehlenden Werten	[–]
$K_{nan,s}$	Anzahl an einzelnen fehlenden Werten	[–]
K_s	Anzahl der Abtastzeitpunkte in einem kürzeren Zeitreihenabschnitt ($K_s < K$)	[–]
m_y	Anzahl der Klassen der Ausgangsgröße y_i	[–]
MAD	Zentralwert der absoluten Abweichungen (engl. <i>Median Absolute Deviation</i>)	[–]
N	Anzahl der Haushalte	[–]
n	Index des Haushalts	[–]
n_a, n_b	Modellordnungen des ARX-Modells	[–]
N_{BD}	Anzahl an Datensätzen mit unterschiedlicher Lastsignalcharakteristik (engl. <i>benchmark dataset</i> , Index: „BD“)	[–]
N_{CP}	Anzahl an Clusterdurchläufen (engl. <i>cluster performances</i> , Index: „CP“)	[–]
$N_{D,b}$	Anzahl der Datentupel, die zu einer bestimmten Klasse b gehören	[–]
$N_{D,oc}$	Anzahl an Datentupeln im Rauschcluster	[–]
N_D	Anzahl der Datentupel	[–]
n_D	Index des Datentupels	[–]
n_k	Totzeit des ARX-Modells	[–]
$N_{P,normmean}$	Anzahl an unterschiedlichen Parameterkombinationen für mittelwertnormierte Tageszeitreihen	[–]

$N_{P,\text{norm}}$	Anzahl an unterschiedlichen Parameterkombinationen für varianznormierte Tageszeitreihen	[–]
N_P	Anzahl an unterschiedlichen Parameterkombinationen für nichtnormierte Tageszeitreihen	[–]
N_{SC}	Anzahl an Strategien zur Wahl der Startcluster	[–]
N_{T_i}	Anzahl der Haushalte einer Tarifgruppe T_i	[–]
$P[k]$	Last zum Zeitpunkt k	[LE]
$p[k]$	Strompreis zum Zeitpunkt k	[GE/E]
$p^{T_i}[k]$	Strompreis, der für alle N_{T_i} Haushalte einer Tarifgruppe T_i zum Zeitpunkt k gilt	[GE/E]
$P_{R,n}^{T_i}[k]$	preisbeeinflusster Anteil der Verbraucherlast des n -ten Haushalts einer Tarifgruppe T_i zum Zeitpunkt k	[LE]
$P_{d,n,\text{normmean}}[k]$	mittelwertnormierte Last des n -ten Haushalts zum Zeitpunkt k am d -ten Tag	[–]
$P_{d,n,\text{norm}}[k]$	varianznormierte Last des n -ten Haushalts zum Zeitpunkt k am d -ten Tag	[–]
$P_{d,n}[k]$	Last des n -ten Haushalts zum Zeitpunkt k am d -ten Tag	[LE]
$p_{EMA}[k]$	exponentiell geglätteter Strompreis zum Zeitpunkt k (engl. <i>exponential moving average</i> , Index: „EMA“)	[GE/E]
p_f	zukünftiger Strompreis (engl. <i>future</i> , Index: „f“)	[GE/E]
$p_n[k]$	Strompreis des n -ten Haushalts zum Zeitpunkt k	[GE/E]
$p_n[k]$	Strompreis des n -ten Haushalts zum Zeitpunkt k	[GE/E]
p_p	vergangener Strompreis (engl. <i>past</i> , Index: „p“)	[GE/E]
$P_{R,n}[k]$	preisbeeinflusste Anteil der Verbraucherlast des n -ten Haushalts zum Zeitpunkt k	[LE]
$P_R[k]$	preisbeeinflusster Anteil der Verbraucherlast zum Zeitpunkt k	[LE]
$P_{U,E}[k]$	exogene Komponente der unbeeinflussten Last zum Zeitpunkt k	[LE]
$P_{U,M}[k]$	Musterkomponente der unbeeinflussten Last zum Zeitpunkt k	[LE]
$P_{U,n}[k]$	unbeeinflusster Anteil der Verbraucherlast des n -ten Haushalts zum Zeitpunkt k	[LE]
$P_{U,P}[k]$	periodische Komponente der unbeeinflussten Last zum Zeitpunkt k	[LE]
$P_{U,S}[k]$	stochastische Komponente der unbeeinflussten Last zum Zeitpunkt k	[LE]
$P_{U,T}[k]$	Trendkomponente der unbeeinflussten Last zum Zeitpunkt k	[LE]
$P_U[k]$	unbeeinflusster Anteil der Verbraucherlast zum Zeitpunkt k	[LE]
$P_U[k]$	unbeeinflusster Anteil der Verbraucherlast zum Zeitpunkt k	[LE]
q	Fuzzifizierungsgrad	[–]
Q_A	relative Transinformation pro Ausgangsentropie $H(y_2)$	[–]

Q_E	relative Transinformation pro Eingangsentropie $H(y_1)$	[–]
$Q_{C,MAE}$	Gütekriterium: mittlerer absoluter Fehler zwischen Zielfunktion $y[k]$ und ihrem vom Modell prognostizierten Wert $\hat{y}[k]$	[–]
$Q_{C,ME}$	Gütekriterium: mittlerer Fehler zwischen Zielfunktion $y[k]$ und ihrem vom Modell prognostizierten Wert $\hat{y}[k]$	[–]
$Q_{C,MSE}$	Gütekriterium: mittlerer quadratischer Fehler zwischen Zielfunktion $y[k]$ und ihrem vom Modell prognostizierten Wert $\hat{y}[k]$	[–]
$Q_{C,NRMSE}$	Gütekriterium: <i>Normalized Root Mean Square Error</i> zwischen Zielfunktion $y[k]$ und ihrem vom Modell prognostizierten Wert $\hat{y}[k]$	[–]
$Q_{C,RMSE}$	Gütekriterium: Wurzel der mittleren quadratischen Abweichung zwischen Zielfunktion $y[k]$ und ihrem vom Modell prognostizierten Wert $\hat{y}[k]$	[–]
$Q_{C,SSE}$	Gütekriterium: Summe der Fehlerquadrate zwischen Zielfunktion $y[k]$ und ihrem vom Modell prognostizierten Wert $\hat{y}[k]$	[–]
r	Gütekriterium: Korrelationskoeffizient	[–]
R^2	Gütekriterium: Bestimmtheitsmaß	[–]
S	Trennungsgrad (<i>compactness and separation validity index</i> oder <i>Xie-Beni-Index</i>)	[–]
s	Anzahl der Einzelmerkmale	[–]
s_y	Anzahl der Ausgangsgrößen	[–]
s_z	Anzahl der Zeitreihen	[–]
t	Zeit (wertekontinuierlich)	[h]
T_i	Tarif	[–]
T_S	Abtastperiodendauer (engl. <i>sampling period</i> , Index: „S“)	[h]
$u[k]$	Einflussfaktoren auf den unbeeinflussten Anteil der Haushaltslast (k -ter Abtastzeitpunkt)	[–]
$u_{ESS}[k]$	exogene Einflussgrößen des elektrischen Energiesystems (k -ter Abtastzeitpunkt)	[–]
$u_m[k]$	meteorologische Einflussfaktoren auf den unbeeinflussten Anteil der Haushaltslast (k -ter Abtastzeitpunkt)	[–]
$u_n[k]$	Einflüsse auf den unbeeinflussten Anteil der Haushaltslast, die auf die technische Ausstattung des Haushalts zurückzuführen sind (k -ter Abtastzeitpunkt)	[–]
$u_{sd}[k]$	soziodemografische Einflüsse auf den unbeeinflussten Anteil der Haushaltslast, wie der sozialen Status, Einkommen, Anzahl und Alter der Haushaltsbewohner (k -ter Abtastzeitpunkt)	[–]
$u_s[k]$	singuläre Faktoren wie Sonderereignisse, die den unbeeinflussten Anteil der Haushaltslast in besonderem Maße beeinflussen (k -ter Abtastzeitpunkt)	[–]

$u_t[k]$	tageszeitabhängige und saisonale Einflussfaktoren auf den unbeflussten Anteil der Haushaltslast (k -ter Abtastzeitpunkt)	[–]
X_ε^n	Zufallsvariable, die die Stärke der Verbrauchsreaktion auf ein Preissignal angibt (n -ter Haushalt)	[–]
X_a^n	normalverteilte Zufallsvariable der Anzahl an Ausreißern (n -ter Haushalt)	[–]
X_E^n	normalverteilte Zufallsvariable des Gesamtenergieinhalts der Lastzeitreihe (n -ter Haushalt)	[–]
X_{ha}^n	gleichverteilte Zufallsvariable, die den numerischen Wert eines Ausreißers festlegt (n -ter Haushalt)	[–]
$X_{nan,g,pos}^n$	gleichverteilte ganzzahlige Zufallsvariable, die die Startindizes für die Bereiche mit fehlenden Werten angibt (n -ter Haushalt)	[–]
$X_{nan,g}^n$	normalverteilte Zufallsvariable der Anzahl an zusammenhängenden fehlenden Werten (n -ter Haushalt)	[–]
$X_{nan,s}^n$	normalverteilte Zufallsvariable der Anzahl an einzelnen fehlenden Werten (n -ter Haushalt)	[–]
$x_{j,n_D}[k]$	k -ter Abtastzeitpunkt des n_D -ten Datentupels der j -ten Zeitreihe im Datensatz	[–]
x_{l,n_D}	n_D -tes Datentupel für das l -te Einzelmerkmal im Datensatz	[–]
y_{i,n_D}	n_D -tes Datentupel der i -ten Ausgangsgröße im Datensatz	[–]
$z[k]$	Störgröße zum Zeitpunkt k	[LE]
z^{-1}	Verschiebeoperator	[–]

B Anhang

Der hier gezeigte Anhang hat den gleichen Aufbau wie die Kapitel 1 bis 4. Die Anhänge B.1 bis B.4 erweitern die Kapitel 1 bis 4 um zusätzliche wertvolle Tabellen, Bilder und Informationen.

B.1 Einleitung

B.1.1 Smart-Grid-Projekte

In Tabelle B.1 ist eine Auflistung von vergangenen und aktuellen *Smart-Grid*-Projekten zu sehen.

Tabelle B.1: Übersicht über *Smart-Grid*-Projekte in der Europäischen Union; Abkürzungen: InSI - Intelligente Systemintegration, NA-ÜN - Netzautomation Übertragungsnetz, NA-VN - Netzautomation Verteilnetz, SG - Smart Grid, SMM - Smart Metering, SpT - Speichertechnologien; Länderkürzel: DE - Deutschland, UK - Großbritannien, IE - Irland, FR - Frankreich, ES - Spanien, PT - Portugal, BE - Belgien, NL - Niederlande, CH - Schweiz, IT - Italien, MT - Malta, CY - Zypern, SL - Slowenien, HU - Ungarn, SK - Slowakei, AT - Österreich, CZ - Tschechien, PL - Polen, DK - Dänemark, FI - Finnland, SE - Schweden, NO - Norwegen.

Projektname	Organisation	Land	Zeitraum	Kategorie
„Mülheim zählt“ Smart Meter Programm	RWE DAG	DE	2009 - 2011	SMM
0.4 kV remote control	-	PL	-	SMM
220 kV SSSC device for power flow control	Red Electrica de Espana S.A.	ES	7/2009 - 7/2014	NA-ÜN
A complete and normalized 61850 substation	Red Electrica de Espana S.A.	ES	10/2009 - 12/2015	NA-ÜN
Activation of 200 MW refuse-generated CHP upward regulation effect	EMD International A/S	DK	7/2009 - 12/2010	InSI
Active Network Management	SG Solutions	UK	4/2010 - 4/2011	NA-VN
Address	Enel Distribuzione	EU-IT	6/2008 - 6/2012	InSI
ADELE Project AA-CAES	RWE Power AG	DE	12/2009 - 12/2013	SpT

Fortsetzung der Tabelle

Projektname	Organisation	Land	Zeitraum	Kategorie
Advanced Systems of Efficient Use of Electrical Energy - SURE	Teces	SL	1/2011 - 1/2014	InSI
AFTER - A Framework for electrical power systems vulnerability identification, defense and Restoration	RSE - Ricerca sul Sistema Energetico	EU-IT	2011 - 2014	NA-ÜN
Agent based control of power systems, ForskEL	CET-DTU	DK	4/2006 - 3/2010	SMM
Almacena	Red Electrica de Espana S.A.	ES	9/2009 - 12/2013	SpT
AMI	Elektro Gorenjska d.d.	SL	-	SMM
AMI	ENERGA-OPERATOR SA	PL	2010 - 2017	SMM
AMIS	Energie AG OÖ Netz GmbH	AT	1/2005 - 6/2012	InSI
Application of SG in photovoltaic power systems, ForskEL	Danfoss Solar Inverters	DK	10/2010 - 12/2013	SMM
Automation and security of Supply	-	DK	-	NA-VN
Automation systems for Demand Response, ForskEL	Danish Energy Industries Federation	DK	3/2006 - 7/2009	SMM
BeAware	Teknillinen Korkeakoulu	EU-FI	5/2010 - 5/2013	DSM
Belgium east loop active network management	-	BE	9/2010 - 6/2011	-
BeyWatch	Telefonica Investigacion y Desarrollo SA	EU-ES	12/2008 - 5/2011	DSM
Building to Grid (B2G)	Salzburg AG	AT	7/2010 - 12/2013	DSM
Central Networks Low Carbon Hub - Optimizing renewable energy resources in Lincolnshire	Central Networks	UK	1/2011 - 12/2014	NA-VN
CET2001 Customer Led Network Revolution	CE Electric	UK	1/2011 - 12/2013	SMM
Charge stands	SydEnergi, (DSO)	DK	1/2010 -	SMM
Charging Infrastructure	Goteborg Energy AB	SE	1/2008 - 12/2010	InSI
Clyde Gateway	Scottish Power	UK	4/2010 - 4/2011	NA-VN
Concept for Management of the Future Electricity System	Energinet.dk, (TSO)	DK	1/2009 - 12/2011	Sonstiges

Fortsetzung der Tabelle

Projektname	Organisation	Land	Zeitraum	Kategorie
Consumer to Grid (C2G)	Salzburg AG	AT	7/2010 - 6/2012	DSM
Consumer web	Vestforsyning A/S (DSO)	DK	1/2010 - 1/2011	DSM
Control and regulation of modern distribution system, ForskEL	Department of Energy Technology- Aalborg University	DK	10/2006 - 12/2010	NA-VN
Cost benefit analysis for the implementation of SMM with pilot project	Ministry of economy of the Slovak republic	SK	1/2011 - 12/2011	SMM
Cryogenic Storage	High view Power Storage	UK	4/2010 - 4/2011	SpT
Customer Value Proposition SG(KEL)	Goteborg Energy AB	SE	1/2008 - 12/2012	Sonstiges
Cyprus SMM demo	Electricity Authority of Cyprus (EAC)	CY	7/2010 - 7/2012	SMM
DA (Distribution Automation)	Enexis BV	NL	2007 - 2010	NA-VN
Data Exchange	National Grid	UK	4/2010 - 4/2011	NA-ÜN
DataHub project	Energinet.dk, (TSO)	DK	4/2009 - 4/2012	SMM
DCN4TSO	Elektro-Slovenija (TSO)	SL	2004 -	NA-ÜN
Decentralized customer-level under frequency load shedding in Switzerland	University of Applied Sciences, Northwestern Switzerland	CH	9/2010 - 2/2012	InSI
Demand response medium sized industry consumers	Danish Technological Institute	DK	1/2009 - 12/2011	InSI
Demonstration project Smart Charging	Enexis	NL	9/2010 - 12/2011	NA-VN
DER-IREC 22@Microgrid	GTD Sistemas de Informacion SA	ES	6/2009 - 11/2011	InSI
Development of a Secure, Economic and Environ-mentally friendly Modern Power System	Department of Energy Technology- Aalborg University	DK	1/2010 - 12/2014	SMM
Development of Early Warning Systems (PMU/WAMS)	Energinet.dk, (TSO)	DK	2006 - 2012	NA-ÜN
DG Demonetz Smart LV Grid	Austrian Institute of Technology	AT	3/2011 - 3/14	InSI
DG Demonetz Validierung	AIT Austria	AT	1/2006 - 2/2013	NA-VN

Fortsetzung der Tabelle

Projektname	Organisation	Land	Zeitraum	Kategorie
Distributed connected wind farms	ESB Networks	IE	5/2009 - 4/2012	Sonstiges
Distribution System planning for SGs, ForskEL	Riso DTU	DK	3/2011 - 7/2012	Sonstiges
DLC+VIT4IP	Kema Nederland BV	EU-NL	1/2010 - 1/2013	Sonstiges
DSO-pilot project	SydEnergi, (DSO)	DK	2009 - 2010	SMM
DSO-Pilot project - Automatic receipt of short circuiting indicators	-	DK	6/2009 - 12/2010	SMM
Easy Street	Enexis	NL	1/2011 - 6/2014	InSI
EcoGrid Denmark, ForskEL	Danish Technological Institute	DK	5/2007 - 4/2009	InSI
EcoGrid EU	Energinet.dk	DK	1/2011 - 12/2014	InSI
E-DeMa	RWE RWN	DE	2009 - 2014	InSI
EDISON	Danish Energy Association	DK	2/2009 - 12/2011	NA-VN
eFlex	DONG Energy, (DSO)	DK	2010 - 2011	NA-VN
Electrical vehicles impacts on the grids	Ores	BE	9/2010 - 12/2011	NA-VN
Electricity demand as frequency controlled re-serves, ENS	CET-DTU	DK	4/2009 - 3/2012	InSI
Electricity demand as frequency controlled re-serves, ForskEL	CET-DTU	DK	4/2006 - 6/2008	InSI
Electricity for road transport, flexible power systems and wind power	Riso DTU	DK	1/2008 - 1/2011	NA-VN
Electricity storage for short term power system service	Materials Research Division at Riso DTU	DK	3/2010 - 12/2010	SpT
Elforsk SG programme	Elforsk AB	SE	1/2011 - 12/2014	SMM
E-mobility	-	IT	12/2008 - 12/2013	InSI
E-mobility	ELMU-EMASZ Nyrt.	HU	-	InSI
EMPORA 1 + 2 - E-Mobile Power Austria	VERBUND AG	AT	1/2010 - 6/2014	NA-VN

Fortsetzung der Tabelle

Projektname	Organisation	Land	Zeitraum	Kategorie
ENERGOZ	Institute of Technology, Slovak Academy of Sciences, Bratislava	SK	5/2010 - 5/2013	InSI
Energy @ home	-	IT	1/2009 - 12/2011	-
Energy Forecast, ForskEL	Ramboll	DK	2/2007 - 3/2010	SMM
E-price	Eindhoven University of Technology	EU-NL	2/2010 - 2/2013	NA-ÜN, NA-VN
ESB Smart Meter Projects	Electricity Supply Board - Networks	IE	1/2009 - 4/2011	SMM
ESTER, Enel integrated System for TESTs on stoRage	ENEL	IT	4/2009 - 4/2013	SpT
E-telligence	EWE AG	DE	11/2009 - 10/2012	InSI
ETM (Distribution Network Automation on 10 kV cable line stations)	ELMU Halozati Kft.	HU	2009 - 2013	NA-VN
EU-DEEP	GDF Suez	EU-FR	1/2004 - 6/2009	InSI
EV Network integration	ESB Networks	IE	12/2009 - 12/2012	DSM
EVCOM	Energinet.dk, (TSO)	DK	1/2008 - 12/2010	SMM
EWIS - European wind integration study	ELIA SYSTEM OPERATOR SA	EU-BE	6/2007 - 10/2009	Sonstiges
ewz-Studie SMM	ewz - Elektrizitätswerk der Stadt Zürich	CH	7/2010 - 12/2012	DSM
Fenix	Iberdrola Distribucion	EU-ES	10/2005 - 10/2009	InSI
Fieldtrail Mobile SG	Enexis	NL	4/2010 - 4/2011	NA-VN
Flex power - perspectives of indirect power system control through dynamic power price	EA energianalyse	DK	6/2010 - 6/2013	InSI
Flexcom, ForskEL	Riso DTU	DK	6/2008 - 5/2010	NA-VN
Fraunhofer FIT	Fraunhofer FIT	DE	2009 - 2012	DSM
From wind power to heat pumps	Energinet.dk, (TSO)	DK	11/2009 - 12/2011	InSI

Fortsetzung der Tabelle

Projektname	Organisation	Land	Zeitraum	Kategorie
G4V - Grid for Vehicles	RWE RHEINLAND WESTFALEN NETZ AG	EU-DE	1/2010 - 6/2011	Sonstiges
GAD	Iberdrola Distribucion	ES	1/2007 - 12/2010	SMM
Generic virtual power plant for optimized micro CHP operation and integration	CET-DTU	DK	9/2007 - 12/2010	SMM
GreenCom	Instituto Superiore Mario Boella, Fraunhofer FIT, Sensing&Control Systems (SCS), In-JeT ApS, Tyndall National Institute - University College Cork, Actua A/S und EnergiMidt A/S	DK	2014 -	SG
Grid Integration of Offshore Windparks	Fraunhofer-Institut für Windenergie und Energiesystemtechnik (IWES)	DE	7/2008 - 6/2011	InSI
Grid4EU	ERDF	EU-FR	4/2011 - 4/2015	InSI
Grid-integration of Electricity Storage	Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V. (FhG)	DE	4/2009 - 3/2011	InSI
GROW-DERS Demonstration of Grid Connected Electricity Systems	KEMA	EU-NL	7/2009 - 7/2011	SpT
Harz.EE-Mobility	University Magdeburg	DE	8/2009 - 7/2011	InSI
Heat Pumps as an active tool in the energy supply system, ForskEL	Inkubator Halden, Norwegian Center of Expertise for Energy and Emission Trading	NO	4/2010 - 7/2012	InSI
HiperDNO	Brunel University	EU-UK	2/2010 - 1/2013	NA-VN
HiT	Salzburg AG	AT	1/2008 - 1/2011	InSI
Hydrogen - Sotavento projects	Gas Natural Fenosa	ES	2005 - 2011	SpT
ICOEUR	TECHNISCHE UNIVERSITAET DORTMUND	EU-DE	1/2009 - 12/2011	NA-ÜN
IMPROSUME - The Impact of Prosumers in a SG based Energy Market	Inkubator Halden, Norwegian Center of Expertise for Energy and Emission Trading	EU-NO	10/2010 - 10/2011	InSI
Information from the electricity grid - remote reading	-	DK	-	SMM

Fortsetzung der Tabelle

Projektname	Organisation	Land	Zeitraum	Kategorie
INOVGRID	EDP Distribucao SA	PT	2007 - 2011	Sonstiges
Installation of remotely read electricity meters	-	DK	2005 - 2010	SMM
INTEGRAL	ECN, Energy research Centre of the Netherlands	EU-NL	2009 - 2010	InSI
Integrated Utilities Business Systems	Enemalta Corporation and Water Services Corporation	MT	9/2008 - 9/2013	SMM
Integration and management of wind power in the Danish electricity system, ForskEL	Department of Energy Technology-Aalborg University	DK	2007 - 6/2009	InSI
Intelligent Energy Management	Linz Strom Netz GmbH	AT	2008 - 2020	SMM
Intelligent Remote Control for Heat Pumps, ForskEL	Nordjysk Elhandel A/S	DK	4/2010 - 3/2011	InSI
Interactive meters, activating price flexible power consumption, ForskEL	DONG Energy A/S	DK	4/2010 - 3/2011	SMM
Internet of Energy	SINTEF	EU-NO	-	InSI
Introduction of emergency Demand Side Response (DSR) programs	PSE Operator S.A.	PL	1/2011 - 3/2012	InSI
iPower	Riso DTU	DK	1/2011 - 12/2016	SMM
IRIN	Verein zur Förderung der wissenschaftlichen Forschung	DE	12/2009 - 5/2011	Sonstiges
Isolves PSSA-M	Austrian Institute of Technology	AT	7/2009 - 1/2012	SMM
Kybernet	INEA d.o.o.	SL	1/2009 - 1/2011	NA-VN
Large-scale demonstration of charging of electric vehicles, ForskEL	ChoosEV A/S	DK	1/2011 - 12/2013	SMM
LASTBEG - Large Scale Tool for Power Balancing in Electric Grid	NANOTECH SAS AIX EN PROVENCE	EU-FR	1/2009 - 10/2009	InSI
LINEAR	VITO	BE	2011 - 2014	InSI
Low Carbon London - A Learning	UK Power Networks (formerly EDF Energy)	UK	1/2011 - 6/2014	SMM
LV Network Templates for a Low-carbon Future	Western Power Distribution	UK	1/2011 - 12/2013	SMM
Manage Smart in SG	Tieto Energy AS	NO	1/2010 - 12/2012	InSI

Fortsetzung der Tabelle

Projektname	Organisation	Land	Zeitraum	Kategorie
Marina power dist. hub with smart-grid functionality, ForskEL	Danish Technological Institute	DK	2/2011 - 1/2012	SMM
Market Based Demand Response	SINTEF Energy Research	NO	2005 - 2008	SMM
Matning 2009	Goteborg Energy AB	SE	3/2006 - 7/2010	SMM
MERGE - Mobile Energy Resources in Grids of Electricity	PUBLIC POWER CORPORATION S.A.	EU-EL	1/2010 - 12/2011	InSI
Micro-Request-Based Aggregation, Forecasting and Scheduling of Energy Demand, Supply and Distribution (MIRACLE)	SAP AG	EU-DE	1/2010 - 1/2013	InSI
Mini Berlin powered by Vattenfall	Vattenfall Europe Innovation GmbH	DE	12/2008 - 11/2010	InSI
Model City Mannheim	MVV Energie	DE	11/2008 - 10/2012	InSI
More Microgrids	ICCS/National Technical University of Athens	EU-GR	1/2006 - 12/2009	InSI, SMM, NA-VN, DSM
More PV2Grid	Fronius International GmbH	AT	3/2010 - 3/2013	InSI
NET-ELAN	Forschungszentrum Jülich GmbH	DE	12/2008 - 11/2011	SpT
Network design and management in a smart city with large deployment of DERR	Ores	BE	12/2010 -	SMM
Netze der Stromversorgung der Zukunft	RWE DAG	DE	2008 - 2011	NA-VN
NextGen	CET-DTU	DK	4/2006 - 6/2010	SMM
NIGHT WIND	NEDERLANDSE ORGANISATIE VOOR TOEGEPAST NATUURWETENSCHAPPELIJK ONDERZOEK	EU-NL	7/2006 - 6/2008	Sonstiges
Open meter	Iberdrola	EU-ES	1/2009 - 6/2011	SMM
Open Node	Atos Origin Sae	EU-ES	1/2010 - 6/2012	InSI

Fortsetzung der Tabelle

Projektname	Organisation	Land	Zeitraum	Kategorie
Opportunities to use Compressed air energy storage for storage of electricity in the electricity system of the future	DTU Mechanical Engineering	DK	12/2005 - 12/2010	SpT
Optimal Power Network design and Operation	Norwegian University of Science and Technology	NO	7/2011 - 6/2014	NA-ÜN
OPTIMATE	TECHNOFI	EU-FR	2009 - 2012	InSI
PEGASE - Pan European Grid Advanced Simulation and state Estimation	TRACTEBEL ENGINEERING S.A.	EU-BE	7/2008 - 6/2012	Sonstiges
Pilot Linky	ERDF	FR	2007 - 2011	SMM
Pilot SMM	Enexis	NL	2007 - 2010	SMM
Pilotprojekt Märkisches Viertel	Vattenfall Europe Distribution Berlin GmbH	DE	1/2010 - 6/2011	SMM
Plug n' play-concept for intelligent indeklimastyring, ForskEL	Neogrid Technologies	DK	1/2011 - 1/2013	InSI
Plugged in Places	Various	UK	4/2010 - 4/2013	DSM
PMU/WAMS target area	Energinet.dk, (TSO)	DK	2006 - 2012	NA-ÜN
Power pit	DONG Energy, (DSO)	DK	2007 - 2009	NA-VN
Price elastic electricity consumption and electricity production in industry	Dansk Energi Analyze a/s	DK	3/2006 - 6/2010	InSI
Price elastic electricity consumption and electricity production in industry	Dansk Energi Analyze a/s	DK	3/2006 - 6/2010	InSI
Price elastic electricity consumption as reserve power - a demonstration project in the horticultural sector	DEG Green Team	DK	4/2006 - 6/2010	InSI
Proactive participation of wind in the electricity markets, ForskEL	EMD International	DK	6/2009 - 12/2010	InSI
Project "Intelligent home"	SEAS-NVE, (DSO)	DK	2009 - 2011	SMM
Project "The Island of Fur on the map"	-	DK	3/2010 - 3/2020	InSI
Project AMR	Vattenfall AB	SE	2006 - 2009	SMM
Prov1Elbil	Danish Technological Institute	DK	1/2009 - 12/2012	InSI

Fortsetzung der Tabelle

Projektname	Organisation	Land	Zeitraum	Kategorie
PV-Island Bornholm, ForskEL	EnergiMidt A/S	DK	1/2010 - 12/2012	InSI
REALISEGRID	RSE	EU-IT	9/2008 - 5/2011	Sonstiges
Real-time demonstration test and evaluation of Bornholm electricity network with high wind power penetration	CET-DTU	DK	10/2009 - 12/2012	SMM
Regenerative Modellregion Harz	Fraunhofer IWES	DE	2008 - 2012	SMM
Regulated power, OUH	-	DK	1/2009 - 6/2009	InSI
Remote Services for CHP, ForskEL	Eurisco	DK	4/2009 - 12/2010	InSI
SACSe	DONG Energy, (DSO)	DK	2008 - 2010	NA-VN
SAFEWIND	RTE	EU-FR	9/2008 - 8/2012	Sonstiges
Second1 - Security concept for DER	EURISCO ApS	DK	3/2010 - 7/2011	SMM
Self-organizing distributed control of a distributed energy system with a high penetration of renew-able energy	DTU Informatics	DK	1/2007 - 6/2010	NA-VN
Service optimization of the distribution network	-	DK	2009 - 2010	InSI
Smart Energy Collective	KEMA	NL	2010 - 2013	NA-VN
Smart green circuits, ESB Networks - SG demonstration project	ESB Networks	IE	1/2010 - 12/2012	NA-VN
SG	SEAS NVE	DK	2008 - 2011	SMM
SG Demonstration System	Arqiva	UK	4/2010 - 4/2011	InSI
SG Fuel Cell CHP on Bornholm, ForskEL	Dantherm Power A/S	DK	1/2011 - 10/2011	InSI
SG Gotland	Vattenfall AB	SE	11/2010 - 12/2015	InSI
SG Task Force project	Energinet.dk, (TSO)	DK	1/2009 - 9/2010	Sonstiges
SGs and energy markets	Cleen Oy	FI	9/2009 - 9/2014	InSI

Fortsetzung der Tabelle

Projektname	Organisation	Land	Zeitraum	Kategorie
Smart Heat Networks	Salzburg AG	AT	3/2010 - 12/2012	InSI
SMM	Fortum	FI	2010 - 2014	SMM
SMM	ZSE Distribucia a.s. - E.ON group	SK	7/2008 - 10/2010	SMM
SMM Multi Utility Pilot	ELMU Halozati Kft. (Electricity, DSO)	HU	2/2011 - 8/2012	SMM
SMM NTA Roll Out	Enexis	NL	7/2010 - 12/2011	SMM
SMM proof of concept	Ores	BE	6/2009 - 9/2012	SMM
Smart neighboring heat supply based on ground heat pumps, ForskEL	Solrod Municipality	DK	1/2011 - 4/2012	InSI
Smart Power System - First trial	ECN, Energy research Centre of the Netherlands	NL	2006 - 2007	InSI
Smart Region	CEZ Distribuce a.s.	CZ	2011 - 2014	NA-VN
Smart Synergy	Salzburg AG	AT	6/2010 - 5/2012	InSI
Smart Watts	Utilicount GmbH & Co. KG, Aachen	DE	11/2008 - 9/2012	SMM
Smart Web Grid	Salzburg AG	AT	3/2011 - 2/2013	InSI
Smartcity Malaga	Endesa	ES	3/2009 - 3/2013	SMM
SmartGen	Sweco Norge AS	EU-NO	10/2010 - 9/2013	InSI
SmartHouse/SmartGrid	SAP Research	EU-DE	-	DSM
STAmi: Advanced Metering Interface	Enel SpA	IT	1/2010 - 1/2011	SMM
Star	Iberdrola Distribucion	ES	4/2010 - 3/2011	SMM
Steinkjer Pilot Project	The Norwegian SG Centre	NO	1/2011 - 12/2015	InSI
Stockholm Royal seaport pre-study phase	Fortum Distribution AB	FI	10/2010 - 3/2011	SMM
Storstad SMM	E.ON Sverige AB	SE	2006 - 2009	SMM
SUMO	Elektro-Slovenija (TSO)	SL	2011 - 2014	Sonstiges

Fortsetzung der Tabelle

Projektname	Organisation	Land	Zeitraum	Kategorie
Supermen	Iskra MIS d.d.	SL	9/2009 - 1/2011	InSI
SUSPLAN	SINTEF ENERGIFORSKNING A/S	EU-NO	9/2008 - 9/2011	Sonstiges
Sustainable urban living	Fortum Electricity solutions and distribution	FI	9/2009 - 12/2013	InSI
Swiss2G	Kraftwerke Oberhasli (KWO)	CH	2010 - 2013	NA-VN
SyM2-Project	E.ON Netz AG, EnBW, RWE DAG	DE	6/2006 - 12/2009	SMM
System output project	-	DK	1/2010 - 12/2010	InSI
Systems with High Level Integration of Renewable Generation Units	Department of Energy Technology - Aalborg University	DK	1/2007 - 12/2009	InSI
Telegestore	Enel Distribuzione	IT	12/2001 - 12/2006	-
Tertiary reserve power with zero CO2 emission	Elektro-Slovenija (TSO)	SL	2011 - 2014	Sonstiges
The cell controller project	Energinet.dk, (TSO)	DK	11/2004 - 10/2011	NA-VN
The metering data processing and central repository concept	PSE Operator S.A.	PL	1/2010 - 3/2011	InSI
Trials with heat pumps on spot agreements	SydEnergi, (DSO)	DK	3/2010 - 12/2011	SMM
TWENTIES	Red Electrica de Espana	EU-ES	2010 - 2013	NA-ÜN
Vehicle to Grid - Interfaces	Salzburg AG	AT	6/2010 - 5/2011	Sonstiges
Vehicle to Grid - Interfaces	Salzburg AG	AT	6/2010 - 5/2011	Sonstiges
Vehicle to Grid - Strategies	Salzburg AG	AT	6/2010 - 12/2012	Sonstiges
Virtual Power Plant	RWE DAG	DE	2008 - 2010	InSI
WAMPAC	Elektro-Slovenija (TSO)	SL	2011 - 2014	Sonstiges
web2energy	HSE AG	EU-DE	1/2010 - 12/2012	SMM, NA-ÜN
WINDGRID	RED ELECTRICA DE ESPANA, S.A.	EU-ES	12/2006 - 12/2009	Sonstiges
Zone concept and smart protection pilot	Fortum Sähkönsiirto Oy	FI	2010 - 2013	NA-VN

Fortsetzung der Tabelle

Projektname	Organisation	Land	Zeitraum	Kategorie
ZUQDE	Salzburg AG	AT	7/2010 - 5/2012	NA-VN

B.1.2 Wichtige Formeln für ökonomische *Demand-Response-Modelle*

Tageszeitliche Lastverschiebungen als Reaktion auf Preisänderungen können mit den Gleichungen (1.2) und (1.3) berechnet werden. Für einen gesamten Tag mit 24 Zeitabschnitten ergibt sich daraus eine 24x24 Matrix für die Kreuzpreiselastizität $\varepsilon[j, k]$, die wie folgt berechnet wird [149, 150]:

$$\begin{bmatrix} \Delta P[1] \\ \vdots \\ \Delta P[k] \\ \vdots \\ \Delta P[24] \end{bmatrix} = \begin{bmatrix} \varepsilon[1,1] & \cdots & \varepsilon[1,k] & \cdots & \varepsilon[1,24] \\ \vdots & \ddots & \vdots & & \vdots \\ \varepsilon[j,1] & \cdots & \varepsilon[j,k] & \cdots & \varepsilon[j,24] \\ \vdots & & \vdots & \ddots & \vdots \\ \varepsilon[24,1] & \cdots & \varepsilon[24,k] & \cdots & \varepsilon[24,24] \end{bmatrix} \times \begin{bmatrix} \Delta p[1] \\ \vdots \\ \Delta p[j] \\ \vdots \\ \Delta p[24] \end{bmatrix}, \quad (\text{B.1})$$

wobei die Eigenpreiselastizitäten $\varepsilon[k, k]$ in der Diagonalen der Matrix stehen.

Abb. B.1 stellt den grundsätzlichen Zusammenhang zwischen Stromangebot und Stromnachfrage dar. Es ist zu erkennen, dass bei einer großen Stromnachfrage (*peak demand*, z.B. P_2) eine kleine Lastreduktion (ΔP) große Auswirkung auf den Strompreis (Δp) hat.

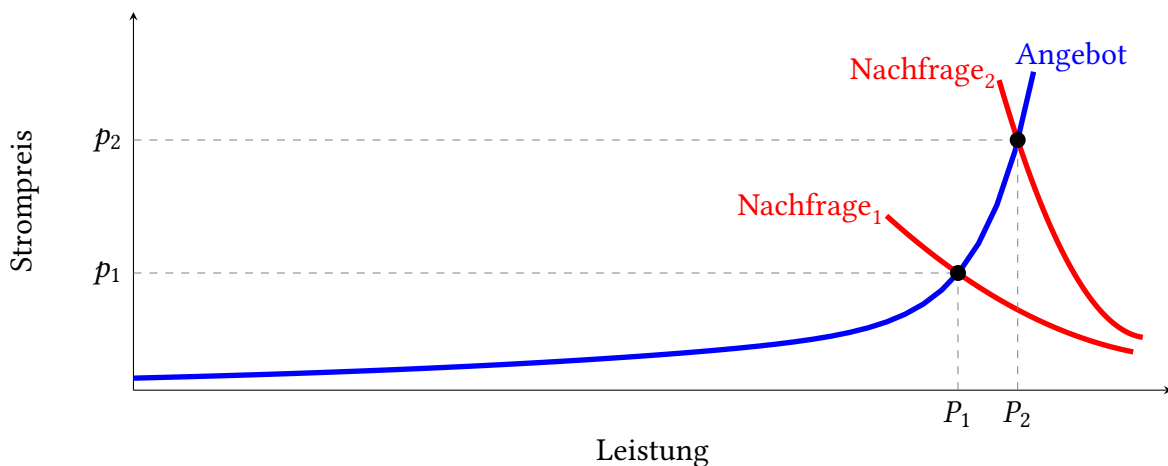


Abbildung B.1: Einfluss des Preises auf die Änderung der Stromnachfrage (modifiziert nach [150, 157, 289]); blaue Angebotskurve: zu einem geringen Strompreis kann auch nur eine geringe Menge (Last) angeboten werden; rote Nachfragekurven: bei einem geringeren Strompreis wird eine größere Menge (Last) nachgefragt

B.1.3 Heiz- und Klimatechniksysteme

In diesem Abschnitt werden ausgewählte Heiz- und Klimatechniksysteme vorgestellt, die in Wohngebäuden eingesetzt werden und zum Flexibilitätspotenzial bei Haushaltsstromkunden beitragen [290–292]. Die Beschreibungen der einzelnen Heiz- und Klimatechniksysteme ergänzen die Aufzählungen in Tabelle B.2.

Elektrische Widerstandsheizung: Bei elektrischen Widerstandsheizungen wird elektrischer Strom durch einen elektrischen Leiter (Heizdraht) mit einem Ohmschen Widerstand geleitet. Die elektrische Arbeit wird dabei direkt in Wärme umgewandelt. Hinsichtlich des Aufbaus und der Wirkweise werden Elektro-Heizungssysteme in folgende vier Heizungssysteme unterteilt:

1. **Elektrische Direktheizungen** sind durch die Gleichzeitigkeit von momentanem Wärmebedarf und entnommener elektrischer Leistung aus dem Netz gekennzeichnet. Die Wärmeenergie steht unmittelbar und damit direkt zur Verfügung, sobald der elektrische Strom durch den Widerstandsdraht geleitet wird [292]. Bei geringer Außentemperatur und damit einhergehendem hohen Wärmebedarf sorgt die Gleichzeitigkeit für hohe Netzbelastungen mit ausgeprägten Lastspitzen. Zu den elektrischen Direktheizungen werden elektrische Direktheizgeräte (Konvektoren, Radiatoren, Strahler, Heizlüfter), Flächen-direktheizungen (Flächenheizung, Deckenstrahlungs-, Fußbodenheizung) und zentrale Direktheizungen (Elektroheizkessel) gezählt [291].
2. **Elektrische Speicherheizungen** sind in der Lage Wärmeenergie zu speichern. Die Aufnahme der elektrischen Energie und Umwandlung in Wärme erfolgt zeitlich entkoppelt von der Wärmeabgabe. Am weitesten verbreitet sind **Nachtspeicherheizungen**, bei denen die Wärmeenergie des elektrischen Widerstandsdrahtes mit hoher Wärmekapazität z.B. in Magnesitsteinen und zumeist nachts gespeichert wird [291, 292]. Die in dem Speicherkern gespeicherte Wärme wird dank Wärmedämmmaterial über einige Stunden gehalten. Die Wärmeabgabe erfolgt, indem die Raumluft über einen eingebauten Ventilator angesaugt, durch das Gerät geleitet und stark erhitzt wird. Eine Luftmischklappe sorgt dafür, dass durch eine automatische Beimischung von nicht erwärmter Raumluft zu der erwärmten Luft eine konstante Ausblastemperatur von ca. 80-100°C erreicht wird [291].

Nachtspeicherheizungen beheizen hauptsächlich einzelne Räume. Dagegen sind **elektrische Zentralspeicherheizungen** große Speicherheizgeräte, die meist im Keller angeordnet sind und anstelle von veralteten Heizkesseln aufgestellt werden. Die Wärme wird bei elektrischen Zentralspeicherheizungen entweder in einem festen Stoff (Gusseisen, Magnesit) oder in Wasser gespeichert und durch Warmwasser-Radiatorheizungen, Warmwasser-Fußbodenheizungen oder Luftheizungssysteme im Wohngebäude verteilt. Deswegen wird zwischen Feststoff- und Wasser-Zentralspeicherheizung unterschieden [291].

Bei der **elektrischen Fußbodenspeicherheizung** werden sogenannte Heizmatten in den Fußboden integriert, wo die elektrische Energie in Wärme umgewandelt wird und zeit-

lich verzögert über die Oberfläche des Fußbodens an den zu beheizenden Raum abgegeben wird. Als Wärmespeicher dient der Fußboden selbst [291].

3. **Elektrische Luftheizung:** Bei den elektrischen Luftheizungen wird zwischen Lüftungs-Direktheizgeräten und Lüftungs-Speicherheizgeräten unterschieden. Beide Gerätearten bestehen aus einem Lüftungsmodul und einem Heizmodul. Lüftungs-Speicherheizgeräte besitzen zusätzlich zum Lüfter und Raumtemperaturregler auch einen integrierten Wärmespeicher [291].
4. **Elektrische Wärmepumpe:** Die elektrische Wärmepumpe ist ein Maschine, die der Umwelt (Grund- und Oberflächenwasser, Erdreich (Sole), Umgebungsluft) Wärme bei einem niederen Temperaturniveau entzieht und dem Heizsystem zuführt. Durch Zufuhr von elektrischer Energie wird der Wärmestrom bei einem höheren Temperaturniveau wieder abgegeben. Somit erzeugen Wärmepumpen keine Wärme, sie verdichten sie vielmehr [291]. Typische Arbeitszahlen von Wärmepumpen als Maßstab für die Effizienz liegen im Bereich von 2,3 bis 3,6, wobei Wärmepumpen, die die Wärme des Bodens nutzen, die höchste Effizienz haben [293].

Wärmepumpen sind wesentlich effizienter als elektrische Widerstandsheizungen. Die durchschnittliche Heizleistung bei installierten Anlagen beträgt für Sole/Wasser-Wärmepumpen 10 kW, für Luft/Wasser-Wärmepumpen 12 kW und für Wasser/Wasser-Wärmepumpen 14 kW [294]. Der Anlagenbestand für Heizungswärmepumpen in Deutschland belief sich im Jahr 2012 auf ca. 500 000 Anlagen, wobei ca. 96% der Anlagen in Wohngebäuden installiert sind [295].

Elektrische Warmwasserbereitung: Anlagen für die elektrische Warmwasserbereitung bieten ähnlich wie elektrische Speicherheizungen die Möglichkeit, den elektrischen Bedarf zu ver gleichmäßigen [296]. Unter der Vielzahl an technischen Geräten zur elektrischen Warmwasserbereitung eignen sich vor allem elektrische Widerstandsheizungen mit integriertem Warmwasserspeicher („Elektro-Standspeicher“) für *Demand Response* [291, 296].

Umwälzpumpe: Die Umwälzpumpe sorgt dafür, dass das Heizwasser im Heizungsrohr als Wärmeträger zirkuliert und ist damit ein wichtiger Bestandteil der Wärmeverteilung in Wohngebäuden [291].

Raumklimasysteme: Raumklimasysteme werden im Allgemeinen in (mobile und stationäre) Raumklimageräte und Raumklimaanlagen unterteilt. Raumklimageräte dienen der Klimatisierung von einzelnen Räumen, wohingegen Raumklimaanlagen eher in großen Gebäuden, wie Krankenhäusern oder Bürogebäuden, eingesetzt werden. Eine weitere Differenzierung und detailliertere Beschreibung von Raumklimasystemen ist in [291] zu finden.

Tabelle B.2: Technisches Lastverlagerungspotenzial verschiedener Anwendungen im Haushalt in Deutschland im Jahr 2013 (aus [72, 225, 290, 291, 296–299], eigene Berechnungen): * - Anwendung ist zeitlich verschiebbar, ** - Anwendung nutzt Speicherwirkung von Gebäuden oder hat internen Speicher.

Anwendung ¹	Marktdurchdringung ² (%)	Gebrauchshäufigkeit	Gebrauchsdauer je Vorgang (h)	Spitzenleistung (kW)	Verbrauch je Vorgang (kWh)	Verschiebepotenzial je N und Monat (kWh)
Waschmaschine*	94.5	12.2/Monat	-	2.3	1	11.6
Wäschetrockner*	39.4	9.7/Monat	-	3.1	2.4	7.9
Geschirrspüler*	67.3	3.8/Woche	-	2.3	1.1	8.7
↳ Waschen, Trocknen, Spülen:		12.4% des Gesamtstromverbrauchs (16.9 TWh)				
Kühlschrank**	99.7	8h/Tag	0.33	0.09	0.031	21.7
Gefrierschrank**	50.5	8h/Tag	0.33	0.12	0.037	21.9
↳ Kühlen und Gefrieren:		16.7% des Gesamtstromverbrauchs (22.8 TWh)				
Nachtspeicher-Elektroheizung**	4	80/Jahr	k.A.	1.5-10	k.A.	56.2
Elektrische Fußbodenheizung**	k.A.	k.A.	k.A.	1-5	k.A.	k.A.
Elektrische Wärmepumpe**	1.4 ³	k.A.	k.A.	12 ⁴	k.A.	k.A.
Elektrische Direktheizung**	1.5	k.A.	k.A.	k.A.	k.A.	8.3
Raumklimaanlage**	1.4	120h/Jahr	6	2.5	15	4.2
↳ Elektrisches Heizen und Kühlen:		12.5% des Gesamtstromverbrauchs (24.4 TWh)				
Umwälzpumpe** (Heizwasser, Brauchwasser)	93	12h/Tag im Winter (8 Monate)	12	0.10	1.2	22.3
Elektrische Warmwasseraufbereitung**	28	k.A.	3.8 ⁵	2-6 ⁶	22.8	14.6
↳ Warmes Wasser:		14.8% des Gesamtstromverbrauchs (20.2 TWh)				

¹Beschreibung ausgewählter Anwendungen und Heiztechniken in Anhang B.1

²Bei den Berechnungen in [298] wurde von 40.032 Mio. Haushalten (Wohnheiten) in Deutschland im Jahr 2013 ausgegangen.

³In 2015 lag der Anteil an Wärmepumpen bezogen auf alle Beheizungssysteme in Wohnungsneubauten bei 20.9%.

⁴Durchschnittliche Heizleistung

⁵Aufheizzeit auf 60°C bei maximaler Heizleistung von 6 kW

⁶Heizleistung bei Elektro-Standspeichern mit 200-400 l Speicherinhalt

Wärmegestehungskosten: Die Wärmegestehungskosten verschiedener Heiztechnologien in Einfamilienhäusern sind in Abb. B.2 gegenübergestellt. Die Daten zur Generierung der Abbildung stammen aus [300–302]. Die in Abb. B.2 angegebenen Kosten beinhalten Investitionskosten und betriebsgebundene Kosten. Bei netzgebundenen Systemen sind außerdem Netzkosten berücksichtigt. Es gilt zu beachten, dass sich betriebs- und netzgebundene Kosten von Jahr zu Jahr verändern können. Die Zahlen in Abb. B.2 beziehen sich auf Durchschnittswerte der Jahre 2006 bis 2009. Die Entwicklung von Rohstoffpreisen, wie beispielsweise Erdgaspreise, sind in Abb. B.2 nicht berücksichtigt.

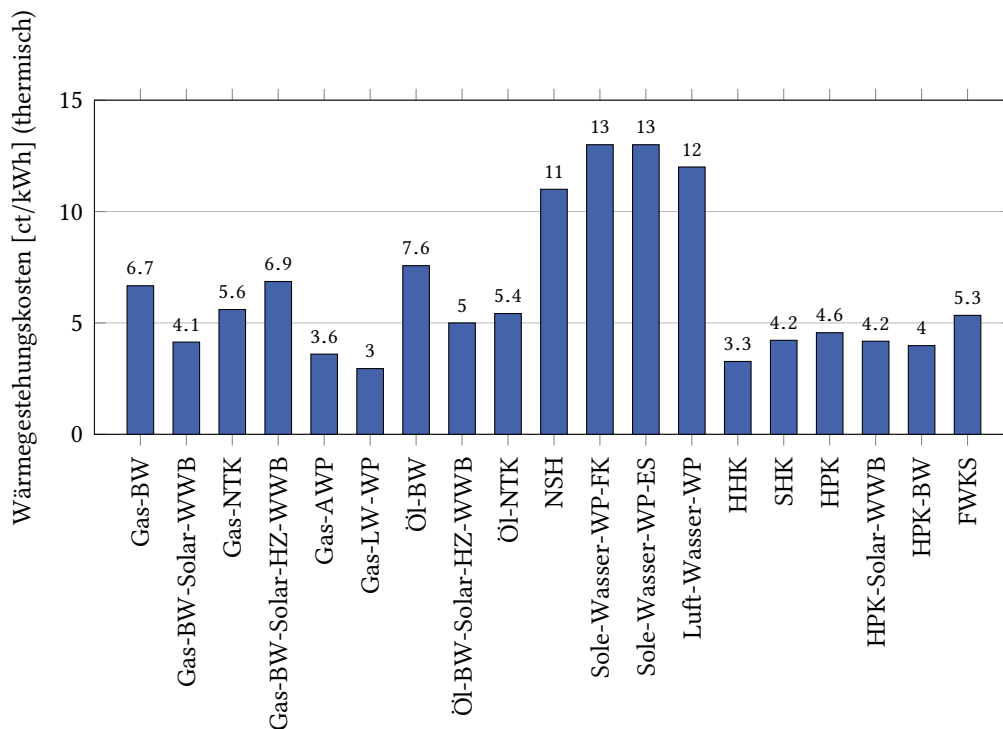


Abbildung B.2: Wärmegestehungskosten verschiedener Heiztechnologien in Einfamilienhäusern (modifiziert nach [300–302]); Abk.: Gas-BW: Erdgas-Brennwertkessel; Gas-BW-Solar-WWB: Erdgas-Brennwertkessel mit Solarkollektoren und -speicher zur Warmwasserbereitung; Gas-NTK: Erdgas-Niedertemperaturkessel; Gas-BW-Solar-HZ-WWB: Erdgas-Brennwertkessel mit Solarkollektoren und -speicher zur Heizungsunterstützung und Warmwasserbereitung; Gas-AWP: Erdgas-Absorptionswärmepumpe; Gas-LW-WP: Erdgasmotorische Luft-Wasser-Wärmepumpe; Öl-BW: Heizöl-Brennwertkessel; Öl-BW-Solar-HZ-WWB: Heizöl-Brennwertkessel mit Solarkollektoren und -speicher zur Heizungsunterstützung und Warmwasserbereitung; Öl-NTK: Heizöl-Niedertemperaturkessel; NSH: Nachtspeicherheizung; Sole-Wasser-WP-FK: Sole-Wasser-Wärmepumpe mit Flächenkollektor; Sole-Wasser-WP-ES: Sole-Wasser-Wärmepumpe mit Erdsonden; Luft-Wasser-WP: Luft-Wasser-Wärmepumpe; HHK: Holzhackschnitzelkessel; SHK: Scheitholzvergaserkessel; HPK: Holzpelletkessel; HPK-Solar-WWB: Holzpelletkessel mit Solarkollektoren und -speicher zur Warmwasserbereitung; HPK-BW: Holzpellet-Brennwertkessel; FWKS: Fernwärmekompaaktstation (Fernwärmekompaaktstationen stellen die Schnittstelle in Gebäuden oder Wohnsiedlungen zur Nutzung der Wärmeenergie aus Wärmenetzen oder Industrieprozessen dar. Die Kompaktstation koppelt Wärmeenergie aus einem (Fern-) Wärmenetz aus und gibt diese beispielsweise an die Gebäude- bzw. Raumheizung ab.).

B.2 Methodik

B.2.1 Anforderungsanalyse

Tabelle B.3 ergänzt die beiden Tabellen 2.1 und 2.2 insofern, als weitere Einflussfaktoren auf das Verbrauchsverhalten bei Haushaltsstromkunden berücksichtigt werden. Bei diesen Einflussfaktoren handelt es sich größtenteils um personenbezogene Daten, die in Feldstudien aus Datenschutzgründen schwer bis gar nicht erhoben werden können oder später aus dem Datensatz entfernt werden.

Tabelle B.3: Fortsetzung der Checklisten (Tabellen 2.1 und 2.2) für *Smart-Meter*-Datensätze für die datengetriebene Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens; Kennzeichnung: (+) = signifikanter Einfluss, (–) = geringer Einfluss, * = Bewertung des Einflusses dieser Eigenschaften anhand der Analysen in [17].

Nr.: Eigenschaft des <i>Smart-Meter</i> -Datensatzes	A1	A2	A3	A4
E14: Anzahl der Haushaltsbewohner*	(+) Die Anzahl der Haushaltsbewohner beeinflusst stark den Stromverbrauch und den Verlauf des Lastprofils. Je höher der Stromverbrauch, desto größer ist das Lastverlagerungspotenzial. Ein größeres Lastverlagerungspotenzial wirkt sich positiv auf die Erfüllung der Aufgabenstellungen A1 bis A4 aus.			
E15: Größe der Wohnung/ Haus*	(+) Die Größe der Wohnung bzw. des Hauses beeinflusst die Höhe des Stromverbrauchs (positive Korrelation zwischen Größe der Wohnung bzw. des Hauses und Stromverbrauch). Je höher der Stromverbrauch, desto größer ist das Lastverlagerungspotenzial. Ein größeres Lastverlagerungspotenzial wirkt sich positiv auf die Erfüllung der Aufgabenstellungen A1 bis A4 aus.			
E16: Standort der Wohnung/ Haus (Stadt, Land)*	(+) Der Standort der Wohnung bzw. des Hauses beeinflusst die Höhe des Stromverbrauchs (positive Korrelation zwischen Standort der Wohnung bzw. des Hauses und Stromverbrauch). Je höher der Stromverbrauch, desto größer ist das Lastverlagerungspotenzial. Ein größeres Lastverlagerungspotenzial wirkt sich positiv auf die Erfüllung der Aufgabenstellungen A1 bis A4 aus.			
E17: Nettoeinkommen der Bewohner*	(+) In [17] wurde das Nettoeinkommen als wichtige Einflussgröße für den Stromverbrauch identifiziert (pos. Korrelation zwischen Nettoeinkommen und Stromverbrauch). Je höher der Stromverbrauch, desto größer ist das Lastverlagerungspotenzial. Ein größeres Lastverlagerungspotenzial wirkt sich positiv auf die Erfüllung der Aufgabenstellungen A1 bis A4 aus.			
E18: Alter der Bewohner*	Bezüglich des Einflusses von E18 auf den Stromverbrauch wurden in [17] differenzierte Aussagen gemacht. Deswegen kann auch keine eindeutige Bewertung hinsichtlich des Einflusses von E18 auf die Erfüllung der Aufgabenstellungen vorgenommen werden.			
E19: Beschäftigungsverhältnis der Bewohner*	Bezüglich des Einflusses von E19 auf den Stromverbrauch wurden in [17] differenzierte Aussagen gemacht. Deswegen kann auch keine eindeutige Bewertung hinsichtlich des Einflusses von E19 auf die Erfüllung der Aufgabenstellungen vorgenommen werden.			

B.2.2 Benchmarkgenerator

Die Preiselastizitätswerte in der nachfolgenden *Lookup*-Tabelle basieren auf dem *MeRegio*-Datensatz¹ [151]. In der *MeRegio*-Feldstudie nahmen ca. 1000 Haushalte und kleinere Gewerbetunden aus dem süddeutschen Raum teil. Es wurde versucht, das Verbrauchsverhalten der Stromkunden sowohl mithilfe eines Preissignals als auch mithilfe eines Steuersignals zu verändern, wobei die Höhe des Preissignals für eine Stunde gültig war. Die Haushalte von etwa 30% der Studienteilnehmern waren mit elektrischen Verbrauchern, dezentralen Erzeugungsanlagen und Stromspeichern ausgestattet, die automatisiert mithilfe einer Steuerbox auf die Anreizsignale reagierten (Automatisches *Demand Response*, siehe Abschnitt 1.2.5). Die übrigen 70% der Haushalte reagierte manuell auf das Preissignal [303]. Weitere Detail zum *MeRegio*-Projekt sind in [303] zu finden.

Tabelle B.4: *Demand-Response*-Simulator: *Lookup*-Tabelle mit den Werten für die Preiselastizität $\epsilon[k]$ in Gleichung (2.13) (aus [17]); Kennzeichnung mit *: keine Werte vorhanden, daher mit Wert 0,00 ersetzt

	NachtKurz	NachtMittel	NachtLang	TagKurz	TagMittel	TagLang
FrühlingWerktag	0,00	-0,03	0,00	-0,33	-0,38	-0,36
FrühlingSamstag	0,00	-0,10	0,00	-0,20	-0,52	-0,43
FrühlingSonntag	-0,14	-0,20	0,00	-0,23	-0,17	-0,29
SommerWerktag	-0,04	-0,10	0,00	-0,21	-0,32	-0,29
SommerSamstag	-0,02	-0,06	0,00	-0,20	-0,45	0,00
SommerSonntag	0,00	-0,16	0,00*	-0,10	-0,17	0,00*
HerbstWerktag	0,00	-0,11	-0,05	-0,25	-0,33	-0,34
HerbstSamstag	-0,08	-0,26	-0,32	-0,15	-0,42	-0,22
HerbstSonntag	0,00	-0,18	-0,05	-0,10	-0,20	-0,27
WinterWerktag	0,00	-0,13	-0,16	-0,30	-0,39	-0,18
WinterSamstag	0,00	-0,24	-0,13	-0,11	-0,52	-0,41
WinterSonntag	-0,19	-0,27	0,00	-0,26	-0,24	-0,35

Jahreszeit: **Frühling**=März-Mai; **Sommer**=Juni-August; **Herbst**=Sep.-Nov.; **Winter**=Dez.-Feb.

Wochentag: **Werktag**=Montag-Freitag; **Samstag**; **Sonntag**

Tageszeit: **Nacht**=22:30 bis 6:30; **Tag**=6:30 bis 22:30

Tarifstufendauer: **Kurz**=1 Stunde; **Mittel**=2-3 Stunden; **Lang**=4-6 Stunden

¹Die Rohdaten aus dem *MeRegio*-Projekt sind nicht frei verfügbar.

B.2.3 Parametereinstellungen des Benchmarkgenerators zur Erzeugung der Benchmarkdatensätze

Benchmarkdatensatz zur systematischen Untersuchung verschiedener Imputationsverfahren

In Tabelle B.5 sind die Parametereinstellungen des Benchmarkgenerators zur Erzeugung des Benchmarkdatensatzes zur systematischen Untersuchung verschiedener Imputationsverfahren.

Tabelle B.5: Parametereinstellungen des Benchmarkgenerators zur Erzeugung des Benchmarkdatensatzes zur systematischen Untersuchung verschiedener Imputationsverfahren (Beschreibung des Benchmarkdatensatzes in Abschnitt 2.3.3.1). Mit den gewählten Parametereinstellungen können 198 unterschiedliche Lastzeitreihen generiert werden. Die Zahl 198 ergibt sich aus der Kombination der drei Werte für β_M bzw. β_S mit den sechs Werten für b_{nan}^n und den elf Werten für $\mu_{nan,g}$ bzw. $\mu_{nan,s}$ mit $\mu_{nan,g} = \mu_{nan,s}$.

Parameter	Symbol	Parameter-variation	Parameterwert bzw. -bereich
Lastsignalgenerator			
Anzahl Haushalte	N	nein	$N = 1$
Zeitreihenlänge	K	nein	$K = 35040$
Abtastperiodendauer	T_S	nein	$T_S = 15 \text{ min}$
Gesamtstromverbrauch	μ_E	nein	$\mu_E = 3200 \text{ kWh}$
Varianz der Gesamtstromverbräuche	σ_E^2	nein	$\sigma_E^2 = 0 \text{ kWh}$
Gewichtungsfaktor der Trendkomponente	β_T	nein	$\beta_T = 0.05$
Gewichtungsfaktor der exogenen Komponente	β_E	nein	$\beta_E = 0.00$
Gewichtungsfaktor der periodischen Komponente	β_P	nein	$\beta_P = 0.05$
Gewichtungsfaktor der Musterkomponente	β_M	ja	$\beta_M \in \{0.90, 0.63, 0.36\}$
Gewichtungsfaktor der stochastische Komponente	β_S	ja	$\beta_S \in \{0.00, 0.27, 0.54\}$
DR-Simulator			
Synthetischer Stromtarif	–	nein	FIXED
Festlegung: R-, SR- oder NR-Haushalt	–	nein	NR
Messfehlergenerator			
Anzahl an Bereichen mit fehlenden Werten	b_{nan}^n	ja	$b_{nan}^n \in \{0, 1, 2, 3, 4, 5\}$
Prozentualer Anteil an Bereichen mit fehlenden Werten	$\mu_{nan,g}$	ja	$\mu_{nan,g} \in \{0, 2, 4, \dots, 20\}$
Varianz von $\mu_{nan,g}$	$\sigma_{nan,g}^2$	nein	$\sigma_{nan,g}^2 = 0$
Prozentualer Anteil an einzelnen fehlenden Werten	$\mu_{nan,s}$	ja	$\mu_{nan,s} \in \{0, 2, 4, \dots, 20\}$
Varianz von $\mu_{nan,s}$	$\sigma_{nan,s}^2$	nein	$\sigma_{nan,s}^2 = 0$
Prozentualer Anteil an Ausreißern	μ_a	nein	$\mu_a = 0$
Varianz von μ_a	σ_a^2	nein	$\sigma_a^2 = 0$
Untere Intervallgrenze des Ausreißermultiplikators	a	nein	$a = 1$
Obere Intervallgrenze des Ausreißermultiplikators	b	nein	$b = 1$

Benchmarkdatensatz zur Ermittlung von optimalen Hampel-Filterparametern

In Tabelle B.6 sind die Parametereinstellungen des Benchmarkgenerators zur Erzeugung des Benchmarkdatensatzes zur Ermittlung von optimalen Hampel-Filterparametern gezeigt.

Tabelle B.6: Parametereinstellungen des Benchmarkgenerators zur Erzeugung des Benchmarkdatensatzes zur Ermittlung von optimalen Hampel-Filterparametern (Beschreibung des Benchmarkdatensatzes in Abschnitt 2.3.3.2). Mit den gewählten Parametereinstellungen können 48 unterschiedliche Lastzeitreihen generiert werden. Die Zahl 48 ergibt sich aus der Kombination der vier Werte für K bzw. T_S mit den drei Werten für β_M bzw. β_S und den vier Werten für μ_a .

Parameter	Symbol	Parameter-variation	Parameterwert bzw. -bereich
Lastsignalgenerator			
Anzahl Haushalte	N	nein	$N = 1$
Zeitreihenlänge	K	ja	$K \in \{105120, 35040, 17520, 8760\}$
Abtastperiodendauer	T_S	ja	$T_S \in \{5, 15, 30, 60\}$ min
Gesamtstromverbrauch	μ_E	nein	$\mu_E = 3200$ kWh
Varianz der Gesamtstromverbräuche	σ_E^2	nein	$\sigma_E^2 = 0$ kWh
Gewichtungsfaktor der Trendkomponente	β_T	nein	$\beta_T = 0.05$
Gewichtungsfaktor der exogenen Komponente	β_E	nein	$\beta_E = 0.00$
Gewichtungsfaktor der periodischen Komponente	β_P	nein	$\beta_P = 0.05$
Gewichtungsfaktor der Musterkomponente	β_M	ja	$\beta_M \in \{0.90, 0.63, 0.36\}$
Gewichtungsfaktor der stochastische Komponente	β_S	ja	$\beta_S \in \{0.00, 0.27, 0.54\}$
DR-Simulator			
Synthetischer Stromtarif	–	nein	FIXED
Festlegung: R-, SR- oder NR-Haushalt	–	nein	NR
Messfehlergenerator			
Anzahl an Bereichen mit fehlenden Werten	b_{nan}^n	nein	$b_{nan}^n = 0$
Prozentualer Anteil an Bereichen mit fehlenden Werten	$\mu_{nan,g}$	nein	$\mu_{nan,g} = 0$
Varianz von $\mu_{nan,g}$	$\sigma_{nan,g}^2$	nein	$\sigma_{nan,g}^2 = 0$
Prozentualer Anteil an einzelnen fehlenden Werten	$\mu_{nan,s}$	nein	$\mu_{nan,s} = 0$
Varianz von $\mu_{nan,s}$	$\sigma_{nan,s}^2$	nein	$\sigma_{nan,s}^2 = 0$
Prozentualer Anteil an Ausreißern	μ_a	ja	$\mu_a \in \{2, 4, \dots, 8\}$
Varianz von μ_a	σ_a^2	nein	$\sigma_a^2 = 0$
Untere Intervallgrenze des Ausreißermultiplikators	a	nein	$a = 1.5$
Obere Intervallgrenze des Ausreißermultiplikators	b	nein	$b = 5$

Benchmarkdatensätze zur Validierung der analyseorientierten Methoden der Pipeline

In Tabelle B.7 sind die Parametereinstellungen des Benchmarkgenerators zur Erzeugung von Benchmarkdatensätzen zur systematischen Untersuchung der analyseorientierten Methoden der Pipeline gezeigt.

Tabelle B.7: Parametereinstellungen des Benchmarkgenerators zur Erzeugung von Benchmarkdatensätzen zur systematischen Untersuchung der analyseorientierten Methoden der Pipeline (Beschreibung der Benchmarkdatensätze in Abschnitt 2.3.3.3). Mit den gewählten Parametereinstellungen können 19 unterschiedliche Benchmarkdatensätze mit jeweils 210 Haushaltslastzeitreihen generiert werden. Die Zahl 19 ergibt sich aus den zehn Parameterwerten für β_M bzw. β_S .

Parameter	Symbol	Parameter-variation	Parameterwert bzw. -bereich
Lastsignalgenerator			
Anzahl Haushalte	N	nein	$N = 210$
Zeitreihenlänge	K	nein	$K = 35040$
Abtastperiodendauer	T_S	nein	$T_S = 15 \text{ min}$
Gesamtstromverbrauch	μ_E	nein	$\mu_E = 3200 \text{ kWh}$
Varianz der Gesamtstromverbräuche	σ_E^2	nein	$\sigma_E^2 = 1000 \text{ kWh}$
Gewichtungsfaktor der Trendkomponente	β_T	nein	$\beta_T = 0.05$
Gewichtungsfaktor der exogenen Komponente	β_E	nein	$\beta_E = 0.00$
Gewichtungsfaktor der periodischen Komponente	β_P	nein	$\beta_P = 0.05$
Gewichtungsfaktor der Musterkomponente	β_M	ja	$\beta_M \in \{0.90, 0.85, \dots, 0.05, 0.00\}$
Gewichtungsfaktor der stochastische Komponente	β_S	ja	$\beta_S \in \{0.00, 0.05, \dots, 0.85, 0.90\}$
DR-Simulator			
Synthetischer Stromtarif	je 70 Haushalte in einer der drei Tarifgruppen „FIXED“, „TOU2“ und „TOU3“		
Festlegung: R-, SR- oder NR-Haushalt	Tarifgruppe „FIXED“: alle Haushalte sind NR-Haushalte; Tarifgruppe „TOU2“: 10 Haushalte sind NR-Haushalte, 10 Haushalte sind SR-Haushalte, 50 Haushalte sind R-Haushalte; Tarifgruppe „TOU3“: 10 Haushalte sind NR-Haushalte, 10 Haushalte sind SR-Haushalte, 50 Haushalte sind R-Haushalte		
Messfehlergenerator			
Anzahl an Bereichen mit fehlenden Werten	b_{nan}^n	nein	$b_{nan}^n = 2$
Prozentualer Anteil an Bereichen mit fehlenden Werten	$\mu_{nan,g}$	nein	$\mu_{nan,g} = 3$
Varianz von $\mu_{nan,g}$	$\sigma_{nan,g}^2$	nein	$\sigma_{nan,g}^2 = 0.75$
Prozentualer Anteil an einzelnen fehlenden Werten	$\mu_{nan,s}$	nein	$\mu_{nan,s} = 3$
Varianz von $\mu_{nan,s}$	$\sigma_{nan,s}^2$	nein	$\sigma_{nan,s}^2 = 0.75$
Prozentualer Anteil an Ausreißern	μ_a	nein	$\mu_a = 3$
Varianz von μ_a	σ_a^2	nein	$\sigma_a^2 = 0.75$
Untere Intervallgrenze des Ausreißermultiplikators	a	nein	$a = 1.5$
Obere Intervallgrenze des Ausreißermultiplikators	b	nein	$b = 5$

B.2.4 Datenvorverarbeitung

Weitere Ergebnisse der vergleichenden Bewertung verschiedener Imputationsverfahren bei fehlerbehafteten Lastzeitreihen

Die Struktur des Datensatzes des ersten Testverfahrens (siehe Abschnitt 2.4.1.1) ist in Tabelle B.5 dargestellt. Des Weiteren sind in Abb. 2.6 Teile der Ergebnisse des Testverfahrens präsentiert. An dieser Stelle folgen weitere Ergebnisse des Testverfahrens, wie die Abb. B.3 und B.4 für $b_{nan}^n = 1, 2, 3, 4$ und Tabelle B.8, die die Basis für die Erstellung der Abbildungen bildet.

Das in Tabelle 1.3 beschriebene autoregressive Modell $AR(n_a)$ der Ordnung n_a ist wie folgt definiert:

$$\hat{y}[k] + \sum_{i_a=1}^{n_a} a_{i_a} y[k - i_a] = e[k], \quad (\text{B.2})$$

wobei $e[k]$ ein gaußverteiltes weißes Rauschen ist. Die Modellstruktur der AR-Modelle kann durch die Angabe der maximalen Modellordnung n_a auf eine Maximalanzahl an Parameterwerten begrenzt werden. Als Datenbasis für die Schätzung der Modellparameter gelten die n_a Abtastzeitpunkte *vor* und *nach* einer *NaN*-Lücke. In einem weiteren Schritt wird die tatsächliche Ordnung des AR-Modells mithilfe des Informationskriteriums von Akaike (engl. *Akaike information criterion*, AIC) ausgewählt. Das AR-Modell schätzt die fehlenden *NaN*-Werte von beiden Seiten und liefert geeignete Ersatzwerte.

Tabelle B.8: Das Ergebnis des Vergleichs verschiedener Imputationsverfahren bei fehlerbehafteten Lastzeitreihen (siehe Versuchsbeschreibung in Abschnitt 2.4.1.1) ist hier als mehrseitige Übersichtstabelle dargestellt. Eine übersichtlichere Darstellung der Werte aus Tabelle B.8 ist in Abb. 2.6, B.3 und B.4 zu finden. Die Bezeichnungen der Zeitreihen in der ersten Spalte geben Aufschluss über die Signalcharakteristik der Zeitreihen. Für die Zeitreihen „HH001“ gilt $\beta_T = 0.05$, $\beta_E = 0.00$, $\beta_P = 0.05$, $\beta_M = 0.90$ und $\beta_S = 0.00$. Für die Zeitreihen „HH002“ gilt $\beta_T = 0.05$, $\beta_E = 0.00$, $\beta_P = 0.00$, $\beta_M = 0.63$ und $\beta_S = 0.27$. Für die Zeitreihen „HH003“ gilt $\beta_T = 0.05$, $\beta_E = 0.00$, $\beta_P = 0.05$, $\beta_M = 0.36$ und $\beta_S = 0.54$.

Zeit- reihe	$\frac{I_{\text{nan.s.}}}{I_{\text{nan.g.}}}$	b_{nan}	r^{NH}	$Q_{\text{C,RMSE}}^{\text{NH}}$	r^{LOCf}	$Q_{\text{C,RMSE}}^{\text{LOCf}}$	r^{NOCB}	$Q_{\text{C,RMSE}}^{\text{NOCB}}$	r^{NEOI}	$Q_{\text{C,RMSE}}^{\text{NEOI}}$	r^{lin}	$Q_{\text{C,RMSE}}^{\text{lin}}$	r^{pchip}	$Q_{\text{C,RMSE}}^{\text{pchip}}$	r^{DCTPL}	$Q_{\text{C,RMSE}}^{\text{DCTPL}}$	r^{AR6}	$Q_{\text{C,RMSE}}^{\text{AR6}}$	r^{AR672}	$Q_{\text{C,RMSE}}^{\text{AR672}}$
HH001	0	0	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH001	0	1	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH001	0	2	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH001	0	3	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH001	0	4	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH001	0	5	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH001	2	0	0.9984	0.0110	0.9999	0.0030	0.9999	0.0030	0.9999	0.0029	1.0000	0.0004	1.0000	0.0002	1.0000	0.0001	1.0000	0.0004	1.0000	0.0005
HH001	2	1	0.9981	0.0120	0.9831	0.0366	0.9926	0.0238	0.9876	0.0308	0.9920	0.0247	0.9926	0.0238	0.9878	0.0306	0.9953	0.0190	0.9953	0.0189
HH001	2	2	0.9972	0.0146	0.9830	0.0360	0.9920	0.0247	0.9865	0.0320	0.9896	0.0281	0.9889	0.0290	0.9864	0.0321	0.9927	0.0236	0.9932	0.0226
HH001	2	3	0.9966	0.0160	0.9824	0.0365	0.9670	NaN	0.9761	0.0428	0.9788	0.0402	0.9784	0.0406	0.9765	0.0424	0.9872	0.0311	0.9885	0.0295
HH001	2	4	0.9955	0.0184	0.9845	0.0343	0.9646	0.0521	0.9713	0.0467	0.9818	0.0372	0.9826	0.0363	0.9720	0.0462	0.9886	0.0293	0.9890	0.0288
HH001	2	5	0.9974	0.0140	0.9871	0.0313	0.9856	0.0332	0.9878	0.0304	0.9881	0.0300	0.9882	0.0300	0.9876	0.0307	0.9924	0.0241	0.9913	0.0257
HH001	4	0	0.9970	0.0152	0.9997	0.0047	0.9997	0.0046	0.9998	0.0043	1.0000	0.0006	1.0000	0.0003	1.0000	0.0002	1.0000	0.0006	1.0000	0.0006
HH001	4	1	0.9950	0.0194	0.9836	0.0352	0.9805	0.0385	0.9818	0.0371	0.9825	0.0364	0.9817	0.0372	0.9821	0.0368	0.9844	0.0343	0.9844	0.0343
HH001	4	2	0.9964	0.0166	0.9772	0.0419	0.9845	0.0344	0.9808	0.0384	0.9815	0.0377	0.9815	0.0377	0.9810	0.0382	0.9906	0.0267	0.9906	0.0267
HH001	4	3	0.9955	0.0186	0.9855	0.0332	0.9791	0.0400	0.9831	0.0358	0.9855	0.0331	0.9842	0.0346	0.9835	0.0354	0.9906	0.0267	0.9906	0.0267
HH001	4	4	0.9945	0.0204	0.9770	0.0418	0.9747	0.0439	0.9744	0.0442	0.9792	0.0397	0.9793	0.0397	0.9746	0.0440	0.9870	0.0313	0.9896	0.0281
HH001	4	5	0.9951	0.0193	0.9592	0.0561	0.9742	0.0441	0.9649	0.0517	0.9689	0.0486	0.9675	0.0496	0.9652	0.0514	0.9825	0.0364	0.9886	0.0294
HH001	6	0	0.9953	0.0189	0.9996	0.0055	0.9996	0.0055	0.9997	0.0051	1.0000	0.0007	1.0000	0.0003	1.0000	0.0002	1.0000	0.0009	1.0000	0.0009
HH001	6	1	0.9840	0.0349	0.9458	0.0606	0.9666	0.0502	0.9807	0.0557	0.9706	0.0472	0.9716	0.0463	0.9602	0.0511	0.9602	0.0545	0.9746	0.0437
HH001	6	2	0.9913	0.0257	0.9515	0.0606	0.9666	0.0502	0.9587	0.0559	0.9624	0.0532	0.9615	0.0539	0.9588	0.0558	0.9734	0.0447	0.9830	0.0358
HH001	6	3	0.9932	0.0228	0.9589	0.0557	0.9682	0.0491	0.9635	0.0525	0.9646	0.0517	0.9651	0.0514	0.9638	0.0523	0.9805	0.0383	0.9831	0.0357
HH001	6	4	0.9920	0.0246	0.9526	0.0607	0.9722	0.0459	0.9640	0.0526	0.9669	0.0504	0.9594	0.0561	0.9641	0.0525	0.9805	0.0383	0.9807	0.0382
HH001	6	5	0.9905	0.0268	0.9609	0.0544	0.9559	0.0584	0.9593	0.0557	0.9706	0.0472	0.9716	0.0463	0.9602	0.0511	0.9781	0.0406	0.9853	0.0333
HH001	8	0	0.9939	0.0215	0.9994	0.0070	0.9993	0.0071	0.9995	0.0063	1.0000	0.0010	1.0000	0.0006	1.0000	0.0004	1.0000	0.0012	1.0000	0.0012
HH001	8	1	0.9883	0.0298	0.9349	0.0606	0.9437	0.0502	0.9864	0.0936	0.8951	0.0898	0.8697	0.1013	0.8869	0.0934	0.9492	0.0614	0.9644	0.0516
HH001	8	2	0.9851	0.0338	0.8954	0.0941	0.9199	0.0814	0.9066	0.0883	0.9130	0.0849	0.9141	0.0844	0.9069	0.0881	0.9690	0.0482	0.9710	0.0467
HH001	8	3	0.9829	0.0361	0.9165	0.0797	0.9187	0.0798	0.9236	0.0764	0.9466	0.0632	0.9415	0.0662	0.9245	0.0759	0.9596	0.0548	0.9749	0.0434
HH001	8	4	0.9871	0.0314	0.9227	0.0769	0.9430	0.0660	0.9289	0.0737	0.9400	0.0675	0.9377	0.0688	0.9299	0.0732	0.9614	0.0537	0.9708	0.0468
HH001	8	5	0.9871	0.0312	0.9343	0.0703	0.9361	0.0697	0.9345	0.0702	0.9425	0.0656	0.9359	0.0695	0.9348	0.0701	0.9593	0.0551	0.9633	0.0524
HH001	10	0	0.9921	0.0244	0.9992	0.0079	0.9992	0.0079	0.9994	0.0069	1.0000	0.0012	1.0000	0.0006	1.0000	0.0005	1.0000	0.0014	1.0000	0.0013
HH001	10	1	0.9809	0.0383	0.9413	0.0666	0.9429	0.0656	0.9478	0.0622	0.9528	0.0592	0.9506	0.0607	0.9483	0.0619	0.9532	0.0590	0.9618	0.0534
HH001	10	2	0.9758	0.0433	0.8639	0.1070	0.9010	0.0898	0.8824	0.0983	0.8922	0.0936	0.8984	0.0908	0.8831	0.0981	0.9429	0.0650	0.9572	0.0565
HH001	10	3	0.9853	0.0334	0.9064	0.0839	0.9037	0.0887	0.9030	0.0869	0.9123	0.0824	0.9143	0.0814	0.9040	0.0865	0.9574	0.0563	0.9627	0.0528

Fortsetzung der Tabelle

Zeit- reihe	$f_{\text{man, s.}}$ $f_{\text{man, g}}$	b_{man}	r_{NH}	$Q_{\text{C, RMSE}}^{\text{NH}}$	r_{LOCF}	$Q_{\text{C, RMSE}}^{\text{LOCF}}$	r_{NOCB}	$Q_{\text{C, RMSE}}^{\text{NOCB}}$	r_{NEOI}	$Q_{\text{C, RMSE}}^{\text{NEOI}}$	r_{lin}	$Q_{\text{C, RMSE}}^{\text{lin}}$	r_{pchip}	$Q_{\text{C, RMSE}}^{\text{pchip}}$	r_{DCTPL}	$Q_{\text{C, RMSE}}^{\text{DCTPL}}$	r_{AR6}	$Q_{\text{C, RMSE}}^{\text{AR6}}$	r_{AR672}	$Q_{\text{C, RMSE}}^{\text{AR672}}$
HH001	10	4	0.9871	0.0313	0.9246	0.0748	0.8936	0.0925	0.9095	0.0832	0.9194	0.0781	0.9185	0.0782	0.9099	0.0831	0.9596	0.0549	0.9703	0.0472
HH001	10	5	0.9864	0.0322	0.9167	0.0809	0.9027	0.0918	0.9085	0.0868	0.9163	0.0827	0.9174	0.0823	0.9092	0.0865	0.9679	0.0491	0.9685	0.0486
HH001	12	0	0.9914	0.0256	0.9990	0.0089	0.9989	0.0090	0.9992	0.0076	1.0000	0.0013	1.0000	0.0007	1.0000	0.0006	1.0000	0.0017	1.0000	0.0015
HH001	12	1	0.9820	0.0372	0.9515	0.0611	0.9339	0.0719	0.9314	0.0716	0.9485	0.0620	0.9443	0.0644	0.9323	0.0711	0.9621	0.0533	0.9629	0.0526
HH001	12	2	0.9814	0.0374	0.8762	0.0985	0.8206	0.1195	0.8457	0.1093	0.8900	0.0906	0.8877	0.0917	0.8473	0.1087	0.9297	0.0718	0.9473	0.0625
HH001	12	3	0.9830	0.0359	0.9368	0.0684	0.9120	0.0821	0.9226	0.0759	0.9307	0.0717	0.9317	0.0711	0.9236	0.0754	0.9407	0.0662	0.9694	0.0480
HH001	12	4	0.9816	0.0373	0.8745	0.0975	0.8487	0.1138	0.8592	0.1062	0.8719	0.1009	0.8715	0.1013	0.8603	0.1058	0.9453	0.0636	0.9625	0.0529
HH001	12	5	0.9771	0.0419	0.8252	0.1260	0.831	0.1064	0.8656	0.1064	0.8723	0.1034	0.8738	0.1027	0.8654	0.1063	0.9459	0.0633	0.9531	0.0591
HH001	14	0	0.9894	0.0284	0.9988	0.0097	0.9988	0.0097	0.9991	0.0081	1.0000	0.0015	1.0000	0.0008	1.0000	0.0006	1.0000	0.0018	1.0000	0.0016
HH001	14	1	0.9785	0.0404	0.9236	0.0748	0.8875	0.0944	0.8946	0.0888	0.9055	0.0840	0.9136	0.0797	0.8951	0.0886	0.9262	0.0736	0.9352	0.0691
HH001	14	2	0.9756	0.0435	0.8648	0.1074	0.9264	0.0749	0.8933	0.0905	0.9018	0.0887	0.9139	0.0822	0.8987	0.0903	0.9474	0.0625	0.9468	0.0628
HH001	14	3	0.9829	0.0361	0.9232	0.0777	0.9460	0.0634	0.9346	0.0705	0.9402	0.0674	0.9414	0.0664	0.9352	0.0702	0.9564	0.0570	0.9568	0.0567
HH001	14	4	0.9766	0.0421	0.9231	NaN	0.8677	0.0992	0.8978	0.0869	0.9091	0.0817	0.9101	0.0813	0.8985	0.0866	0.9335	0.0699	0.9502	0.0608
HH001	14	5	0.9789	0.0400	0.8933	NaN	0.9140	0.0806	0.9008	0.0868	0.9095	0.0828	0.9039	0.0858	0.9016	0.0865	0.9414	0.0658	0.9552	0.0577
HH001	16	0	0.9882	0.0300	0.9984	NaN	0.9983	0.0113	0.9989	0.0091	1.0000	0.0017	1.0000	0.0008	1.0000	0.0007	0.9999	0.0021	1.0000	0.0019
HH001	16	1	0.9746	0.0443	0.9082	NaN	0.7630	0.1469	0.8443	0.1110	0.8803	0.0958	0.9020	0.0848	0.8449	0.1108	0.9179	0.0775	0.9318	0.0708
HH001	16	2	0.9694	0.0485	0.7870	0.1423	0.8413	0.1110	0.7971	0.1298	0.8520	0.1072	0.8595	0.1046	0.7985	0.1294	0.9112	0.0804	0.9250	0.0741
HH001	16	3	0.9785	0.0403	0.8470	0.1098	0.8584	0.1049	0.8530	0.1072	0.8540	0.1068	0.8542	0.1068	0.8538	0.1070	0.9175	0.0776	0.9330	0.0702
HH001	16	4	0.9769	0.0419	0.7912	0.1379	0.8771	0.0959	0.8299	0.1183	0.8388	0.1144	0.8408	0.1136	0.8303	0.1182	0.9176	0.0776	0.9321	0.0707
HH001	16	5	0.9772	0.0417	0.8796	0.0940	0.8490	0.1064	0.8636	0.1005	0.8687	0.0984	0.8715	0.0973	0.8646	0.1002	0.9379	0.0677	0.9579	0.0560
HH001	18	0	0.9862	0.0324	0.9981	NaN	0.9980	NaN	0.9988	0.0096	1.0000	0.0018	1.0000	0.0009	1.0000	0.0007	0.9999	0.0025	0.9999	0.0022
HH001	18	1	0.9795	0.0394	0.9041	0.0869	0.9316	0.0721	0.9293	0.0725	0.9389	0.0674	0.9404	0.0664	0.9304	0.0719	0.9430	0.0649	0.9431	0.0649
HH001	18	2	0.9711	0.0468	0.7212	NaN	0.8799	0.0931	0.7746	0.1320	0.8013	0.1233	0.7549	0.1408	0.7755	0.1318	0.8953	0.0870	0.9127	0.0798
HH001	18	3	0.9723	0.0464	0.8149	0.1249	0.7725	0.1441	0.7943	0.1344	0.7964	0.1334	0.7924	0.1350	0.7949	0.1343	0.9336	0.0700	0.9393	0.0670
HH001	18	4	0.9612	0.0547	0.8281	0.1138	0.8223	0.1204	0.8245	0.1169	0.8504	0.1069	0.8433	0.1100	0.8256	0.1166	0.8947	0.0871	0.9223	0.0754
HH001	18	5	0.9672	0.0505	0.8495	0.1082	0.7887	0.1342	0.8167	0.1220	0.8736	0.0985	0.8788	0.0956	0.8195	0.1211	0.9198	0.0766	0.9408	0.0661
HH001	20	0	0.9841	0.0348	0.9980	0.0124	0.9980	0.0122	0.9987	0.0098	1.0000	0.0019	1.0000	0.0010	1.0000	0.0006	0.9999	0.0029	0.9999	0.0026
HH001	20	1	0.9653	0.0516	0.7972	0.1380	0.8119	0.1309	0.8040	0.1347	0.8049	0.1343	0.8055	0.1342	0.8049	0.1344	0.9184	0.0772	0.9181	0.0773
HH001	20	2	0.9700	0.0475	0.7660	0.1548	0.9098	0.0811	0.8246	0.1213	0.8548	0.1077	0.8478	0.1116	0.8255	0.1210	0.9138	0.0793	0.9256	0.0739
HH001	20	3	0.9608	0.0551	0.8171	0.1164	0.7844	0.1325	0.7978	0.1250	0.8073	0.1216	0.8062	0.1221	0.7991	0.1247	0.8883	0.0896	0.8989	0.0855
HH001	20	4	0.9693	0.0489	0.8795	0.0958	0.8976	0.0865	0.8899	0.0903	0.8937	0.0886	0.8934	0.0888	0.8910	0.0898	0.9240	0.0747	0.9410	0.0660
HH001	20	5	0.9622	0.0537	0.8402	0.1107	0.8268	0.1143	0.8267	0.1126	0.8562	0.1016	0.8502	0.1042	0.8281	0.1121	0.8861	0.0904	0.9221	0.0755
HH002	0	0	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH002	0	1	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH002	0	2	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH002	0	3	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH002	0	4	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH002	0	5	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH002	2	0	0.9975	0.0100	0.9998	0.0029	0.9998	0.0029	0.9998	0.0028	1.0000	0.0010	1.0000	0.0010	1.0000	0.0009	1.0000	0.0012	1.0000	0.0011

Fortsetzung der Tabelle

Zeit- reihe	$\mu_{\text{man},s}$, $\mu_{\text{man},g}$	b_{man}^n	r_{NH}	r_{NH} $Q_{\text{C},\text{RMSE}}$	r_{LOCF} $Q_{\text{C},\text{RMSE}}$	r_{NOCB} $Q_{\text{C},\text{RMSE}}$	r_{NEOI} $Q_{\text{C},\text{RMSE}}$	r_{lin} $Q_{\text{C},\text{RMSE}}$	r_{lin} $Q_{\text{C},\text{RMSE}}$	r_{pchip} $Q_{\text{C},\text{RMSE}}$	r_{pchip} $Q_{\text{C},\text{RMSE}}$	r_{DCTPL} $Q_{\text{C},\text{RMSE}}$	r_{DCTPLS} $Q_{\text{C},\text{RMSE}}$	r_{AR6} $Q_{\text{C},\text{RMSE}}$	r_{AR6} $Q_{\text{C},\text{RMSE}}$	r_{AR672} $Q_{\text{C},\text{RMSE}}$	r_{AR672} $Q_{\text{C},\text{RMSE}}$
HH002	2	1	0.9927	0.0173	0.9765	0.0313	0.9771	0.0309	0.9772	0.0308	0.9772	0.0308	0.9773	0.9858	0.0240	0.9879	0.0222
HH002	2	2	0.9939	0.0158	0.9812	0.0276	0.9829	0.0263	0.9840	0.0255	0.9842	0.0253	0.9832	0.9885	0.0216	0.9888	0.0213
HH002	2	3	0.9945	0.0150	0.9912	0.0189	0.9910	0.0192	0.9935	0.0162	0.9936	0.0161	0.9917	0.9950	0.0143	0.9948	0.0146
HH002	2	4	0.9945	0.0149	0.9837	0.0258	0.9856	0.0242	0.9876	0.0224	0.9882	0.0219	0.9860	0.9901	0.0201	0.9901	0.0200
HH002	2	5	0.9952	0.0139	0.9848	0.0249	0.9870	0.0230	0.9880	0.0221	0.9884	0.0217	0.9873	0.9909	0.0192	0.9911	0.0190
HH002	4	0	0.9949	0.0144	0.9996	0.0040	0.9997	0.0037	0.9999	0.0015	1.0000	0.0014	1.0000	0.9999	0.0017	0.9999	0.0017
HH002	4	1	0.9908	0.0195	0.9864	0.0236	0.9727	0.0337	0.9809	0.0281	0.9727	0.0341	0.9733	0.9895	0.0206	0.9895	0.0206
HH002	4	2	0.9898	0.0204	0.9621	0.0397	0.9712	0.0343	0.9748	0.0321	0.9751	0.0319	0.9715	0.9814	0.0274	0.9813	0.0275
HH002	4	3	0.9911	0.0190	0.9805	0.0282	0.9796	0.0288	0.9839	0.0256	0.9826	0.0266	0.9806	0.9876	0.0225	0.9873	0.0227
HH002	4	4	0.9846	0.0251	0.9578	0.0416	0.9579	0.0415	0.9619	0.0394	0.9605	0.0401	0.9582	0.9776	0.0301	0.9781	0.0297
HH002	4	5	0.9884	0.0218	0.9757	0.0314	0.9737	0.0329	0.9819	0.0271	0.9807	0.0280	0.9773	0.9856	0.0242	0.9871	0.0229
HH002	6	0	0.9916	0.0185	0.9994	0.0051	0.9995	0.0047	0.9999	0.0019	0.9999	0.0017	0.9999	0.9999	0.0021	0.9999	0.0020
HH002	6	1	0.9874	0.0226	0.9838	0.0256	0.9841	0.0255	0.9845	0.0251	0.9844	0.0252	0.9845	0.9852	0.0244	0.9851	0.0246
HH002	6	2	0.9847	0.0250	0.9109	0.0619	0.9238	0.0566	0.9571	0.0416	0.9527	0.0440	0.9254	0.9713	0.0340	0.9710	0.0342
HH002	6	3	0.9866	0.0233	0.9695	0.0353	0.9486	0.0461	0.9636	0.0385	0.9572	0.0418	0.9580	0.9783	0.0296	0.9782	0.0297
HH002	6	4	0.9848	0.0249	0.9632	0.0388	0.9679	0.0360	0.9674	0.0364	0.9675	0.0364	0.9662	0.9746	0.0320	0.9743	0.0322
HH002	6	5	0.9813	0.0278	0.9455	0.0480	0.9542	0.0434	0.9607	0.0400	0.9569	0.0421	0.9542	0.9750	0.0318	0.9768	0.0306
HH002	8	0	0.9898	0.0204	0.9990	0.0063	0.9991	0.0062	0.9999	0.0023	0.9999	0.0021	0.9999	0.9998	0.0026	0.9998	0.0025
HH002	8	1	0.9806	0.0281	0.9315	0.0541	0.8604	0.0827	0.9354	0.0516	0.9488	0.0457	0.8897	0.9668	0.0366	0.9657	0.0372
HH002	8	2	0.9778	0.0301	0.9308	0.0534	0.9083	0.0643	0.9379	0.0505	0.9383	0.0505	0.9215	0.9597	0.0401	0.9598	0.0401
HH002	8	3	0.9813	0.0276	0.9593	0.0406	0.9570	0.0421	0.9609	0.0400	0.9616	0.0396	0.9591	0.9789	0.0292	0.9790	0.0291
HH002	8	4	0.9814	0.0276	0.9627	0.0390	0.9578	0.0416	0.9642	0.0380	0.9571	0.0416	0.9599	0.9762	0.0310	0.9764	0.0309
HH002	8	5	0.9796	0.0288	0.9435	0.0480	0.9598	0.0403	0.9539	0.0433	0.9497	0.0453	0.9536	0.9670	0.0364	0.9676	0.0360
HH002	10	0	0.9869	0.0231	0.9989	0.0068	0.9989	0.0068	0.9998	0.0025	0.9999	0.0023	0.9999	0.9998	0.0028	0.9998	0.0027
HH002	10	1	0.9772	0.0305	0.8930	0.0695	0.9498	0.0452	0.9404	0.0492	0.9353	0.0514	0.9204	0.9604	0.0398	0.9591	0.0405
HH002	10	2	0.9752	0.0318	0.9290	NaN	0.9674	0.0362	0.9541	0.0434	0.9516	0.0448	0.9480	0.9741	0.0323	0.9740	0.0324
HH002	10	3	0.9752	0.0317	0.9484	0.0455	0.9046	0.0619	0.9465	0.0462	0.9496	0.0448	0.9335	0.9605	0.0398	0.9610	0.0395
HH002	10	4	0.9746	0.0322	0.9514	0.0442	0.9265	0.0557	0.9512	0.0442	0.9485	0.0455	0.9395	0.9656	0.0371	0.9652	0.0374
HH002	10	5	0.9714	0.0342	0.8967	0.0645	0.9029	0.0627	0.9094	0.0602	0.9016	0.0632	0.9003	0.9492	0.0450	0.9511	0.0441
HH002	12	0	0.9841	0.0254	0.9985	0.0077	0.9989	0.0067	0.9998	0.0028	0.9998	0.0026	0.9998	0.9997	0.0032	0.9998	0.0031
HH002	12	1	0.9746	0.0323	0.9647	0.0379	0.9281	0.0591	0.9565	0.0426	0.9441	0.0499	0.9427	0.9605	0.0399	0.9692	0.0352
HH002	12	2	0.9616	0.0397	0.7991	0.0949	0.8619	0.0751	0.8962	0.0644	0.9004	0.0630	0.8434	0.9321	0.0518	0.9331	0.0514
HH002	12	3	0.9704	0.0346	0.8967	0.0660	0.9018	0.0637	0.9122	0.0601	0.9105	0.0609	0.8976	0.9313	0.0521	0.9403	0.0486
HH002	12	4	0.9656	0.0378	0.8533	0.0808	0.8581	0.0760	0.8798	0.0699	0.8822	0.0691	0.8557	0.9367	0.0500	0.9374	0.0497
HH002	12	5	0.9649	0.0377	0.8931	NaN	0.8829	0.0696	0.9088	0.0601	0.9020	0.0624	0.8890	0.9279	0.0533	0.9303	0.0524
HH002	14	0	0.9821	0.0270	0.9981	0.0088	0.9981	0.0088	0.9998	0.0029	0.9998	0.0028	0.9998	0.9997	0.0037	0.9997	0.0035
HH002	14	1	0.9623	0.0395	0.7766	0.1061	0.9359	0.0511	0.8645	0.0748	0.8289	0.0901	0.8654	0.9501	0.0447	0.9524	0.0436
HH002	14	2	0.9653	0.0375	0.9306	NaN	0.9333	0.0514	0.9375	0.0497	0.9380	0.0496	0.9342	0.9411	0.0483	0.9415	0.0481
HH002	14	3	0.9621	0.0393	0.8196	0.0851	0.8318	0.0896	0.8687	0.0734	0.8531	0.0765	0.8203	0.9324	0.0517	0.9367	0.0501

Fortsetzung der Tabelle

Zeitreihe	$f_{\text{man},s} / f_{\text{man},g}$	b_{man}	r_{NH}	$Q_{\text{C,RMSE}}^{\text{NH}}$	r_{LOCF}	$Q_{\text{C,RMSE}}^{\text{LOCF}}$	r_{NOCB}	$Q_{\text{C,RMSE}}^{\text{NOCB}}$	r_{NEOI}	$Q_{\text{C,RMSE}}^{\text{NEOI}}$	r_{lin}	$Q_{\text{C,RMSE}}^{\text{lin}}$	r_{pchip}	$Q_{\text{C,RMSE}}^{\text{pchip}}$	r_{DCTPL}	$Q_{\text{C,RMSE}}^{\text{DCTPL}}$	r_{AR6}	$Q_{\text{C,RMSE}}^{\text{AR6}}$	r_{AR672}	$Q_{\text{C,RMSE}}^{\text{AR672}}$
HH002	14	4	0.9678	0.0363	0.9137	0.0603	0.9388	0.0495	0.9255	0.0548	0.9337	0.0515	0.9322	0.0521	0.9263	0.0545	0.9561	0.0419	0.9599	0.0400
HH002	14	5	0.9638	0.0382	0.8580	NaN	0.7520	0.1028	0.8011	0.0901	0.8170	0.0857	0.7932	0.0921	0.8028	0.0897	0.9192	0.0563	0.9209	0.0557
HH002	16	0	0.9791	0.0292	0.9978	NaN	0.9978	0.0094	0.9985	0.0079	0.9997	0.0034	0.9998	0.0031	0.9998	0.0030	0.9996	0.0040	0.9997	0.0037
HH002	16	1	0.9558	0.0425	0.7107	NaN	0.8478	0.0827	0.7456	0.1051	0.8274	0.0826	0.8048	0.0886	0.7462	0.1049	0.9062	0.0604	0.9075	0.0600
HH002	16	2	0.9666	0.0369	0.9101	0.0603	0.8852	0.0752	0.8948	0.0681	0.9066	0.0639	0.9119	0.0618	0.8962	0.0677	0.9498	0.0447	0.9504	0.0444
HH002	16	3	0.9655	0.0374	0.9008	0.0647	0.8459	0.0860	0.8702	0.0755	0.8878	0.0692	0.8938	0.0671	0.8707	0.0754	0.9507	0.0443	0.9519	0.0438
HH002	16	4	0.9573	0.0420	0.8066	0.0938	0.8450	0.0832	0.8276	0.0881	0.8304	0.0872	0.8253	0.0890	0.8284	0.0880	0.9307	0.0523	0.9339	0.0511
HH002	16	5	0.9629	0.0388	0.8914	NaN	0.9036	NaN	0.8971	0.0649	0.9139	0.0591	0.9114	0.0600	0.9046	0.0646	0.9342	0.0510	0.9402	0.0487
HH002	18	0	0.9766	0.0309	0.9974	0.0102	0.9973	0.0104	0.9982	0.0087	0.9997	0.0035	0.9997	0.0033	0.9980	0.0032	0.9996	0.0041	0.9996	0.0039
HH002	18	1	0.9609	0.0399	0.9147	0.0604	0.9481	0.0460	0.9298	0.0530	0.9446	0.0472	0.9488	0.0451	0.9314	0.0525	0.9540	0.0428	0.9539	0.0429
HH002	18	2	0.9516	0.0447	0.8099	0.0967	0.8787	0.0698	0.8277	0.0870	0.8473	0.0815	0.8369	0.0846	0.8291	0.0867	0.9135	0.0583	0.9213	0.0555
HH002	18	3	0.9565	0.0420	0.7974	0.0892	0.8744	0.0704	0.8355	0.0802	0.8540	0.0751	0.8490	0.0764	0.8366	0.0799	0.8948	0.0638	0.8969	0.0632
HH002	18	4	0.9547	0.0433	0.7270	NaN	0.8985	NaN	0.8054	0.0924	0.8473	0.0799	0.8182	0.0911	0.8073	0.0919	0.9200	0.0560	0.9208	0.0557
HH002	18	5	0.9548	0.0426	0.8694	0.0720	0.8377	0.0845	0.8497	0.0791	0.8658	0.0741	0.8696	0.0730	0.8509	0.0788	0.9051	0.0608	0.9059	0.0605
HH002	20	0	0.9735	0.0329	0.9969	0.0112	0.9970	0.0110	0.9980	0.0090	0.9996	0.0039	0.9997	0.0036	0.9997	0.0035	0.9995	0.0047	0.9995	0.0043
HH002	20	1	0.9417	0.0497	0.8947	0.0666	0.9300	NaN	0.9196	0.0571	0.9244	0.0555	0.9243	0.0555	0.9213	0.0565	0.9308	0.0526	0.9354	0.0506
HH002	20	2	0.9565	0.0505	0.7431	0.1041	0.8712	0.0721	0.8155	0.0854	0.8575	0.0741	0.8479	0.0767	0.8170	0.0850	0.8838	0.0671	0.8910	0.0649
HH002	20	3	0.9557	0.0425	0.8450	0.0817	0.8380	0.0813	0.8434	0.0809	0.8465	0.0799	0.8468	0.0799	0.8445	0.0806	0.9355	0.0505	0.9350	0.0507
HH002	20	4	0.9514	0.0443	0.8491	0.0782	0.7543	0.1007	0.8035	0.0897	0.8445	0.0783	0.8443	0.0783	0.8042	0.0895	0.8969	0.0633	0.9043	0.0610
HH002	20	5	0.9411	0.0493	0.8247	NaN	0.8450	0.0780	0.8291	0.0815	0.8497	0.0763	0.8445	0.0778	0.8304	0.0812	0.8892	0.0654	0.8915	0.0647
HH003	0	0	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH003	0	1	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH003	0	2	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH003	0	3	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH003	0	4	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH003	0	5	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
HH003	2	0	0.9938	0.0103	0.9996	0.0024	0.9996	0.0026	0.9996	0.0026	0.9999	0.0013	0.9999	0.0012	0.9999	0.0012	0.9999	0.0014	0.9999	0.0013
HH003	2	1	0.9838	0.0167	0.9904	0.0127	0.9902	0.0129	0.9903	0.0128	0.9906	0.0126	0.9905	0.0127	0.9906	0.0126	0.9893	0.0134	0.9907	0.0126
HH003	2	2	0.9882	0.0142	0.9883	0.0141	0.9933	0.0106	0.9902	0.0129	0.9912	0.0122	0.9913	0.0121	0.9902	0.0129	0.9940	0.0101	0.9941	0.0100
HH003	2	3	0.9853	0.0159	0.9784	0.0193	0.9748	0.0208	0.9767	0.0200	0.9778	0.0195	0.9786	0.0192	0.9771	0.0198	0.9895	0.0133	0.9904	0.0128
HH003	2	4	0.9884	0.0141	0.9914	NaN	0.9816	0.0177	0.9865	0.0151	0.9896	0.0133	0.9874	0.0146	0.9870	0.0149	0.9940	0.0101	0.9941	0.0100
HH003	2	5	0.9874	0.0146	0.9913	0.0121	0.9886	0.0139	0.9900	0.0130	0.9913	0.0121	0.9914	0.0121	0.9900	0.0130	0.9927	0.0111	0.9928	0.0110
HH003	4	0	0.9882	0.0142	0.9992	0.0037	0.9992	0.0037	0.9992	0.0036	0.9998	0.0019	0.9998	0.0018	0.9998	0.0018	0.9998	0.0021	0.9998	0.0020
HH003	4	1	0.9765	0.0200	0.9861	0.0154	0.9866	NaN	0.9867	0.0150	0.9884	0.0140	0.9879	0.0143	0.9873	0.0147	0.9883	0.0141	0.9886	0.0139
HH003	4	2	0.9727	0.0216	0.9852	0.0158	0.9673	0.0239	0.9767	0.0200	0.9795	0.0187	0.9743	0.0210	0.9771	0.0198	0.9853	0.0156	0.9860	0.0154
HH003	4	3	0.9752	0.0206	0.9874	0.0146	0.9806	0.0182	0.9844	0.0163	0.9872	0.0147	0.9866	0.0151	0.9851	0.0159	0.9872	0.0147	0.9891	0.0136
HH003	4	4	0.9751	0.0206	0.9685	0.0233	0.9701	0.0225	0.9694	0.0228	0.9738	0.0211	0.9739	0.0210	0.9704	0.0224	0.9801	0.0183	0.9808	0.0180
HH003	4	5	0.9743	0.0209	0.9756	0.0203	0.9520	0.0291	0.9628	0.0253	0.9680	0.0234	0.9668	0.0238	0.9626	0.0254	0.9824	0.0172	0.9834	0.0167
HH003	6	0	0.9813	0.0178	0.9987	0.0048	0.9987	0.0047	0.9988	0.0044	0.9997	0.0023	0.9997	0.0023	0.9997	0.0022	0.9996	0.0025	0.9997	0.0024

Fortsetzung der Tabelle

Zeit- reihe	$\mu_{\text{man},s}$, $\mu_{\text{man},g}$	b_{man}	r_{NH}	r_{NH} $Q_C, RMSE$	r_{LOC}	$Q_C, RMSE$	r_{NOCB}	$Q_C, RMSE$	r_{NEOI}	$Q_C, RMSE$	r_{lin}	$Q_C, RMSE$	r_{pchip}	$Q_C, RMSE$	r_{DCTL}	$Q_C, RMSE$	r_{AR6}	$Q_C, RMSE$	r_{AR672}	$Q_C, RMSE$
HH003	6	1	0.9639	0.0248	0.9829	0.0170	0.9722	0.0223	0.9773	0.0198	0.9793	0.0189	0.9763	0.0204	0.9780	0.0195	0.9841	0.0164	0.9839	0.0165
HH003	6	2	0.9654	0.0243	0.9588	0.0273	0.9797	0.0185	0.9712	0.0222	0.9767	0.0200	0.9737	0.0204	0.9723	0.0218	0.9816	0.0176	0.9822	0.0173
HH003	6	3	0.9639	0.0247	0.9561	0.0272	0.9379	0.0325	0.9474	0.0298	0.9640	0.0246	0.9643	0.0245	0.9486	0.0295	0.9710	0.0220	0.9711	0.0220
HH003	6	4	0.9619	0.0255	0.9452	0.0311	0.9675	0.0234	0.9516	0.0288	0.9614	0.0256	0.9603	0.0260	0.9524	0.0286	0.9757	0.0202	0.9753	0.0204
HH003	6	5	0.9620	0.0254	0.9375	0.0326	0.9542	0.0280	0.9436	0.0310	0.9625	0.0251	0.9594	0.0261	0.9449	0.0306	0.9734	0.0211	0.9754	0.0203
HH003	8	0	0.9735	0.0212	0.9982	0.0055	0.9982	NaN	0.9984	0.0051	0.9996	0.0028	0.9996	0.0027	0.9996	0.0026	0.9995	0.0029	0.9996	0.0028
HH003	8	1	0.9497	0.0294	0.9657	0.0240	0.9616	NaN	0.9644	0.0245	0.9654	0.0241	0.9654	0.0241	0.9656	0.0241	0.9668	0.0236	0.9681	0.0231
HH003	8	2	0.9521	0.0283	0.9456	0.0303	0.9472	0.0298	0.9467	0.0300	0.9476	0.0297	0.9476	0.0297	0.9477	0.0297	0.9526	0.0281	0.9543	0.0276
HH003	8	3	0.9523	0.0284	0.9152	0.0377	0.9526	NaN	0.9340	0.0333	0.9436	0.0307	0.9352	0.0329	0.9353	0.0329	0.9583	0.0264	0.9600	0.0258
HH003	8	4	0.9513	0.0288	0.9610	0.0256	0.9026	0.0433	0.9306	0.0352	0.9380	0.0331	0.9485	0.0298	0.9316	0.0350	0.9629	0.0249	0.9681	0.0231
HH003	8	5	0.9440	0.0310	0.8901	0.0452	0.9107	0.0403	0.9029	0.0423	0.9075	0.0411	0.9078	0.0409	0.9034	0.0422	0.9548	0.0274	0.9577	0.0265
HH003	10	0	0.9677	0.0234	0.9978	0.0061	0.9978	0.0061	0.9981	0.0057	0.9994	0.0031	0.9995	0.0030	0.9995	0.0033	0.9994	0.0033	0.9994	0.0032
HH003	10	1	0.9389	0.0323	0.9666	0.0237	0.9311	0.0363	0.9500	0.0299	0.9558	0.0280	0.9563	0.0279	0.9511	0.0295	0.9700	0.0224	0.9708	0.0221
HH003	10	2	0.9388	0.0321	0.9214	NaN	0.9350	0.0341	0.9274	0.0355	0.9415	0.0316	0.9312	0.0347	0.9285	0.0353	0.9614	0.0254	0.9613	0.0254
HH003	10	3	0.9381	0.0327	0.9513	0.0291	0.9271	0.0358	0.9395	0.0320	0.9554	0.0273	0.9527	0.0282	0.9407	0.0317	0.9631	0.0248	0.9650	0.0242
HH003	10	4	0.9353	0.0330	0.9438	0.0306	0.8942	0.0423	0.9154	0.0375	0.9284	0.0344	0.9252	0.0351	0.9162	0.0373	0.9516	0.0283	0.9532	0.0279
HH003	10	5	0.9341	0.0336	0.9502	0.0288	0.9010	0.0411	0.9270	0.0350	0.9418	0.0312	0.9264	0.0352	0.9289	0.0345	0.9619	0.0252	0.9644	0.0244
HH003	12	0	0.9616	0.0256	0.9972	0.0069	0.9972	0.0069	0.9977	0.0063	0.9993	0.0035	0.9994	0.0033	0.9994	0.0033	0.9992	0.0038	0.9992	0.0036
HH003	12	1	0.9255	0.0356	0.8378	0.0533	0.8024	0.0597	0.8212	0.0564	0.8225	0.0562	0.8227	0.0561	0.8227	0.0562	0.9232	0.0357	0.9290	0.0342
HH003	12	2	0.9215	0.0364	0.8945	0.0419	0.9274	0.0345	0.9074	0.0390	0.9156	0.0373	0.9147	0.0375	0.9090	0.0387	0.9244	0.0353	0.9285	0.0343
HH003	12	3	0.9221	0.0363	0.8872	0.0436	0.8422	0.0542	0.8595	0.0498	0.9069	0.0395	0.8817	0.0460	0.8615	0.0494	0.9376	0.0321	0.9416	0.0310
HH003	12	4	0.9298	0.0347	0.9254	0.0360	0.8796	0.0459	0.9034	0.0410	0.9086	0.0398	0.9077	0.0401	0.9047	0.0407	0.9455	0.0302	0.9527	0.0280
HH003	12	5	0.9222	0.0367	0.9304	0.0339	0.9057	NaN	0.9153	0.0377	0.9234	0.0357	0.9285	0.0344	0.9169	0.0373	0.9463	0.0298	0.9477	0.0294
HH003	14	0	0.9558	0.0274	0.9964	0.0078	0.9964	0.0078	0.9972	0.0069	0.9991	0.0040	0.9992	0.0038	0.9992	0.0037	0.9989	0.0043	0.9990	0.0040
HH003	14	1	0.9076	0.0395	0.8065	0.0621	0.9415	0.0311	0.8462	0.0522	0.8796	0.0458	0.8903	0.0437	0.8479	0.0519	0.9461	0.0298	0.9464	0.0298
HH003	14	2	0.9134	0.0385	0.8975	0.0424	0.8964	0.0430	0.9000	0.0419	0.9053	0.0406	0.8971	0.0427	0.9016	0.0415	0.9290	0.0343	0.9453	0.0301
HH003	14	3	0.9219	0.0364	0.9156	0.0373	0.9186	0.0365	0.9194	0.0364	0.9209	0.0360	0.9208	0.0360	0.9212	0.0360	0.9406	0.0313	0.9414	0.0311
HH003	14	4	0.9118	0.0384	0.8818	0.0458	0.8963	NaN	0.8848	0.0438	0.9057	0.0394	0.9019	0.0401	0.8862	0.0435	0.9177	0.0367	0.9246	0.0351
HH003	14	5	0.9105	0.0387	0.8614	0.0495	0.8914	0.0431	0.8793	0.0456	0.8876	0.0438	0.8873	0.0439	0.8810	0.0453	0.9234	0.0354	0.9258	0.0348
HH003	16	0	0.9476	0.0298	0.9958	0.0084	0.9959	0.0084	0.9968	0.0074	0.9989	0.0043	0.9990	0.0040	0.9991	0.0040	0.9987	0.0046	0.9989	0.0043
HH003	16	1	0.9050	0.0399	0.9101	0.0392	0.9255	0.0351	0.9107	0.0383	0.9216	0.0359	0.9166	0.0372	0.9126	0.0379	0.9299	0.0340	0.9340	0.0329
HH003	16	2	0.8967	0.0420	0.9358	0.0326	0.8742	NaN	0.9036	0.0406	0.9223	0.0362	0.9167	0.0382	0.9058	0.0401	0.9236	0.0356	0.9424	0.0308
HH003	16	3	0.9096	0.0392	0.8578	NaN	0.8568	0.0524	0.8563	0.0501	0.8966	0.0418	0.8876	0.0437	0.8587	0.0497	0.9317	0.0336	0.9345	0.0328
HH003	16	4	0.9110	0.0387	0.8715	0.0482	0.8989	0.0416	0.8805	0.0456	0.8937	0.0427	0.8943	0.0426	0.8823	0.0452	0.9278	0.0345	0.9303	0.0338
HH003	16	5	0.9035	0.0407	0.8726	0.0476	0.8328	0.0566	0.8390	0.0537	0.8850	0.0443	0.8556	0.0511	0.8429	0.0531	0.9341	0.0329	0.9370	0.0322
HH003	18	0	0.9415	0.0315	0.9948	0.0094	0.9950	0.0093	0.9963	0.0080	0.9988	0.0046	0.9989	0.0044	0.9989	0.0043	0.9985	0.0051	0.9987	0.0047
HH003	18	1	0.8962	0.0435	0.8921	0.0483	0.9387	0.0321	0.9093	0.0402	0.9298	0.0350	0.9143	0.0407	0.9115	0.0397	0.9448	0.0303	0.9463	0.0298
HH003	18	2	0.8606	0.0499	0.6514	NaN	0.7093	NaN	0.6812	0.0766	0.6881	0.0757	0.6712	0.0782	0.6836	0.0764	0.8863	0.0428	0.8893	0.0422
HH003	18	3	0.8891	0.0440	0.8883	0.0438	0.8222	0.0598	0.8534	0.0521	0.8617	0.0505	0.8601	0.0508	0.8557	0.0518	0.9204	0.0364	0.9293	0.0341

Fortsetzung der Tabelle

Zeitreihe	$n_{\text{man},s} / n_{\text{man},g}$	$b_{\text{man}}^{\text{NH}}$	r^{NH}	$Q_{C, \text{RMSE}}^{\text{NH}}$	r^{LOCF}	$Q_{C, \text{RMSE}}^{\text{LOCF}}$	r^{NOCB}	$Q_{C, \text{RMSE}}^{\text{NOCB}}$	r^{NEOI}	$Q_{C, \text{RMSE}}^{\text{NEOI}}$	r^{lin}	$Q_{C, \text{RMSE}}^{\text{lin}}$	r^{pchip}	$Q_{C, \text{RMSE}}^{\text{pchip}}$	r^{DCTPL}	$Q_{C, \text{RMSE}}^{\text{DCTPLS}}$	r^{AR6}	$Q_{C, \text{RMSE}}^{\text{AR6}}$	r^{AR672}	$Q_{C, \text{RMSE}}^{\text{AR672}}$
HH003	18	4	0.8879	0.0433	0.8595	0.0491	0.8095	0.0569	0.8382	0.0526	0.8549	0.0493	0.8425	0.0513	0.8402	0.0522	0.8769	0.0445	0.8991	0.0404
HH003	18	5	0.8861	0.0437	0.7361	0.0725	0.8171	0.0594	0.7642	0.0668	0.8139	0.0573	0.7914	0.0618	0.7660	0.0665	0.9085	0.0385	0.9149	0.0372
HH003	20	0	0.9355	0.0331	0.9942	NaN	0.9941	0.0100	0.9938	0.0085	0.9986	0.0049	0.9987	0.0047	0.9987	0.0046	0.9983	0.0055	0.9985	0.0051
HH003	20	1	0.8837	0.0441	0.7616	NaN	0.6947	0.0783	0.7215	0.0730	0.7268	0.0722	0.7293	0.0719	0.7229	0.0728	0.8849	0.0430	0.8854	0.0429
HH003	20	2	0.8810	0.0454	0.7899	NaN	0.9290	NaN	0.8491	0.0515	0.8857	0.0445	0.8405	0.0544	0.8516	0.0511	0.9283	0.0343	0.9338	0.0330
HH003	20	3	0.8904	0.0432	0.9227	0.0355	0.8415	0.0518	0.8830	0.0439	0.8957	0.0415	0.9096	0.0384	0.8856	0.0434	0.9308	0.0337	0.9310	0.0337
HH003	20	4	0.8859	0.0441	0.8416	NaN	0.8935	0.0420	0.8645	0.0482	0.8824	0.0446	0.8684	0.0478	0.8668	0.0478	0.9060	0.0390	0.9115	0.0379
HH003	20	5	0.8820	0.0449	0.8564	NaN	0.8820	0.0436	0.8713	0.0455	0.8787	0.0442	0.8753	0.0448	0.8739	0.0451	0.8966	0.0409	0.9050	0.0392

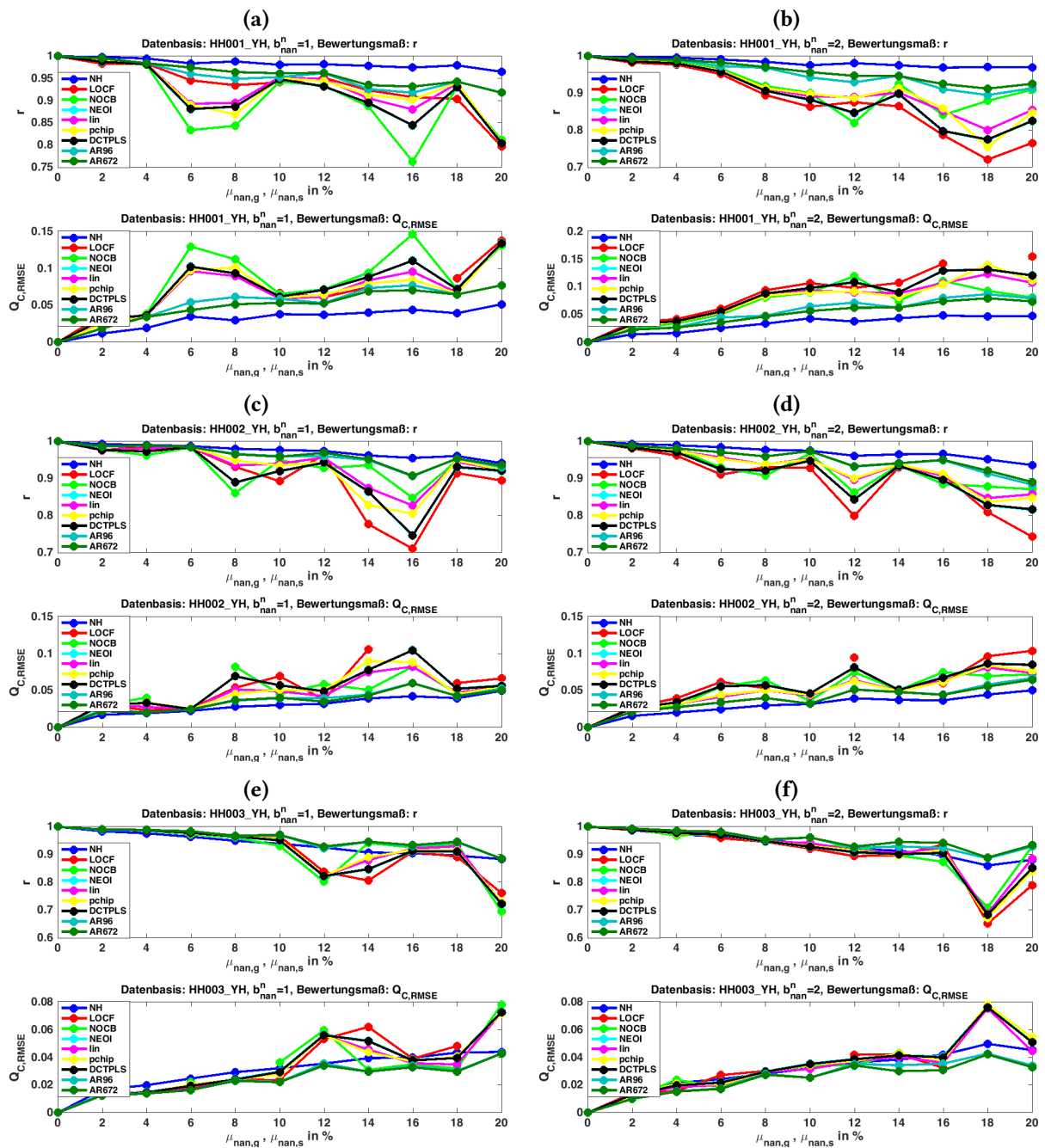


Abbildung B.3: Vergleich verschiedener Imputationsverfahren bei fehlerbehafteten Lastzeitreihen für $b_{nan}^1 = 1$ und $b_{nan}^2 = 2$ (Werte aus Tabelle B.8, Ergebnisse für $b_{nan}^1 = 0, 3, 4, 5$ in Abb. 2.6 und B.4, Beschreibung der Imputationsverfahren in Tabelle 1.3); Abkürzungen: NH - Nearest Historical Average, LOCF - Last-Observation-Carried-Forward, NOCB - Next-Observation-Carried-Backward, NEOI - nächstgelegene Nicht-NaN-Wert, lin - lineare Interpolation zwischen benachbarten Nicht-NaN-Werten, pchip - Piecewise Cubic Hermite Interpolating Polynomial, DCTPLS - Penalized-Least-Square-Regression auf Basis einer diskreten Kosinustransformation, {AR96, AR672} - AR-Modell mit unterschiedlichen Ordnungen.

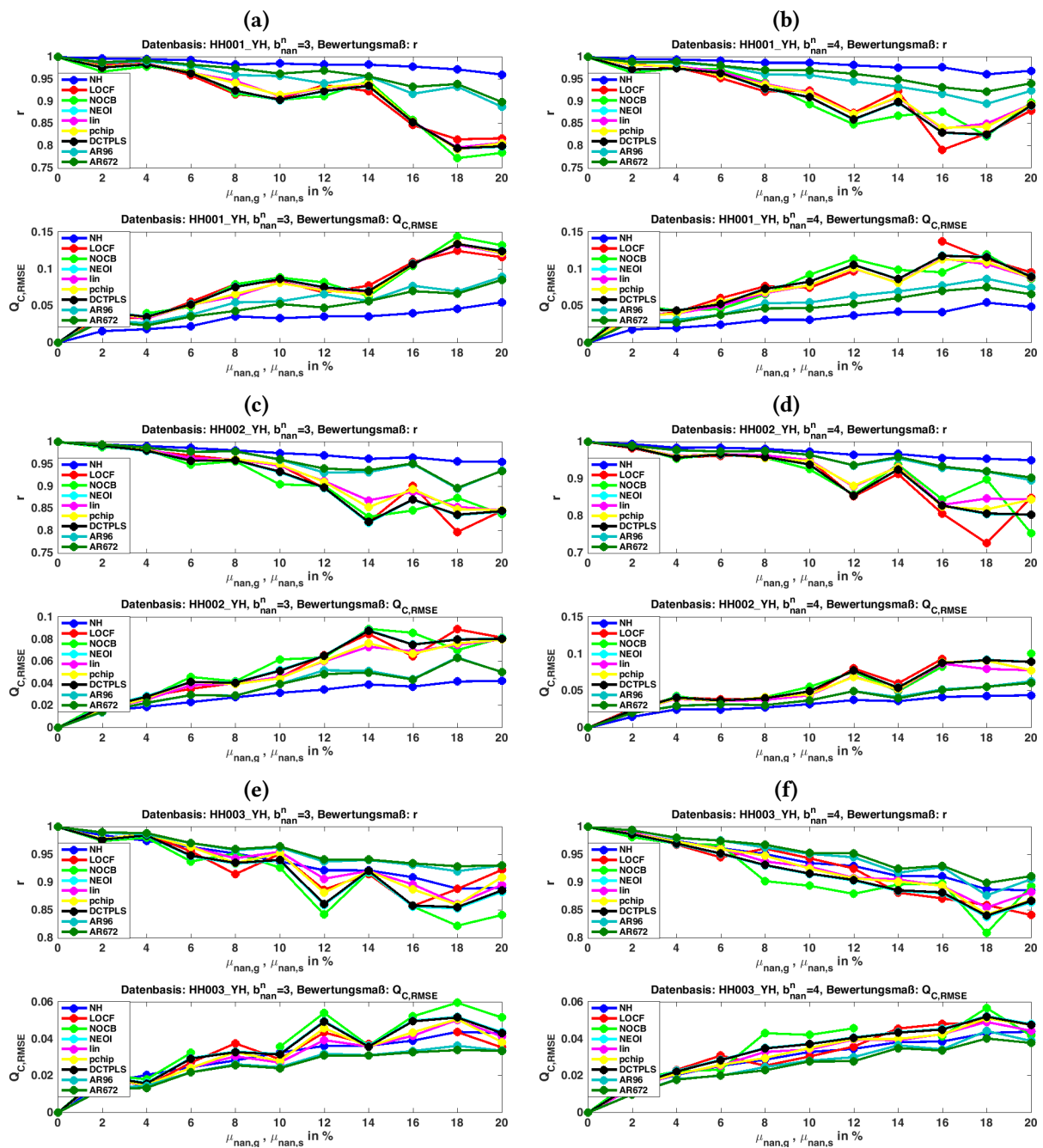


Abbildung B.4: Vergleich verschiedener Imputationsverfahren bei fehlerbehafteten Lastzeitreihen für $b_{nan}^n = 3$ und $b_{nan}^n = 4$ (Werte aus Tabelle B.8, Ergebnisse für $b_{nan}^n = 0, 1, 2, 5$ in Abb. 2.6 und B.3, Beschreibung der Imputationsverfahren in Tabelle 1.3); Abkürzungen: NH - Nearest Historical Average, LOCF - Last-Observation-Carried-Forward, NOCB - Next-Observation-Carried-Backward, NEOI - nächstgelegene Nicht-NaN-Wert, lin - lineare Interpolation zwischen benachbarten Nicht-NaN-Werten, pchip - Piecewise Cubic Hermite Interpolating Polynomial, DCTPLS - Penalized-Least-Square-Regression auf Basis einer diskreten Kosinustransformation, {AR96,AR672} - AR-Modell mit unterschiedlichen Ordnungen.

Weitere Ergebnisse der vergleichenden Bewertung verschiedener Verfahren zur Ausreißerbehandlung bei fehlerbehafteten Lastzeitreihen

Die Struktur des zusätzlichen Testdatensatzes ist in Tabelle B.9 dargestellt. Des Weiteren sind in Tabelle B.9 die Ergebnisse des Testverfahrens (siehe Abschnitt 2.4.1.2) zur Wahl der Hampel-Filterparameter h und $C_{h,\alpha}^{Ham}$ dargestellt. Als Bewertungsmaß dienen die Wurzel des mittleren quadratischen Fehlers $Q_{C, RMSE}$ und der Korrelationskoeffizient r zwischen der originalen (d.h., ohne Ausreißer) und der gefilterten Jahreslastzeitreihe. Die Berechnungsvorschriften für $Q_{C, RMSE}$ und r stehen in Tabelle 2.11.

Tabelle B.9: Das Ergebnis der Ermittlung von optimalen Hampel-Filterparametern h und $C_{h,\alpha}^{Ham}$ bei ausreißer-behafteten Lastzeitreihen (siehe Beschreibung des Testverfahrens in Abschnitt 2.4.1.2) ist hier als mehrseitige Übersichtstabelle dargestellt. Aus Platzgründen wurden die konstanten Parameter $\mu_E = 3.2$, $\sigma_E^2 = 0$, $b_{nan}^n = 0$, $\mu_{nan,g} = 0$, $\sigma_{nan,g}^2 = 0$, $\mu_{nan,s} = 0$, $\sigma_{nan,s}^2 = 0$, $\sigma_a^2 = 0$, $a = 1.5$, $b = 5$, $\beta_T = 5\%$, $\beta_P = 5\%$ und $\beta_E = 0\%$ nicht in die Tabelle aufgenommen. Die Parametereinstellungen des Benchmarkgenerators zur Erzeugung der Lastzeitreihen stehen in Tabelle B.6.

Zeit- reihe	K	T _S [h]	β _M [%]	β _S [%]	μ _a [%]	h	C _{h,α} ^{Ham}	r (vor Hampel)	r (nach Hampel)	Q _{C, RMSE} (vor Hampel)	Q _{C, RMSE} (nach Hampel)
1	105120	5	90	0	2	3	2	0.8084	1.0000	0.1437	0.0041
2	35040	15	90	0	2	3	2	0.8127	0.9997	0.1414	0.0067
3	17520	30	90	0	2	4	2	0.7966	0.9960	0.1497	0.0167
4	8760	60	90	0	2	3	2	0.7943	0.9868	0.1501	0.0300
5	105120	5	63	27	2	3	3	0.7130	0.9999	0.1335	0.0058
6	35040	15	63	27	2	4	2	0.7147	0.9991	0.1424	0.0079
7	17520	30	63	27	2	3	2	0.7075	0.9974	0.1384	0.0095
8	8760	60	63	27	2	3	2	0.7149	0.9876	0.1317	0.0213
9	105132	5	36	54	2	3	3	0.5817	0.9989	0.1306	0.0042
10	35044	15	36	54	2	3	3	0.6023	0.9975	0.1331	0.0068
11	17522	30	36	54	2	3	2	0.5657	0.9978	0.1329	0.0062
12	8761	60	36	54	2	3	2	0.6083	0.9897	0.1318	0.0138
13	105132	5	90	0	4	4	2	0.7087	0.9999	0.2056	0.0024
14	35044	15	90	0	4	3	2	0.7207	0.9986	0.1925	0.0098
15	17522	30	90	0	4	3	2	0.7140	0.9904	0.2009	0.0257
16	8761	60	90	0	4	3	2	0.7025	0.9744	0.2053	0.0420
17	105144	5	63	27	4	4	3	0.6207	0.9998	0.1838	0.0030
18	35048	15	63	27	4	4	2	0.6019	0.9982	0.1950	0.0080
19	17524	30	63	27	4	4	2	0.6014	0.9922	0.1961	0.0174
20	8762	60	63	27	4	3	2	0.6131	0.9806	0.1879	0.0260
21	105144	5	36	54	4	4	3	0.4731	0.9976	0.1879	0.0064
22	35048	15	36	54	4	4	2	0.4660	0.9963	0.1849	0.0080
23	17524	30	36	54	4	3	2	0.5045	0.9931	0.1847	0.0126
24	8762	60	36	54	4	3	2	0.4794	0.9806	0.1830	0.0176
25	105156	5	90	0	6	7	2	0.6474	0.9997	0.2616	0.0106
26	35052	15	90	0	6	4	2	0.6553	0.9967	0.2497	0.0167
27	17526	30	90	0	6	3	2	0.6594	0.9821	0.2455	0.0355

Fortsetzung der Tabelle

Zeit- reihe	K	T _S [h]	β _M [%]	β _S [%]	μ _a [%]	h	C _{h,α} ^{Ham}	r (vor Hampel)	r (nach Hampel)	Q _{C, RMSE} (vor Hampel)	Q _{C, RMSE} (nach Hampel)
28	8763	60	90	0	6	6	2	0.6535	0.9414	0.2514	0.0637
29	105156	5	63	27	6	5	2	0.5521	0.9995	0.2295	0.0054
30	35052	15	63	27	6	4	2	0.5399	0.9977	0.2392	0.0103
31	17526	30	63	27	6	5	2	0.5642	0.9839	0.2353	0.0240
32	8763	60	63	27	6	3	2	0.5584	0.9404	0.2323	0.0480
33	105168	5	36	54	6	6	2	0.4076	0.9947	0.2331	0.0098
34	35056	15	36	54	6	5	2	0.4245	0.9935	0.2325	0.0118
35	17528	30	36	54	6	4	2	0.4184	0.9889	0.2240	0.0142
36	8764	60	36	54	6	3	2	0.4480	0.9691	0.2223	0.0244
37	105168	5	90	0	8	6	2	0.6175	0.9995	0.2867	0.0056
38	35056	15	90	0	8	5	2	0.6254	0.9939	0.2816	0.0209
39	17528	30	90	0	8	5	2	0.6144	0.9725	0.2914	0.0441
40	8764	60	90	0	8	4	2	0.6068	0.9299	0.2917	0.0703
41	105180	5	63	27	8	6	2	0.5058	0.9992	0.2775	0.0066
42	35060	15	63	27	8	4	2	0.4927	0.9934	0.2888	0.0164
43	17530	30	63	27	8	4	2	0.5048	0.9777	0.2765	0.0282
44	8765	60	63	27	8	4	2	0.5243	0.9132	0.2509	0.0561
45	105180	5	36	54	8	5	2	0.3763	0.9946	0.2679	0.0096
46	35060	15	36	54	8	5	2	0.3725	0.9919	0.2665	0.0126
47	17530	30	36	54	8	4	2	0.3880	0.9847	0.2735	0.0160
48	8765	60	36	54	8	5	2	0.3867	0.9298	0.2598	0.0348

B.2.5 Zeitreihennormierung

In Tabelle B.10 ist eine Übersicht über gängige Normierungsverfahren bei Zeitreihen dargestellt.

Tabelle B.10: Übersicht über Normierungsmethoden bei Energiezeitreihen: Diese Tabelle ergänzt die Beschreibung der Normierung von Zeitreihen in Abschnitt 2.4.3 (modifiziert nach [85, 209, 235, 274, 276, 304]).

Methode (Kürzel) Berechnung	Beschreibung
Offset-Verschiebung (OV) $x_{n_D,OV}[k] = x_{n_D}[k] - \bar{x}_{n_D}$	<ul style="list-style-type: none"> Verschiebung der Zeitreihen $x_{n_D}[k]$ um Mittelwert $\bar{x}_{n_D} = \sum_{k=1}^K x_{n_D}[k]$ geeignet bei Zeitreihen mit unterschiedlichen Offsets Beseitigung von Offsets, Angleichung der Mittelwerte ($\bar{x}_{n_D,OV} = 0$) Varianzen der Zeitreihen bleiben bestehen negative Zeitreihenwerte möglich
Z-Score-Normalisierung (ZS) $x_{n_D,ZS}[k] = \frac{1}{\sigma_x} \cdot (x_{n_D}[k] - \bar{x}_{n_D})$	<ul style="list-style-type: none"> Verschiebung der Zeitreihen $x_{n_D}[k]$ um Mittelwert \bar{x}_{n_D} und Normierung mittels Standardabweichung $\sigma_x = \sqrt{\frac{1}{K-1} \sum_{k=1}^K (x_{n_D}[k] - \bar{x}_{n_D})^2}$ geeignet bei Zeitreihen mit unterschiedlichen Offsets und Amplituden Beseitigung von Offsets, Angleichung der Mittelwerte ($\bar{x}_{n_D,VN} = 0$) Varianzen der Zeitreihen werden angeglichen negative Zeitreihenwerte möglich
Mittelwertdivision (MD) $x_{n_D,MD}[k] = \frac{x_{n_D}[k]}{\bar{x}_{n_D}}$	<ul style="list-style-type: none"> Normierung der Zeitreihen $x_{n_D}[k]$ durch Division des Mittelwerts \bar{x}_{n_D} geeignet bei Zeitreihen mit unterschiedlichen Amplituden Beseitigung von Offsets, Angleichung der Mittelwerte ($\bar{x}_{n_D,MD} = 1$) Varianzen der Zeitreihen bleiben bestehen keine negativen Zeitreihenwerte
Min/Max-Normalisierung (MM) $x_{n_D,MM}[k] = \frac{x_{n_D}[k] - x_{n_D,\min}}{x_{n_D,\max} - x_{n_D,\min}}$	<ul style="list-style-type: none"> [0, 1]-Normierung der Zeitreihe Angleichung der Wertebereiche der Zeitreihen keine Beseitigung von Offsets Varianzen der Zeitreihen bleiben bestehen ungeeignet bei ausreißerbehafteten Zeitreihen keine negativen Zeitreihenwerte
Trendbereinigung (TB) $x_{n_D,TB}[k] = x_{n_D}[k] - \hat{x}_{n_D,T}[k]$	<ul style="list-style-type: none"> Trendbereinigung durch Verschiebung der Zeitreihe anhand der geschätzten, linearen Trendlinie $\hat{x}_{n_D,T}[k] = \beta_0 + \beta_1 \cdot k$ geeignet bei Zeitreihen mit unterschiedlichen Offsets und Trends Beseitigung von Offsets, Angleichung der Mittelwerte ($\bar{x}_{n_D,TB} = 0$) Beseitigung von Trends in Zeitreihen Varianzen der Zeitreihen bleiben bestehen ungeeignet bei Tageslastzeitreihen, da oftmals kein signifikanter Trend negative Zeitreihenwerte möglich

B.2.6 Zeitreihenaggregation

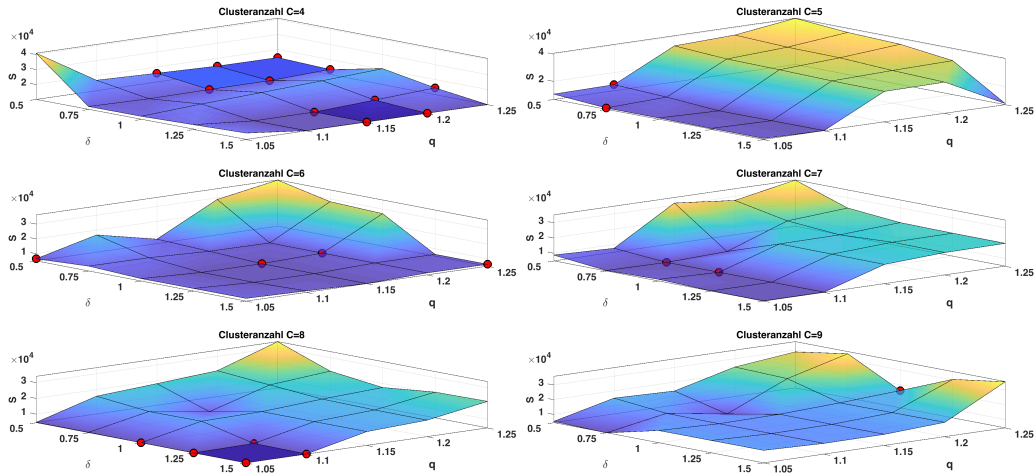
In Tabelle B.11 sind alle Ausgangsgrößen dargestellt, die für die verschiedenen *Smart-Meter*-Datensätze aus Tabelle 2.3 möglich sind.

Tabelle B.11: Mögliche Ausgangsgrößen für unterschiedliche Aggregationen der Datensätze; Kennzeichnung: –: nicht vorhanden, ✓: vorhanden

Ausgangsgröße	GH	GT	VH	VT	DH	DT	SH	ST
Jahr	–	–	–	–	✓	✓	✓	✓
Monat	–	–	–	–	✓	✓	✓	✓
Tag	–	–	–	–	✓	✓	✓	✓
Stunde	–	–	–	–	–	–	✓	✓
Wochentag	–	–	–	–	✓	✓	✓	✓
Werktag (ja/nein)	–	–	–	–	✓	✓	✓	✓
Tagesnummer	–	–	–	–	✓	✓	✓	✓
Tarifgruppe	✓	✓	✓	✓	✓	✓	✓	✓
Haushaltsnummer	✓	–	✓	–	✓	–	✓	–

B.2.7 Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern

(a) Sensitivität des Trennungsgrades



(b) Sensitivität des prozentualen Anteils der Datentupel im Rauschcluster

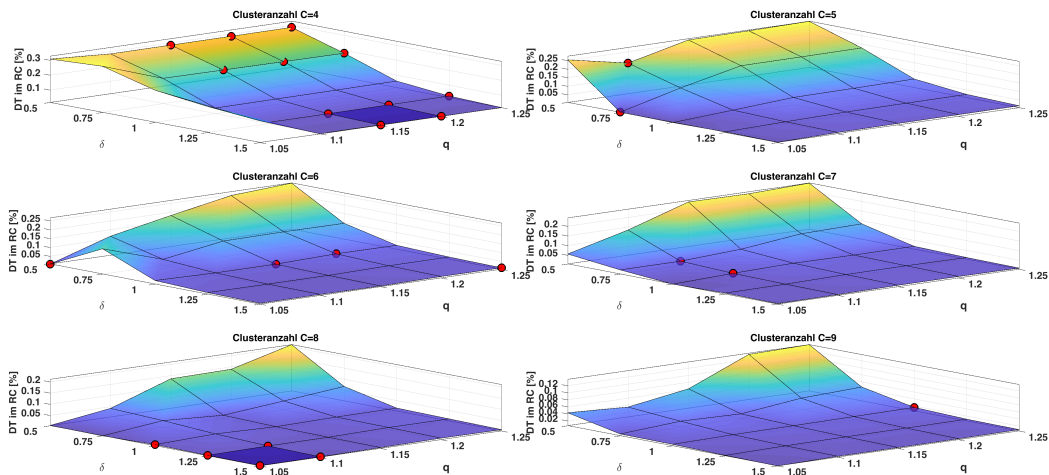


Abbildung B.5: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH_{norm} des Benchmarkdatensatzes „Settlement210H365D_70M20S“ ((a) Trennungsgrad, (b) prozentualer Anteil der Datentupel im Rauschcluster). Für Fall DH siehe Abb. 2.18, für Fall DH_{normmean} siehe Abb. B.7. Ergänzende Visualisierungen in Abb. 2.19. Die roten (alle Cluster enthalten Datentupel) und weißen (mindestens ein Cluster enthält keine Datentupel) Kreise kennzeichnen diejenigen δ, q -Parameterkombinationen, für die sich eine optimale Clusteranzahl einstellt.

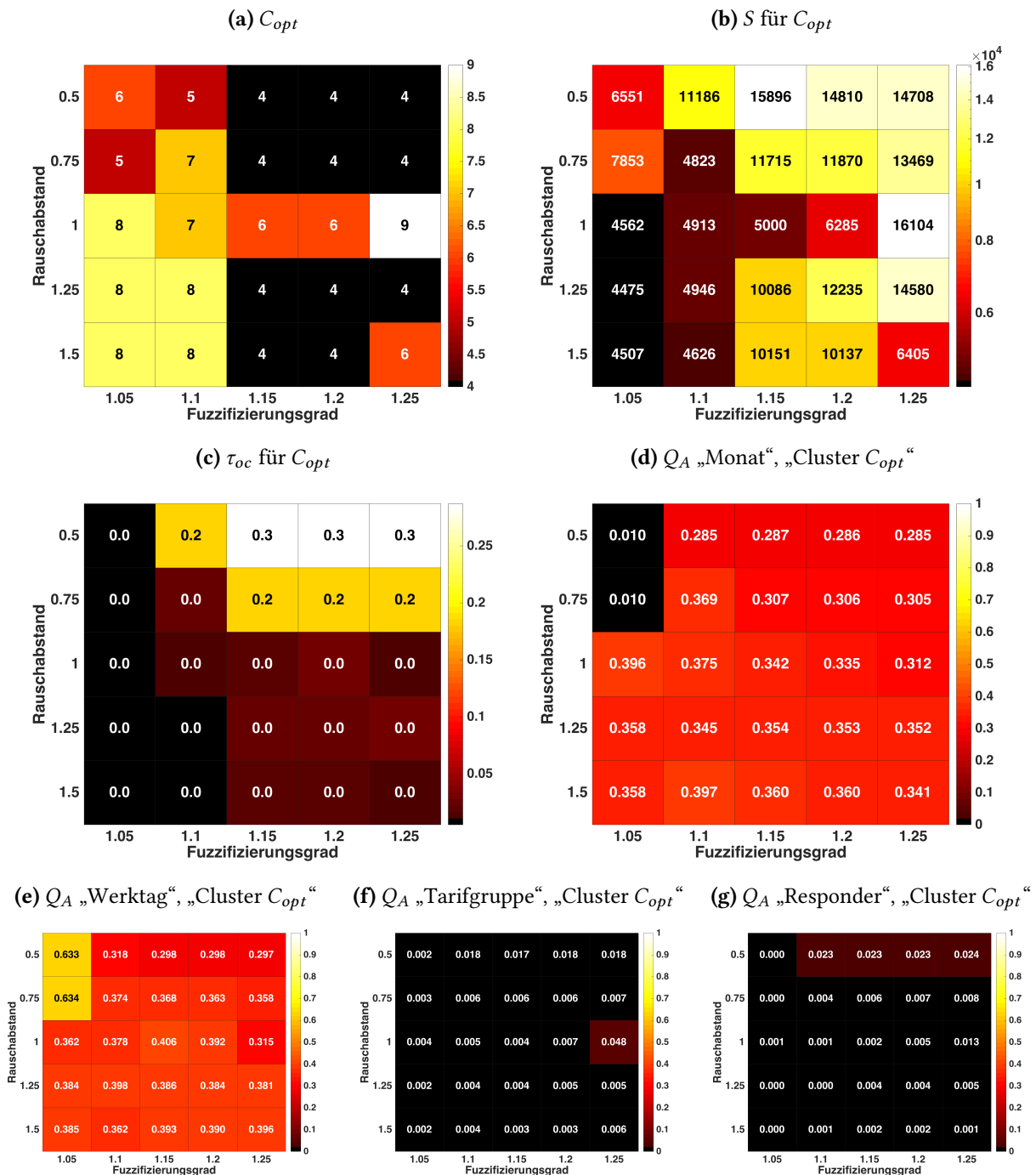
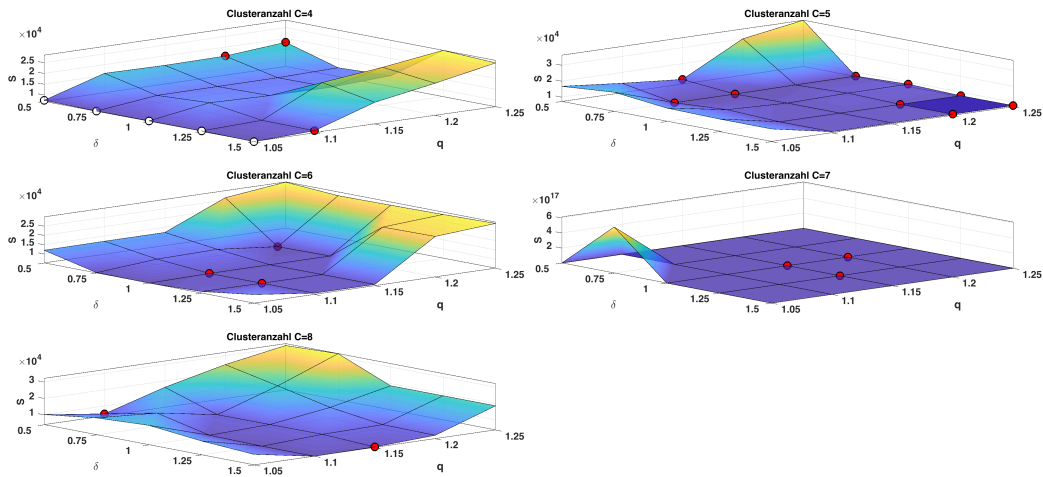


Abbildung B.6: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH_{norm} des Benchmarkdatensatzes „Settlement210H365D_70M20S“ ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(g) relative Transinformation pro Ausgangsentropie). Für Fall DH siehe Abb. 2.19, für Fall $DH_{normmean}$ siehe Abb. B.8. Ergänzende Visualisierungen zu Abb. B.5. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

(a) Sensitivität des Trennungsgrades



(b) Sensitivität des prozentualen Anteils der Datentupel im Rauschcluster

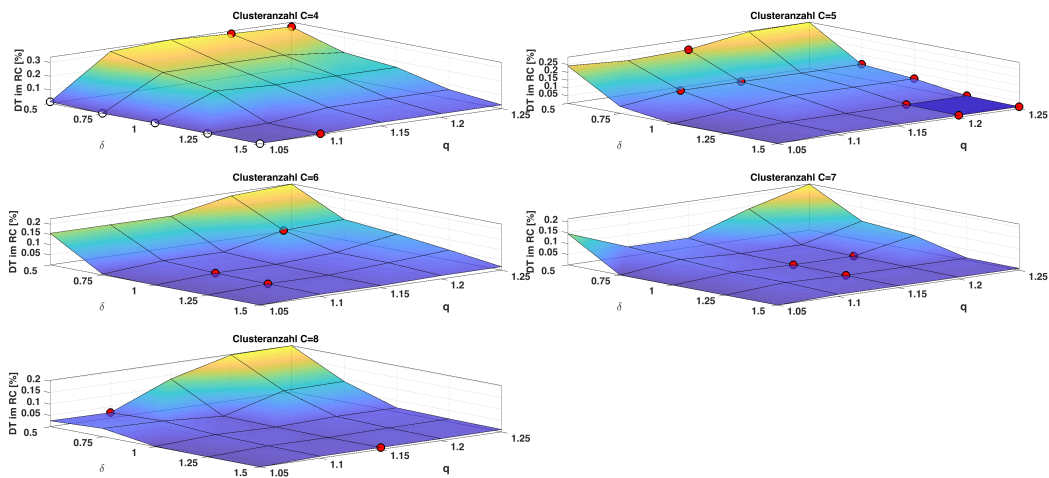


Abbildung B.7: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}$ des Benchmarkdatensatzes „Settlement210H365D_70M20S“ ((a) Trennungsgrad, (b) prozentualer Anteil an Datentupeln im Rauschcluster). Für Fall DH siehe Abb. 2.18, für Fall DH_{norm} siehe Abb. B.5. Ergänzende Visualisierungen in Abb. B.8. Die roten (alle Cluster enthalten Datentupel) und weißen (mindestens ein Cluster enthält keine Datentupel) Kreise kennzeichnen diejenigen δ, q -Parameterkombinationen, für die sich eine optimale Clusteranzahl einstellt.

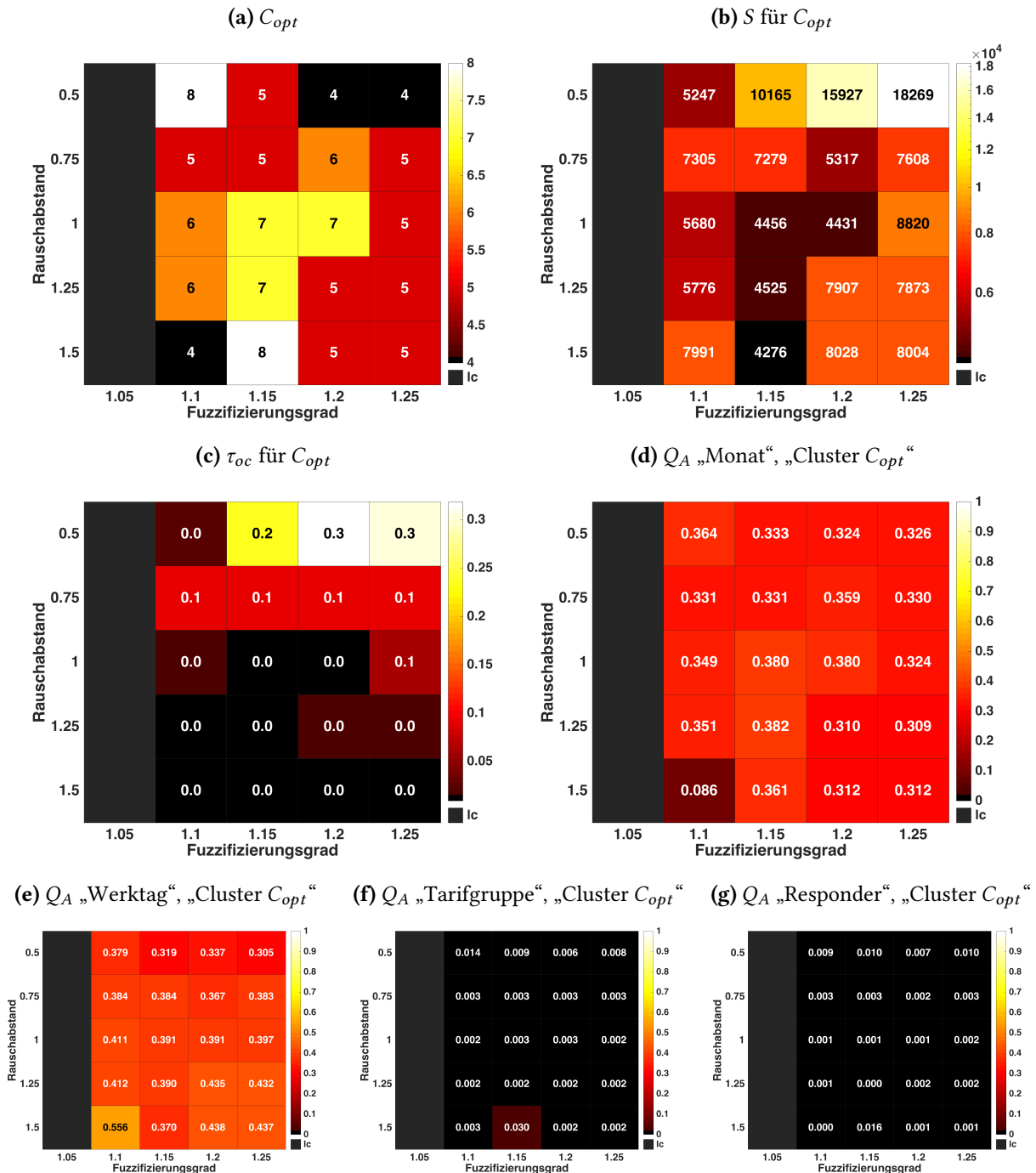
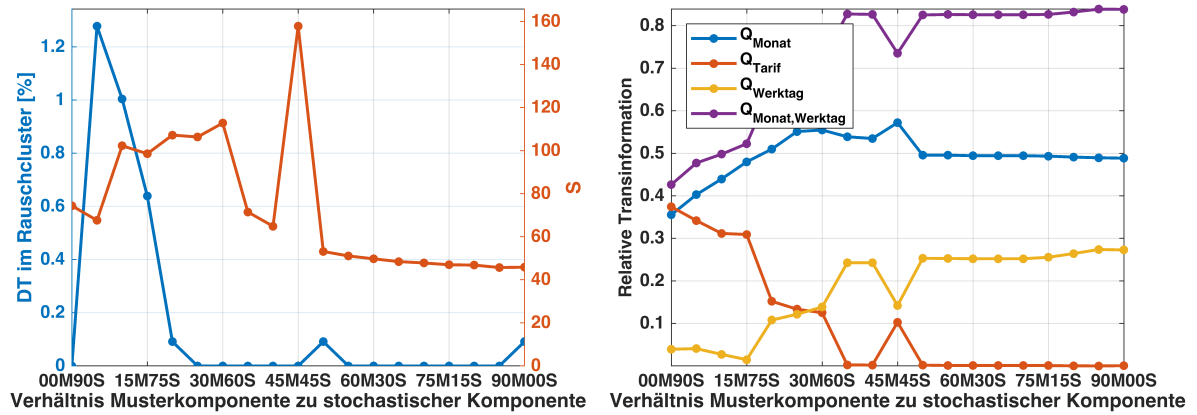
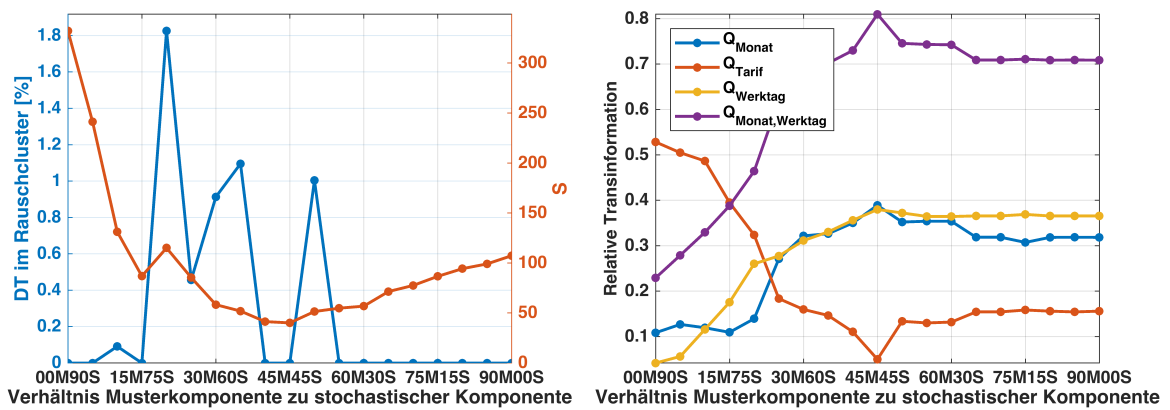


Abbildung B.8: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}$ des Benchmarkdatensatzes „Settlement210H365D_70M20S“ ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(g) relative Transinformation pro Ausgangsentropie). Für Fall DH siehe Abb. 2.19, für Fall DH_{norm} siehe Abb. B.6. Ergänzende Visualisierungen zu Abb. B.7. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

(a) DT-Fälle: Clusterparameter $C = 6$, $q = 1.1$ und $\delta = 1.25$



(b) DT_{norm} -Fälle: Clusterparameter $C = 8$, $q = 1.05$ und $\delta = 1.0$



(c) DT_{normmean} -Fälle: Clusterparameter $C = 7$, $q = 1.15$ und $\delta = 1.0$

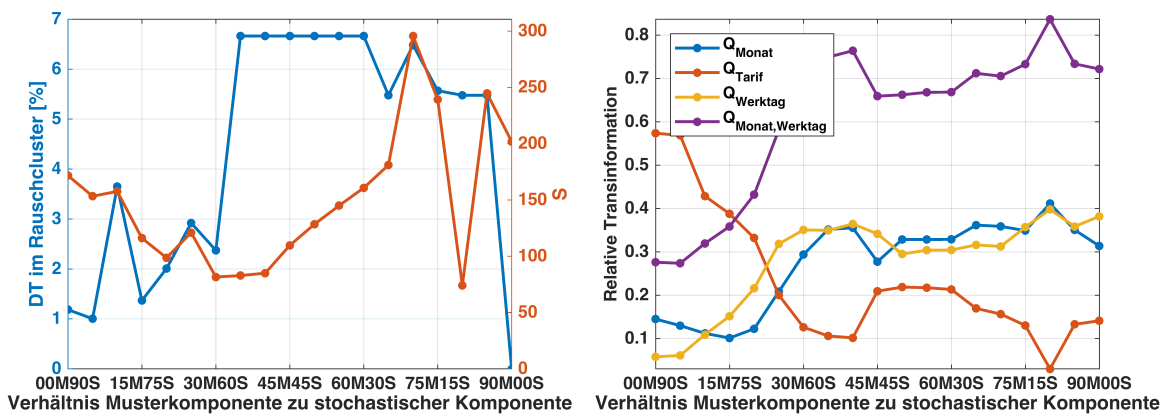


Abbildung B.9: Sensitivität der Clusterergebnisse bezüglich der Lastsignalcharakteristik für (a) DT-Fälle, (b) DT_{norm} -Fälle und (c) DT_{normmean} -Fälle. Die hier verwendeten Clusterparameter sind die gleichen wie in Abb. 2.20.

B.2.8 Sensitivität der Clusterergebnisse bezüglich der Wahl der Startcluster

Eine übersichtliche Darstellung des Vergleichs der verschiedenen Strategien zur Wahl der Startcluster ist in Abb. 2.21, B.10 und B.11 zu finden. Die Konfiguration dieser Sensitivitätsanalyse zeigt Tabelle 2.6.

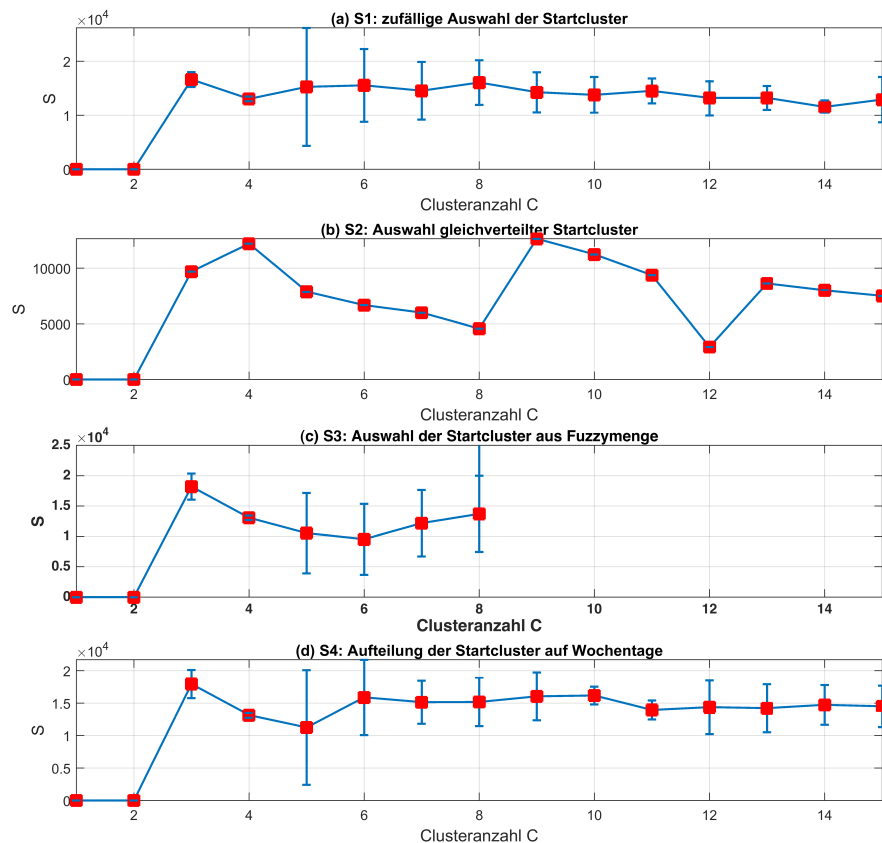


Abbildung B.10: Sensitivität der Clusterergebnisse bezüglich der Wahl der Startcluster für den DH_{norm} -Fall des Benchmarkdatensatzes „Settlement210H365D_70M20S“ (für Fall DH siehe Abb. 2.21, für Fall DH_{normmean} siehe Abb. B.11): das Clustering wurde 10-fach mit den Clusterparametern $C = 3, \dots, 15$, $q = 1.05$ und $\delta = 1.0$ durchgeführt. Die blaue Linie kennzeichnet den mittleren Verlauf der Werte für S bei 10-facher Wiederholung des Clusterings. Die Balken geben die Streuung der Werte für S um den Mittelwert an. In Abb. B.10(c) wurde die Skalierung der y-Achse so gewählt, dass das lokale Minimum bei $C = 6$ gut erkennbar ist. Die Werte für $S(C > 8)$ sind aus diesem Grund nicht in der Abbildung sichtbar.

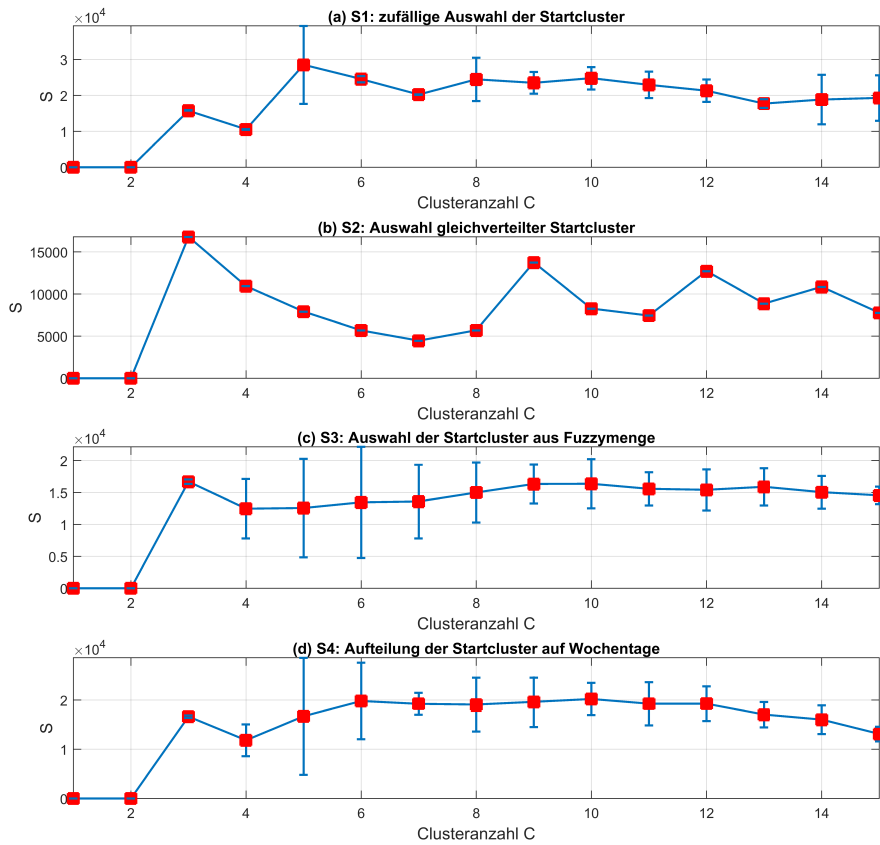


Abbildung B.11: Sensitivität der Clusterergebnisse bezüglich der Wahl der Startcluster für den DH_{normmean} -Fall des Benchmarkdatensatzes „Settlement210H365D_70M20S“ (für Fall DH siehe Abb. 2.21, für Fall DH_{norm} siehe Abb. B.10): das Clustering wurde 10-fach mit den Clusterparametern $C = 3, \dots, 15$, $q = 1.15$ und $\delta = 1.0$ durchgeführt. Die blaue Linie kennzeichnet den mittleren Verlauf der Werte für S bei 10-facher Wiederholung des Clusterings. Die Balken geben die Streuung der Werte für S um den Mittelwert an.

B.3 Implementierung

Definition der MATLAB Skripte und Funktionen

Es existieren zwei Typen von Programmeinheiten in MATLAB: Skripte und Funktionen. Skripte laufen im MATLAB-Arbeitsbereich (auch *Workspace* genannt) ab und es werden keine Übergabeparameter verwendet. Somit sind einem Skript alle im *Workspace* definierten Variablen bekannt. Bei einer Funktion werden sogenannte Übergabeparameter verwendet, um die Kommunikation mit anderen Funktionen oder Skripten zu realisieren.

Definition des MATLAB *Timetable*-Datenformats

Das *Timetable*-Datenformat in MATLAB stellt eine Tabelle dar, bei der jede Zeile mit einer Zeitangabe verknüpft ist. Die *Timetable* speichert spaltenorientierte Datenvariablen, die unterschiedliche Datentypen (z.B. *string* oder *double*) und Größen haben. Des Weiteren ermöglicht das *Timetable*-Format die Benutzung zeitspezifischer Funktionen, um beispielsweise mehrere *Timetables* anhand des Zeitstempels zu verknüpfen oder zu synchronisieren.

Skripte und Funktionen der SciXMiner-Erweiterungen *Load Benchmark Generator* und *Energy*

In Listing B.1 ist die Funktion `load_benchmark_generator_settlement.m` zur Erzeugung von Benchmarkdatensatzes dargestellt.

Listing B.1: Funktion `load_benchmark_generator_settlement.m`

```

1  % Read data from datastore and generate matlab table "SETTLEMENT"
2  filename = fullfile(op,...
3      'Settlements',...
4      [parameter.gui.load_benchmark_generator.General.sn, '.csv']);
5  % store csv-file as datastore
6  ds = tabularTextDatastore(filename, 'TreatAsMissing', 'NA');
7  % read one file at a time
8  ds.ReadSize = 'file';
9  % read data from datastore and generate matlab table "SETTLEMENT"
10 SETTLEMENT = read(ds);
11 % reset datastore before re-reading
12 reset(ds);
13
14 % Generate prjz-file for each customer in settlement
15 for iC = 1:size(SETTLEMENT,1)
16
17     % display
18     disp(' ')
19     disp('-----')
20     disp(['Start with household number ', num2str(iC), ' of ', num2str(size(SETTLEMENT,1))])
21
22     % Step 1.1: General, settings, and inputs
23     defineSettingsSettlement;
24
25     % Step 1.2: Set all paths to all folders and subfolders
26     setPaths;
27
28     % Step 2: Load data -> signal generator or real data importer
29     % Step 2.1: Signal generator (C) Stefan Klaiber
30     if strcmp(INPUT.General.sigGenORdatImp, 'Signal generator')

```

```

31     TT = runSigGen(INPUT);
32
33     % Step 2.2: Real data importer
34     elseif strcmp(INPUT.General.sigGenORdatImp, 'Data importer')
35         TT = runDatImp(INPUT);
36     end
37
38     % Step 3: Price importer
39     TT = runPriImp(TT, INPUT);
40
41     % Step 4: Demand Response simulator
42     TT = runDRSim(TT, INPUT);
43
44     % Step 5: Data manipulator
45     TT = runDatMan(TT, INPUT);
46
47     % Step 6: Save all results and settings as mat-files and csv-files
48     runSaveMAT;
49     runSaveCSV
50
51     % Step 7.: Temporary plot
52     if INPUT.General.tempPlots==1
53         temporaryPlot(TT, INPUT);
54     end
55
56     % Step 8: Generate SciXminer project form timetable data
57     timetable2scixminer(TT, INPUT);
58
59     % delete temporary variables
60     clearvars TT
61
62 end
63
64 % Automatically fusion projects that are in "SettlementX" folder
65 if INPUT.General.projectFusion==1
66     mode_new_project=2;
67     next_function_parameter = [INPUT.General.pathResults, '/', ...
68         INPUT.General.nameSettlement, '/', INPUT.TT2SciXminer.selOV];
69     project_fusion;
70 end
71
72 % delete temporary variables
73 clearvars SETTLEMENT
74 clearvars op ic ds

```

In Listing B.2 ist das Skript `defineNames.m` zur Auswahl der Bezeichner dargestellt.

Listing B.2: Auszug aus dem Skript `defineNames.m`

```

1 %% defining possible names for relevant time series
2 namesTSLoad = {'P_mess', 'P_{mess}', 'Load'};
3 namesTSPrice = {'price', 'Preis', 'p', 'Strompreis'};
4 namesTSDayOfYear = {'DayOfYear'};
5 namesTSYear = {'Year', 'JAHR'};
6 namesTSMonth = {'MonthMM', 'MONAT', 'Monat'};
7 namesTSDayOfMonth = {'DayOfMonth', 'TAG'};
8 namesTSHour = {'Hour', 'Stunde', 'STUNDE'};
9 namesTSMinute = {'Minute', 'MINUTE'};
10 namesTSSecond = {'Second', 'Sekunde', 'SEKUNDE'};
11 % determine load time series and price time series for each loaded project
12 bezTSLoad = deblank(var_bez(ismember(var_bez, namesTSLoad), :));
13 bezTSPrice = deblank(var_bez(ismember(var_bez, namesTSPrice), :));
14 bezTSDayOfYear = deblank(var_bez(ismember(var_bez, namesTSDayOfYear), :));
15 bezTSYear = deblank(var_bez(ismember(var_bez, namesTSYear), :));
16 bezTSMonth = deblank(var_bez(ismember(var_bez, namesTSMonth), :));
17 bezTSDayOfMonth = deblank(var_bez(ismember(var_bez, namesTSDayOfMonth), :));
18 bezTSHour = deblank(var_bez(ismember(var_bez, namesTSHour), :));

```

```

19 bezTSMinute      = deblank(var_bez(ismember(var_bez,namesTSMinute),:));
20 bezTSSecond     = deblank(var_bez(ismember(var_bez,namesTSSecond),:));
21 % determine load and price feature for each loaded project
22 bezEMLoad       = deblank(dorgbez(ismember(dorgbez,namesTSLoad),:));
23 bezEMPrice      = deblank(dorgbez(ismember(dorgbez,namesTSPrice),:));
24 % delete temporary variables
25 clearvars namesTSLoad namesTSPrice
26 clearvars namesTSDayOfYear namesTSYear namesTSMonth namesTSDayOfMonth
27 clearvars namesTSHour namesTSMinute namesTSSecond
28
29 %% defining possible names for relevant output variables
30 namesOVYear      = {'Year','JAHR'};
31 namesOVMonth     = {'Month','MONAT'};
32 namesOVMonthMM   = {'MonthMM','MONATMM'};
33 namesOVDayOfMonth = {'DayOfMonth','TAG'};
34 namesOVWeekday   = {'DayOfWeek','WOCHENTAG'};
35 namesOVWorkdayYN = {'Werktag (ja/nein)','WERKTAG (ja/nein)'};
36 namesClassWorkdayYN = {'ja'};
37 namesOVWorkday   = {'bd','Typtag'};
38 namesClassWorkday = {'Business day','Werktage'};
39 namesOVDayNumber = {'DayOfYear','Tagesnummer'};
40 namesOVGroup     = {'Group','Vertrag'};
41 namesOVHousehold = {'File','CustomerID'};
42 namesOVPrice     = {'price','Preis','p','Strompreis'};
43 namesOVResponder = {'Responder'};
44 % determine output variables for each loaded project
45 bezOVYear        = deblank(bez_code(ismember(bez_code,namesOVYear),:));
46 bezOVMonth       = deblank(bez_code(ismember(bez_code,namesOVMonth),:));
47 bezOVMonthMM     = deblank(bez_code(ismember(bez_code,namesOVMonthMM),:));
48 bezOVDayOfMonth  = deblank(bez_code(ismember(bez_code,namesOVDayOfMonth),:));
49 bezOVWeekday     = deblank(bez_code(ismember(bez_code,namesOVWeekday),:));

```

In Listing B.3 ist ein Auszug aus der Funktion `vsa_en.m` zum Entwurf des VSA-Modells dargestellt.

Listing B.3: Auszug aus der Funktion `vsa_en.m`

```

1 % preallocate model evaluation matrix
2 Q = NaN(size(Kh,1),6);
3
4 % loop for each model order
5 for iKh = 1:size(Kh,1)
6
7     % optimization algorithm: settings for lsqnonlin algorithm
8     options = optimset('Algorithm','trust-region-reflective');
9
10    % optimization: start values for [a Emin Esoll Emax]
11    par0 = [0.1, 0.1, 0.1, 0.1];
12    % optimization: lower bounds for [a Emin Esoll Emax]
13    lb = [0, 0, 0, 0];
14    % optimization: upper bounds for [a Emin Esoll Emax]
15    ub = [inf, 20, 20, 20];
16
17    % some other parameters
18    constants.Kh      = Kh(iKh);
19    constants.t_react = parMod.t_react;
20    constants.dTs     = parMod.dTs;
21    constants.price_criteria = parMod.price_criteria;
22    constants.price_calc   = parMod.price_calc;
23
24    % run the lsqnonlin algorithm
25    [parOpt, resnorm] = lsqnonlin(@compute_error, par0, lb, ub, options, ...
26        x_real, y_real, constants);
27
28    % get to best parameters:
29    % best parameter value for a
30    a_fit = parOpt(1);

```

```
31     % best parameter value for Emin
32     Emin_fit = parOpt(2);
33     % best parameter value for Esoll
34     Esoll_fit = parOpt(3);
35     % best parameter value for Emax
36     Emax_fit = parOpt(4);
37     % residual
38     Q_sse = resnorm;
39     % write model evaluation matrix
40     Q(iKh,1) = a_fit;
41     Q(iKh,2) = Kh(iKh);
42     Q(iKh,3) = Emin_fit;
43     Q(iKh,4) = Esoll_fit;
44     Q(iKh,5) = Emax_fit;
45     Q(iKh,6) = Q_sse;
46 end
47
48 % find position of best model with respect to minimal residual
49 pos = find(ismember(Q(:,6),min(Q(:,6)))==1);
```

In Listing B.4 ist ein Auszug aus der Funktion `vsa_an.m` zur Anwendung des VSA-Modells dargestellt.

Listing B.4: Auszug aus der Funktion `vsa_an.m`

```
1 %% calculate the demand response for each timestep
2 ydach = vsa(d_val, parMod);
```

B.4 Anwendungen

Das neue Datenanalyse- und Modellierungskonzept wird in Anhang B.4 an realen *Smart-Meter*-Daten angewendet. In Ergänzung zu Kapitel 4 werden in Anhang B.4 weitere Auswertegrafiken und -tabellen aufgeführt.

B.4.1 Olympic-Peninsula-Project-Datensatz

Die Struktur des OPP-Datensatzes ist in Tabelle B.12 zusammengefasst.

Tabelle B.12: Struktur des OPP-Datensatzes gemäß Tabelle 2.3

Aggregationsniveau	Daten-tupel (N_D)	Zeitreihenlänge (K)	Zeitreihen (s_z), Einzelmerkmale (s)	Abtastperiodendauer (T_S) in Stunden
Fall GH	84	34656	$s_z = 2$ ($p[k]$ in ct/kWh, $P_n[k]$ in kW)	0.25
Fall GT	3	34656	$s_z = 2$ ($p[k]$ in ct/kWh, $P_n[k]$ in kW)	0.25
Fall DH	21318	96	$s_z = 2$ ($p[k]$ in ct/kWh, $P_n[k]$ in kW)	0.25
Fall DH ^{wo}	14069	96	$s_z = 2$ ($p[k]$ in ct/kWh, $P_n[k]$ in kW)	0.25
Fall DT	931	96	$s_z = 2$ ($p[k]$ in ct/kWh, $P_n[k]$ in kW)	0.25
Fall DT ^{wo}	613	96	$s_z = 2$ ($p[k]$ in ct/kWh, $P_n[k]$ in kW)	0.25
Fall SH	2046528	–	$s = 2$ (p in ct/kWh, P_n in kW)	–
Fall SH ^{wo}	1350624	–	$s = 2$ (p in ct/kWh, P_n in kW)	–
Fall ST	89376	–	$s = 2$ (p in ct/kWh, P_n in kW)	–
Fall ST ^{wo}	58848	–	$s = 2$ (p in ct/kWh, P_n in kW)	–

B.4.1.1 Bestimmung von optimalen Clusterparametern für Fall DH_{norm}

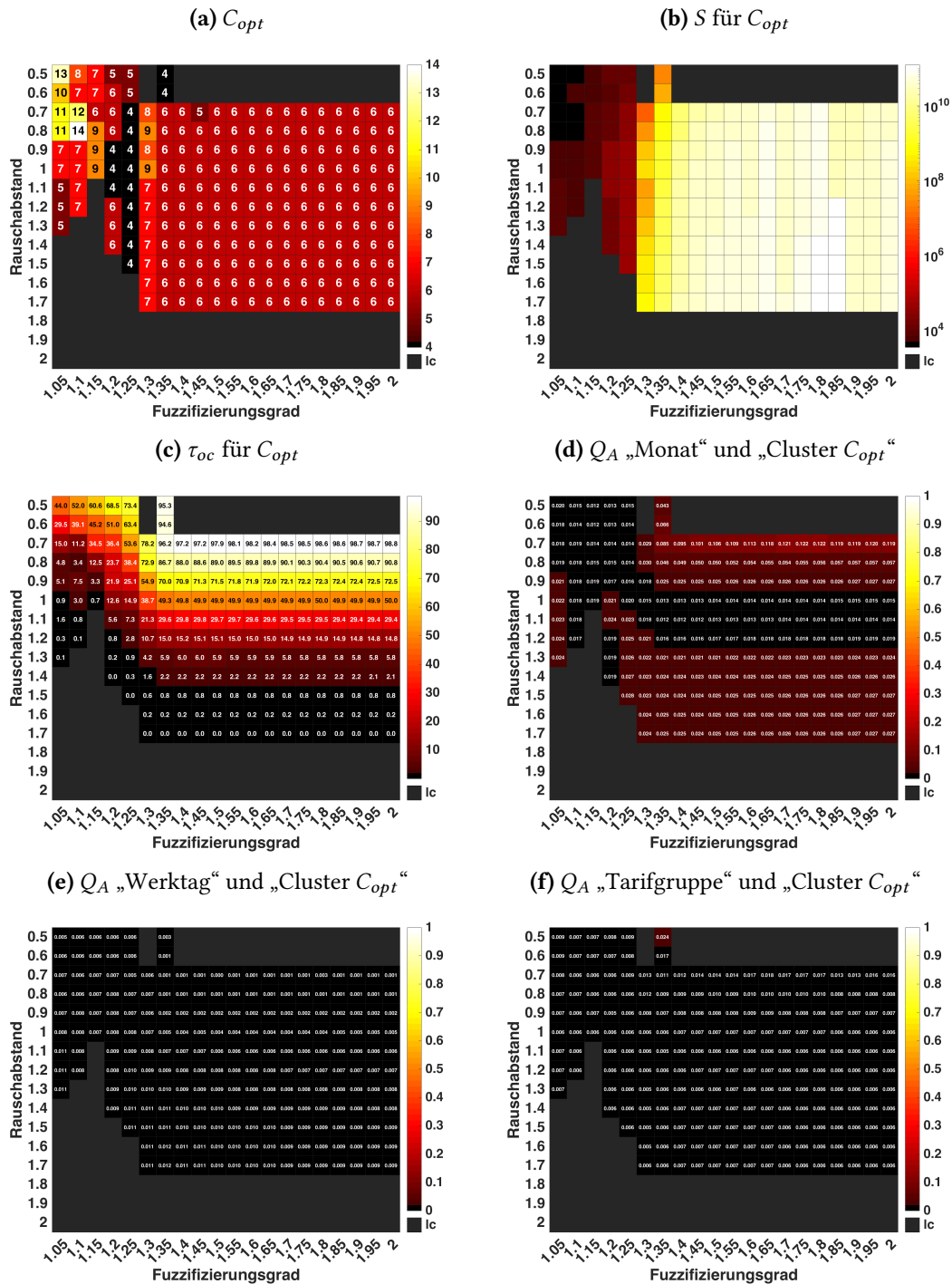


Abbildung B.12: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH_{norm} des OPP-Datensatzes (a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(f) relative Transinformation pro Ausgangsentropie). Für Fall DH siehe Abb. 4.2, für Fall $DH_{normmean}$ siehe Abb. B.13. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.1.2 Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}$

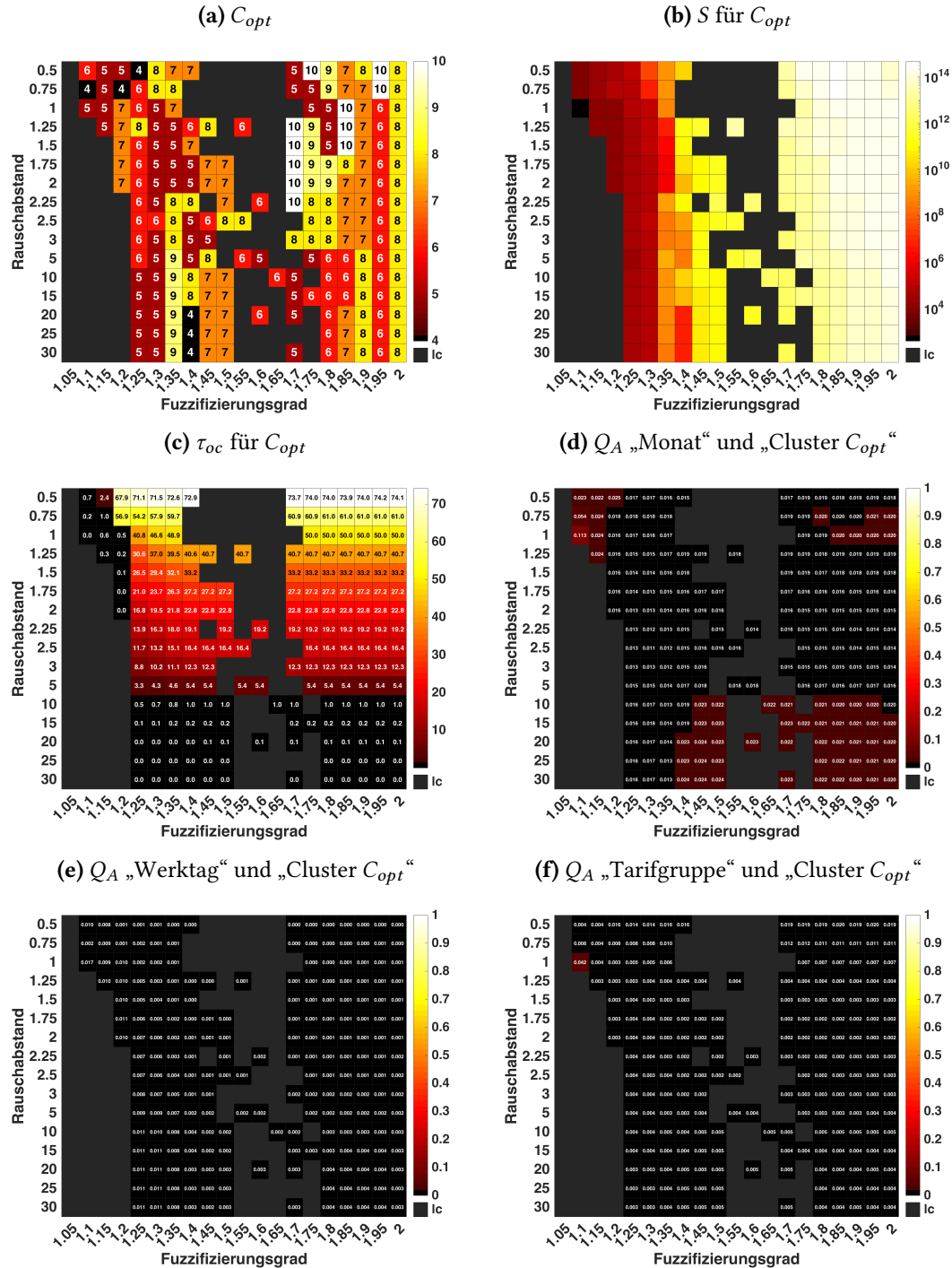


Abbildung B.13: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}$ des OPP-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(f) relative Transinformation pro Ausgangsentropie). Für Fall DH siehe Abb. 4.2, für Fall DH_{norm} siehe Abb. B.12. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.1.3 Bestimmung von optimalen Clusterparametern für Fall DH^{wo}

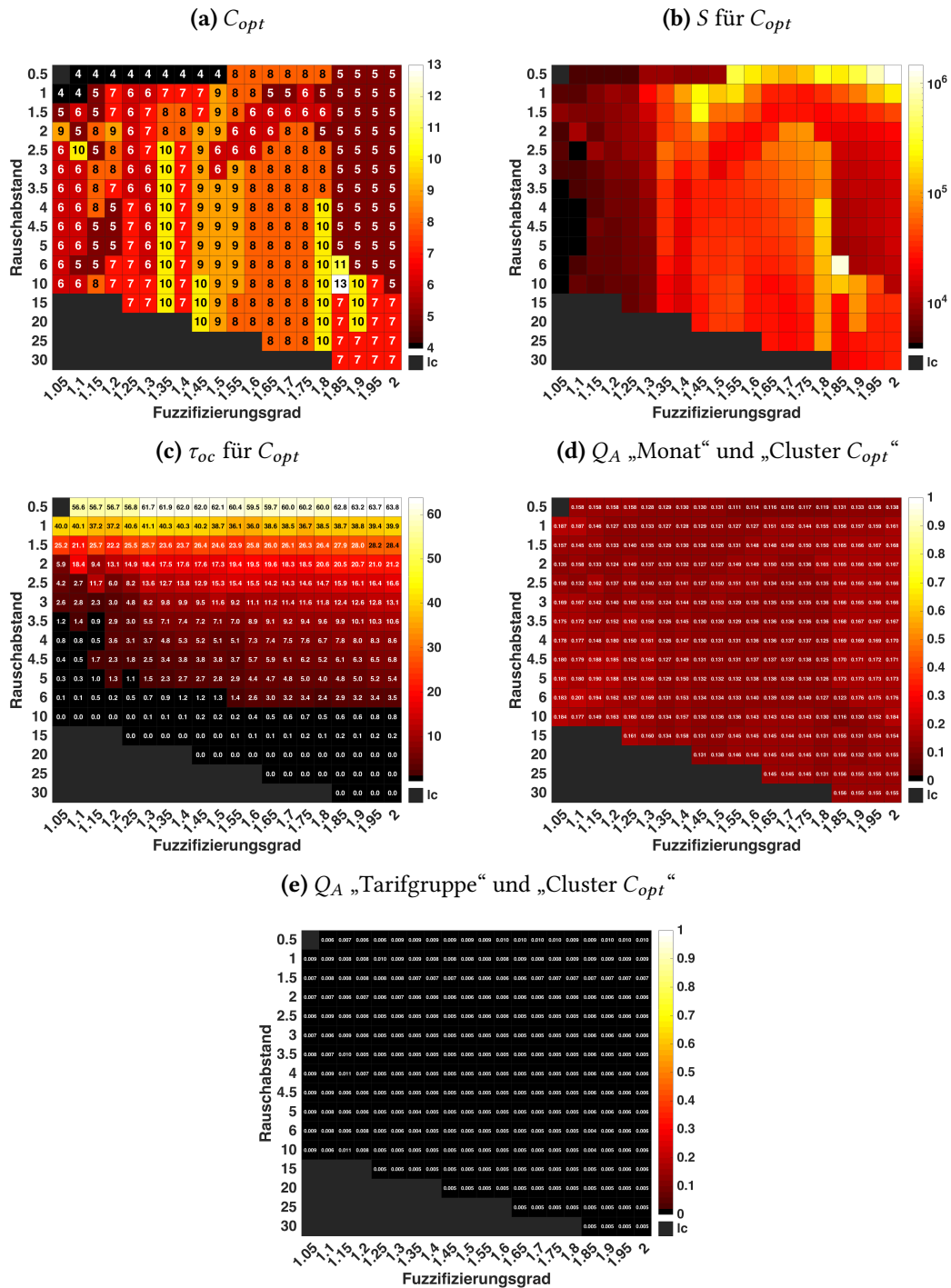


Abbildung B.14: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH^{wo} des OPP-Datensatzes (a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(e) relative Transinformation pro Ausgangsentropie). Für Fall DH^{wo} siehe Abb. B.15, für Fall DH^{wo}_{norm} siehe Abb. B.16. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.1.4 Bestimmung von optimalen Clusterparametern für Fall DH_{norm}^{wo}

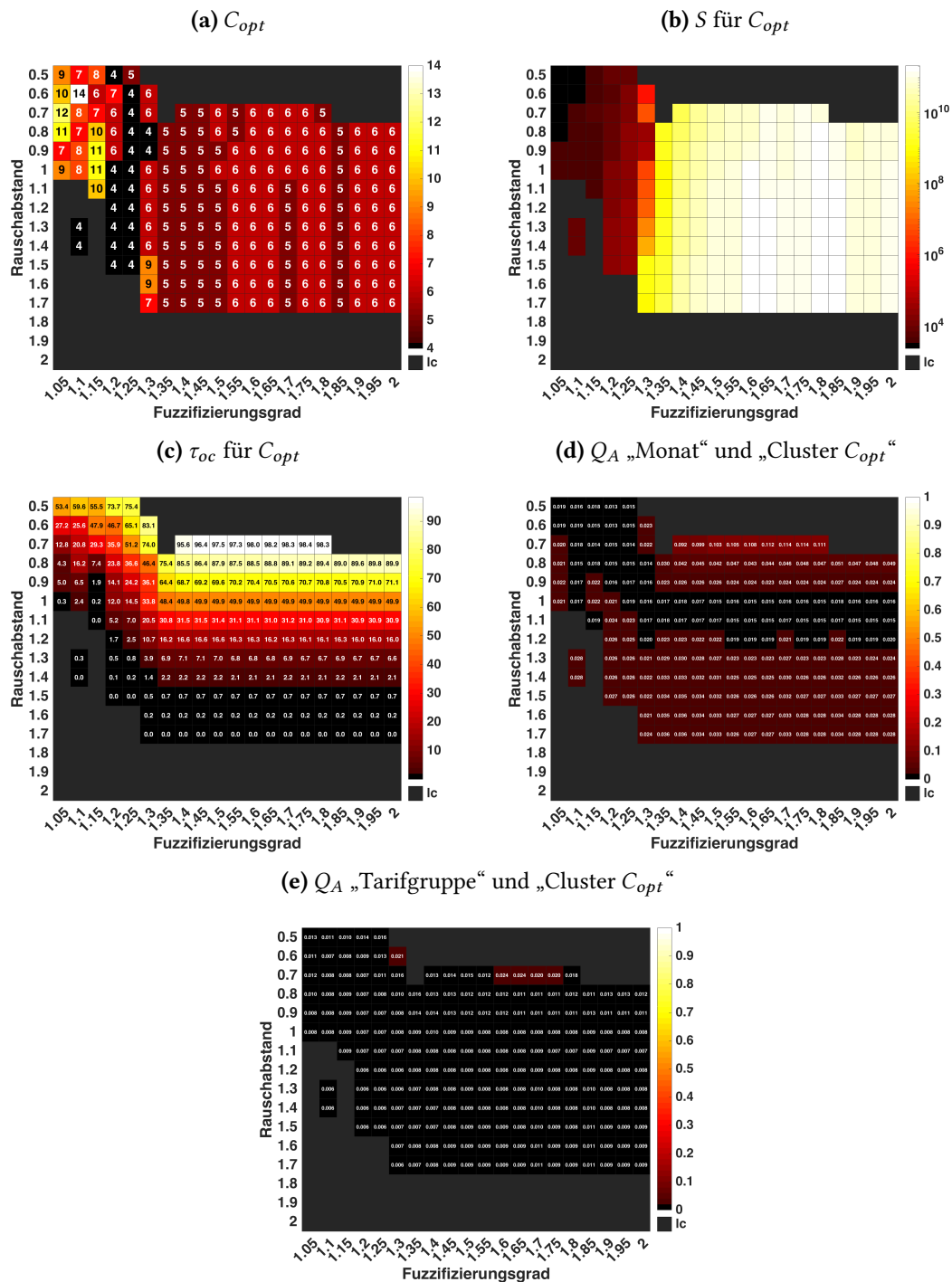


Abbildung B.15: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH_{norm}^{wo} des OPP-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(e) relative Transformation pro Ausgangsentropie). Für Fall DH_{norm}^{wo} siehe Abb. B.14, für Fall $DH_{norm}^{wo,mean}$ siehe Abb. B.16. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.1.5 Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}^{wo}$

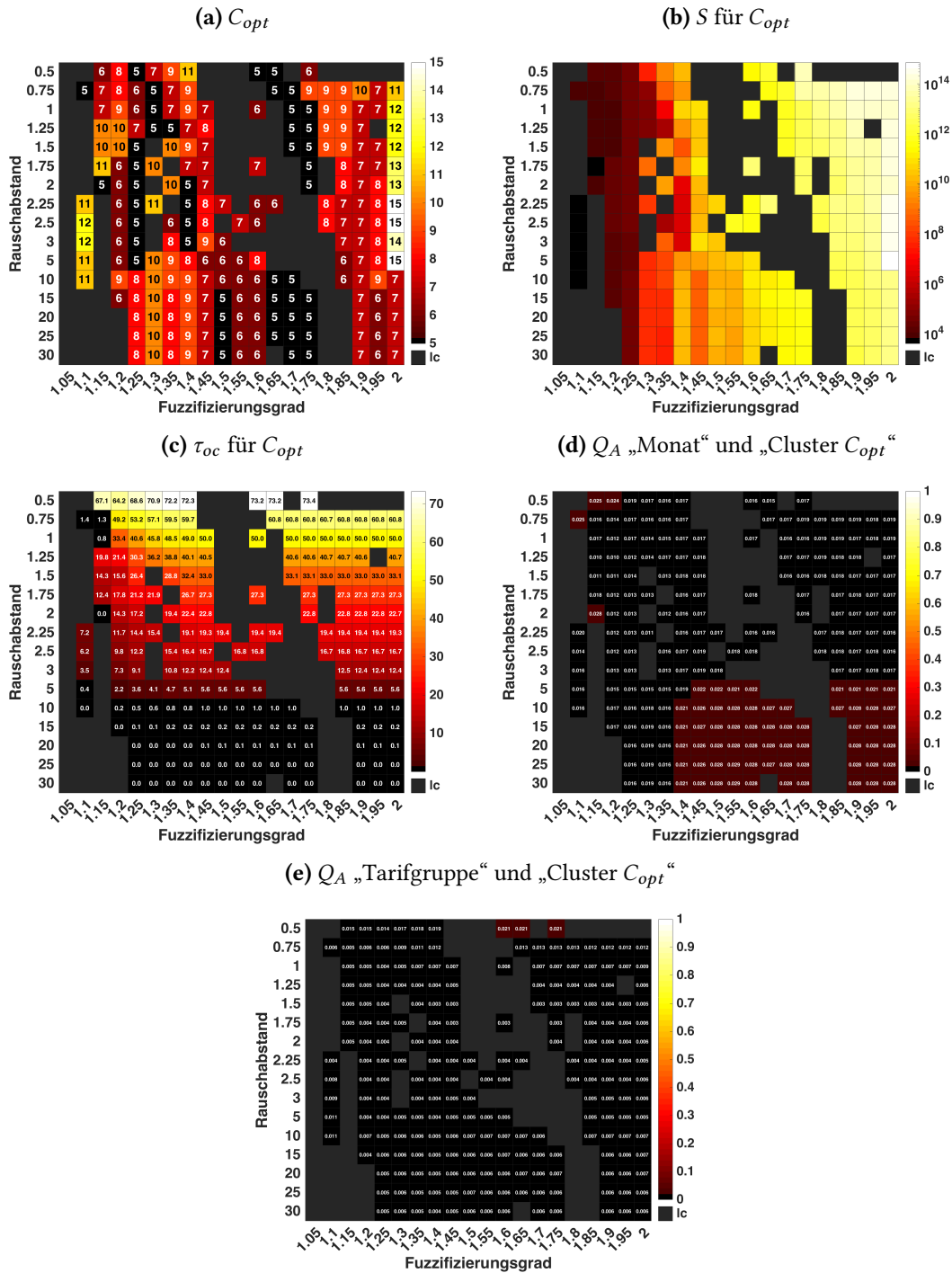


Abbildung B.16: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}^{wo}$ des OPP-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(e) relative Transinformation pro Ausgangsentropie pro Ausgangsentropie). Für Fall DH^{wo} siehe Abb. B.14, für Fall DH_{norm}^{wo} siehe Abb. B.15. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.1.6 Bestimmung von optimalen Clusterparametern für Fall DT

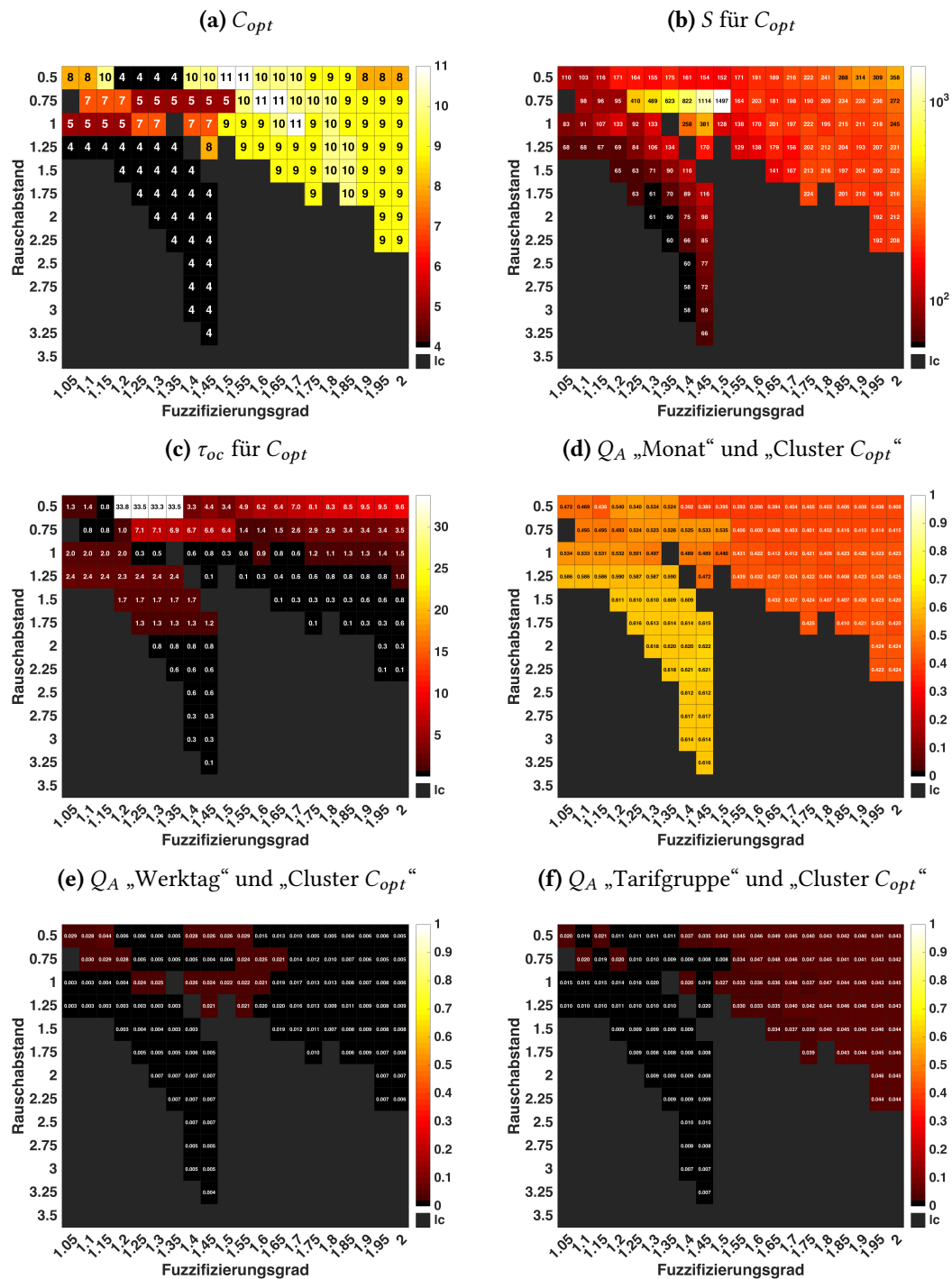


Abbildung B.17: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DT des OPP-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(f) relative Transinformation pro Ausgangsentropie). Für Fall DT_{norm} siehe Abb. B.18, für Fall DT_{normmean} siehe Abb. B.19. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.1.7 Bestimmung von optimalen Clusterparametern für Fall DT_{norm}

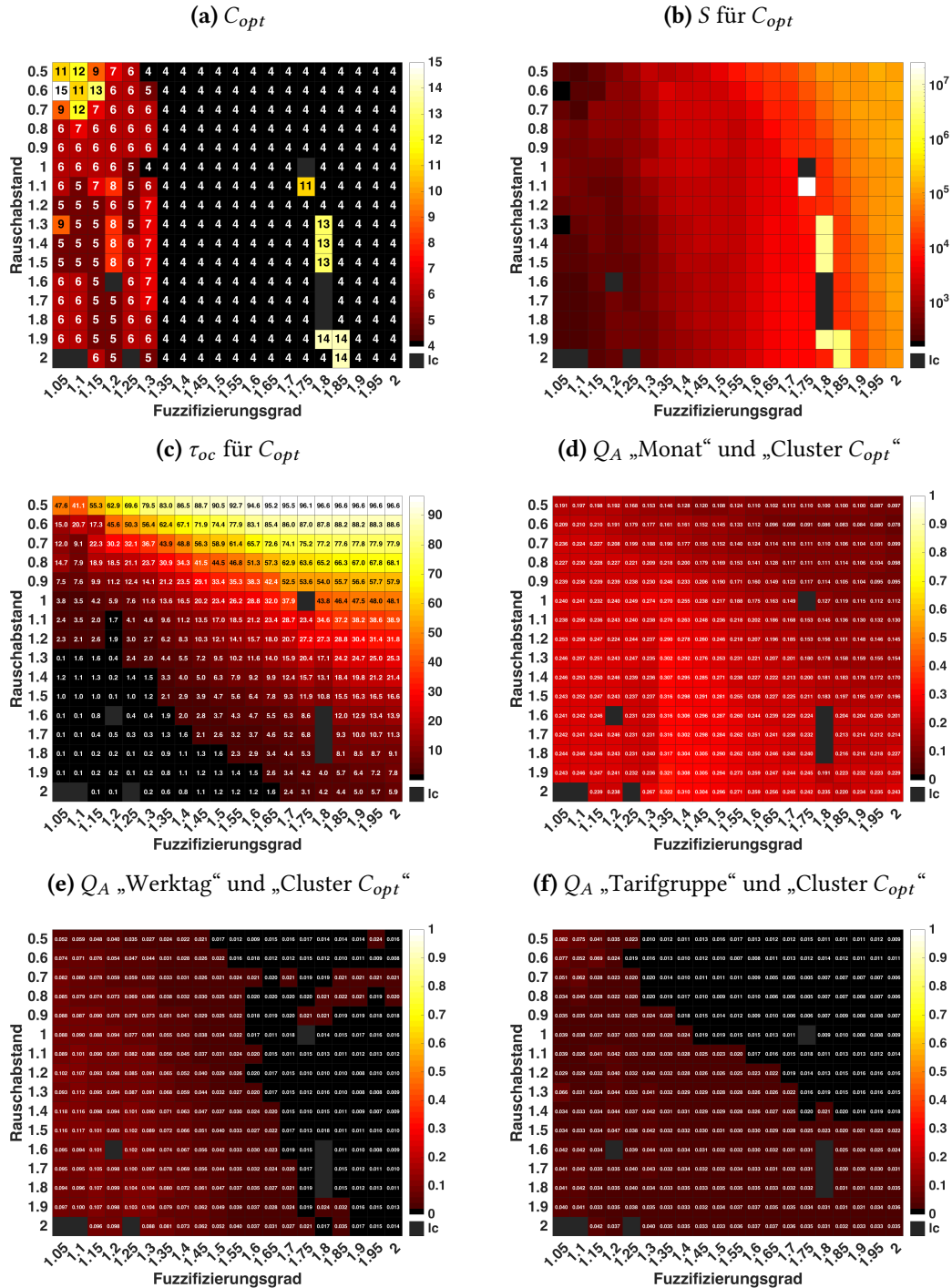


Abbildung B.18: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DT_{norm} des OPP-Datensatzes (a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(f) relative Transinformation pro Ausgangsentropie). Für Fall DT siehe Abb. B.17, für Fall DT_{normmean} siehe Abb. B.19. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.1.8 Bestimmung von optimalen Clusterparametern für Fall DT_{normmean}

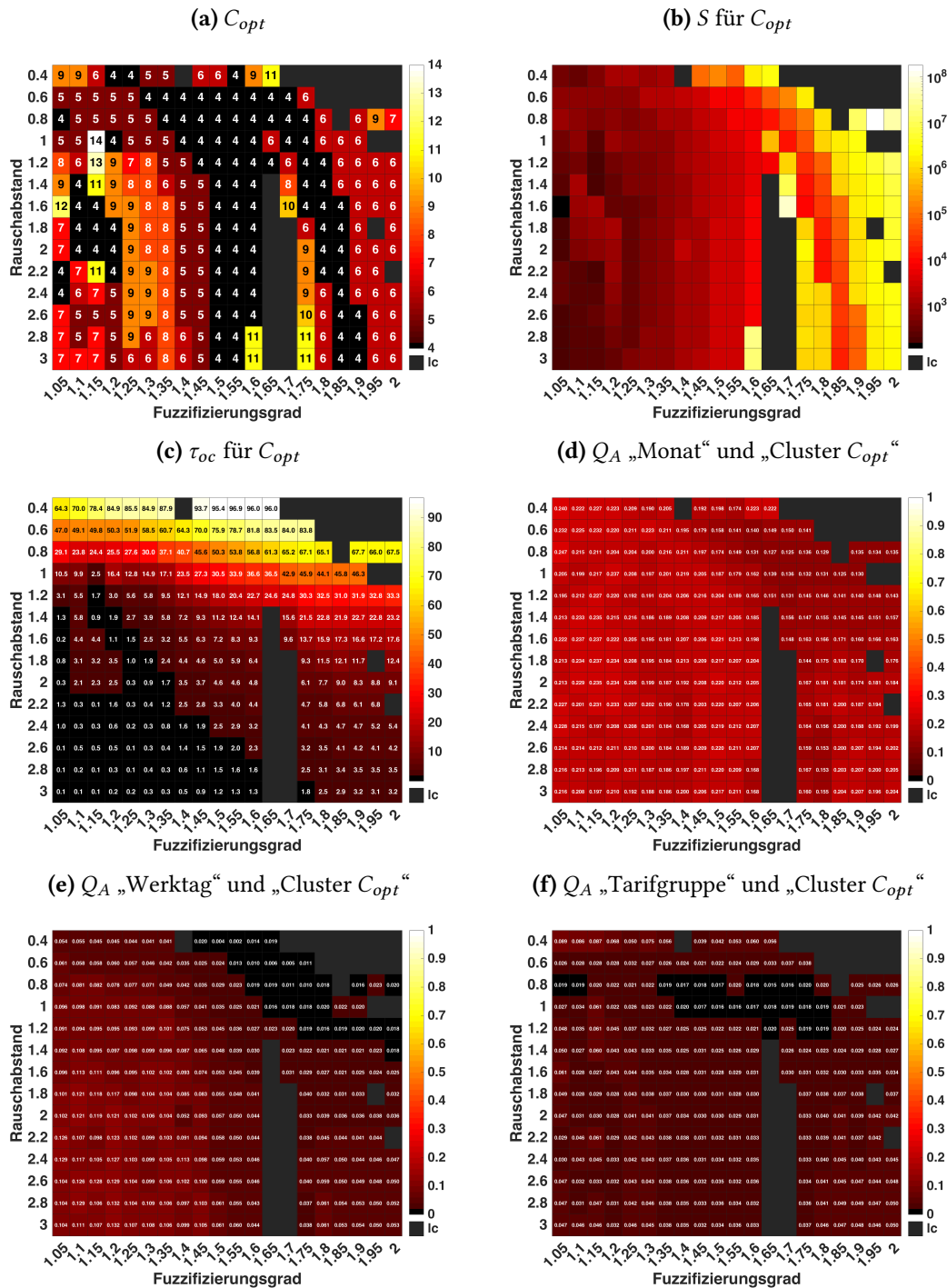


Abbildung B.19: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DT_{normmean} des OPP-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(f) relative Transinformation pro Ausgangsentropie). Für Fall DT siehe Abb. B.17, für Fall DT_{norm} siehe Abb. B.18. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.1.9 Bestimmung von optimalen Clusterparametern für Fall DT^{wo}

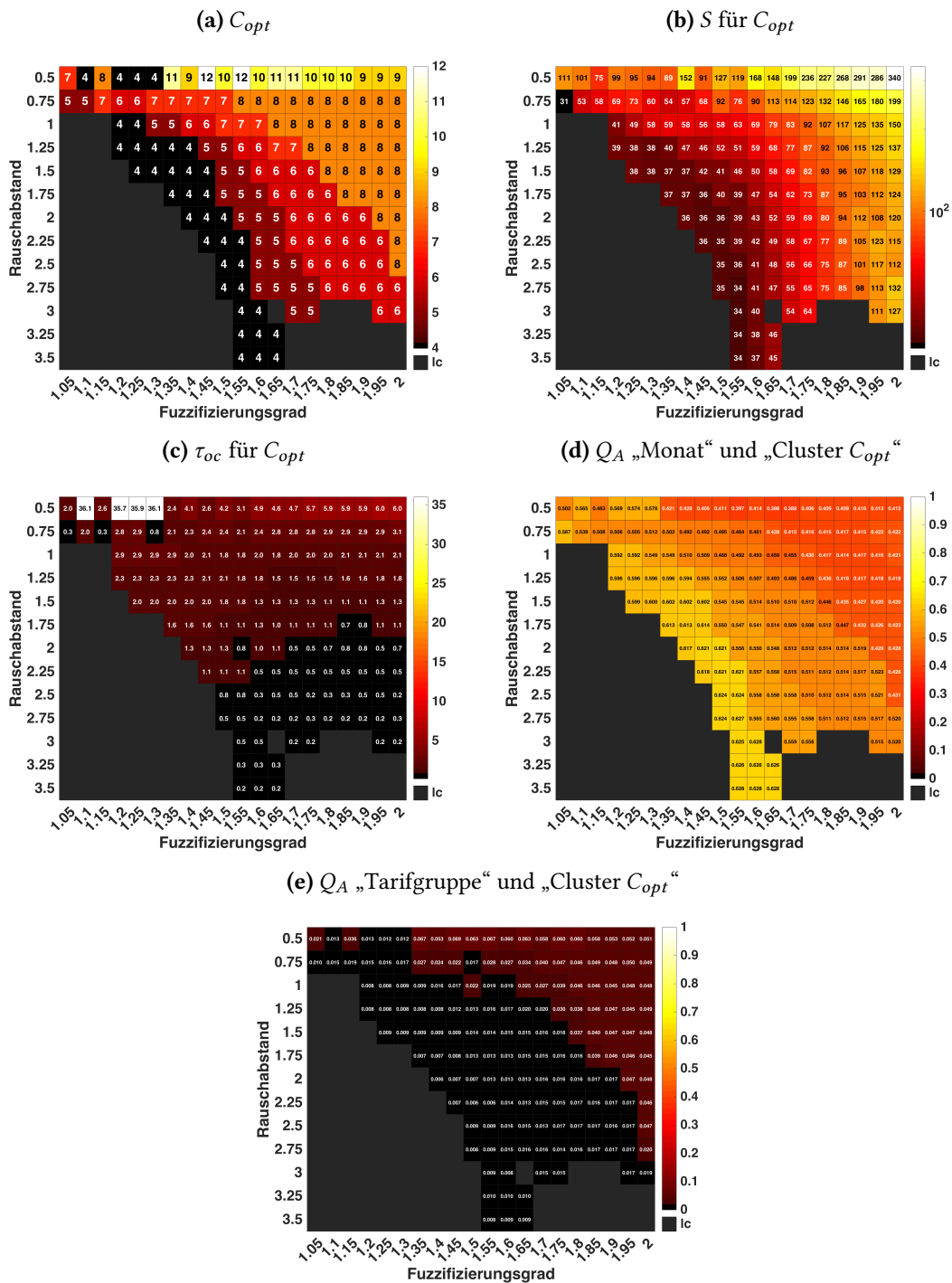


Abbildung B.20: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DT^{wo} des OPP-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(e) relative Transformation pro Ausgangsentropie). Für Fall DT^{wo}_{norm} siehe Abb. B.21, für Fall DT^{wo}_{normmean} siehe Abb. B.22. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.1.10 Bestimmung von optimalen Clusterparametern für Fall DT_{norm}^{wo}

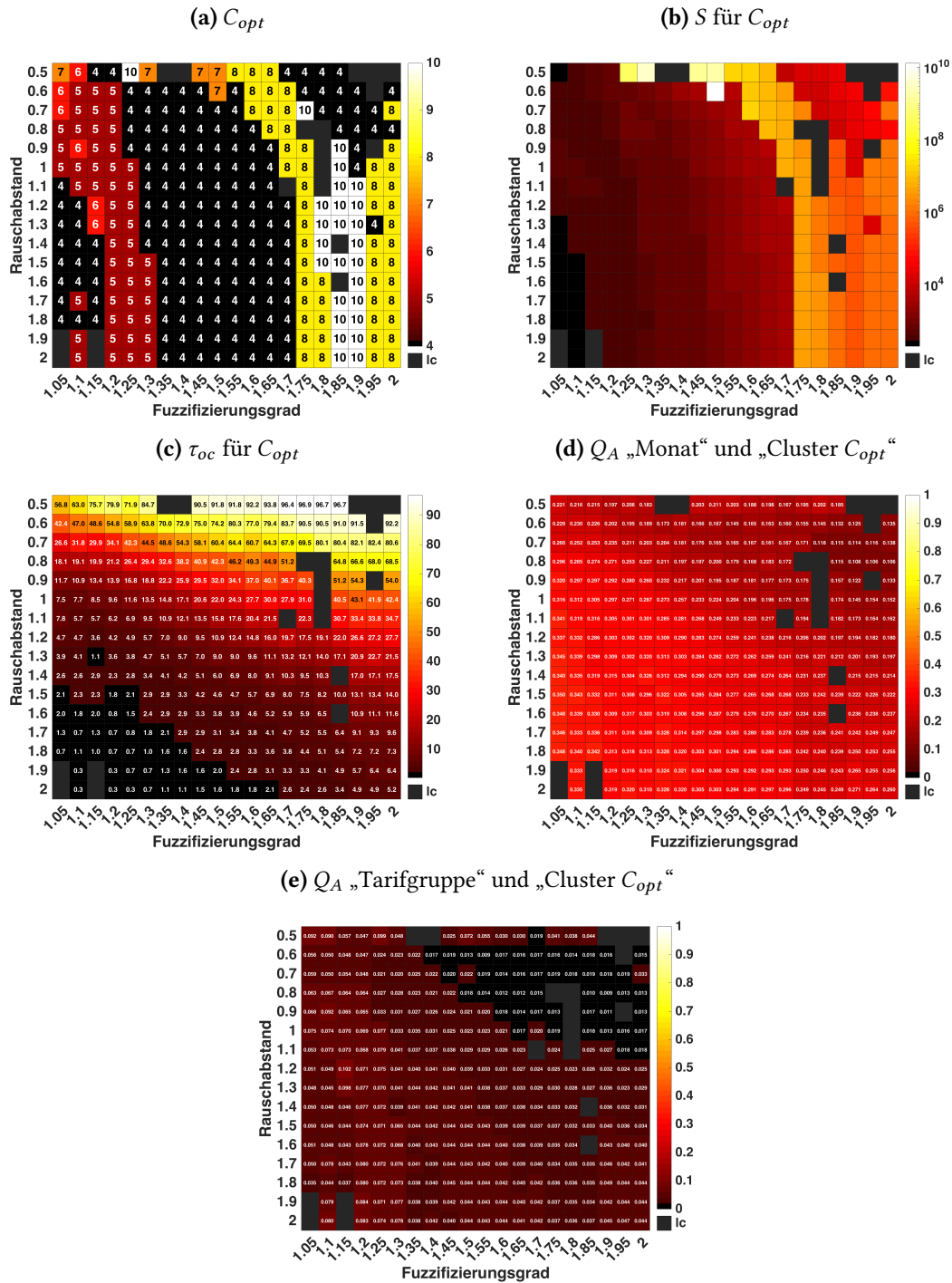


Abbildung B.21: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DT_{norm}^{wo} des OPP-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(e) relative Transinformation pro Ausgangsentropie). Für Fall DT_{norm}^{wo} siehe Abb. B.20, für Fall $DT_{norm}^{wo,mean}$ siehe Abb. B.22. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.1.11 Bestimmung von optimalen Clusterparametern für Fall $DT_{normmean}^{wo}$

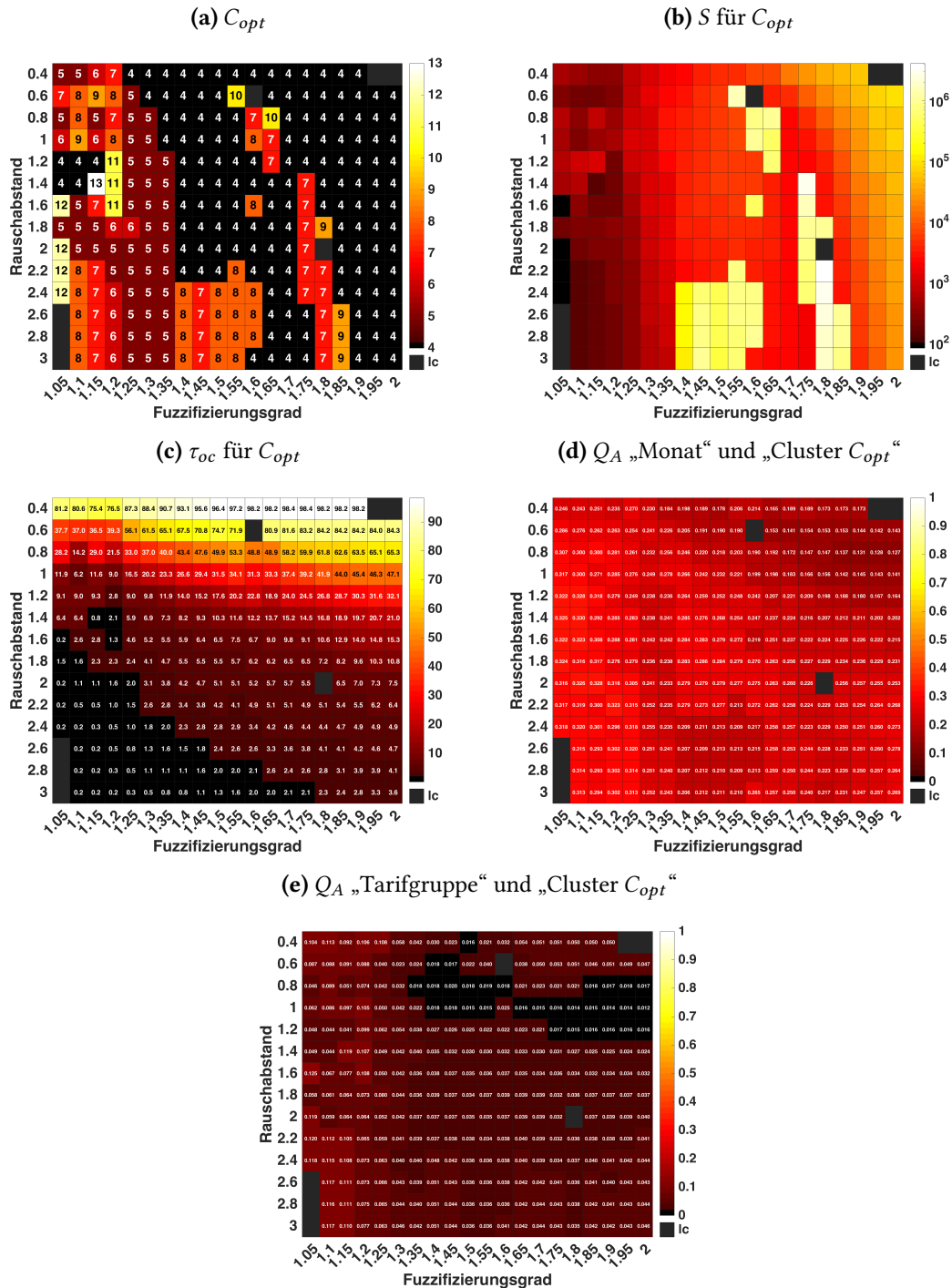


Abbildung B.22: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DT_{normmean}^{wo}$ des OPP-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(e) relative Transinformation pro Ausgangsentropie). Für Fall DT_{norm}^{wo} siehe Abb. B.20, für Fall DT_{norm}^{wo} siehe Abb. B.21. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.1.12 Bestimmung von optimalen Clusterparametern für Fall $DH_{TOU,Standard}^{wo}$

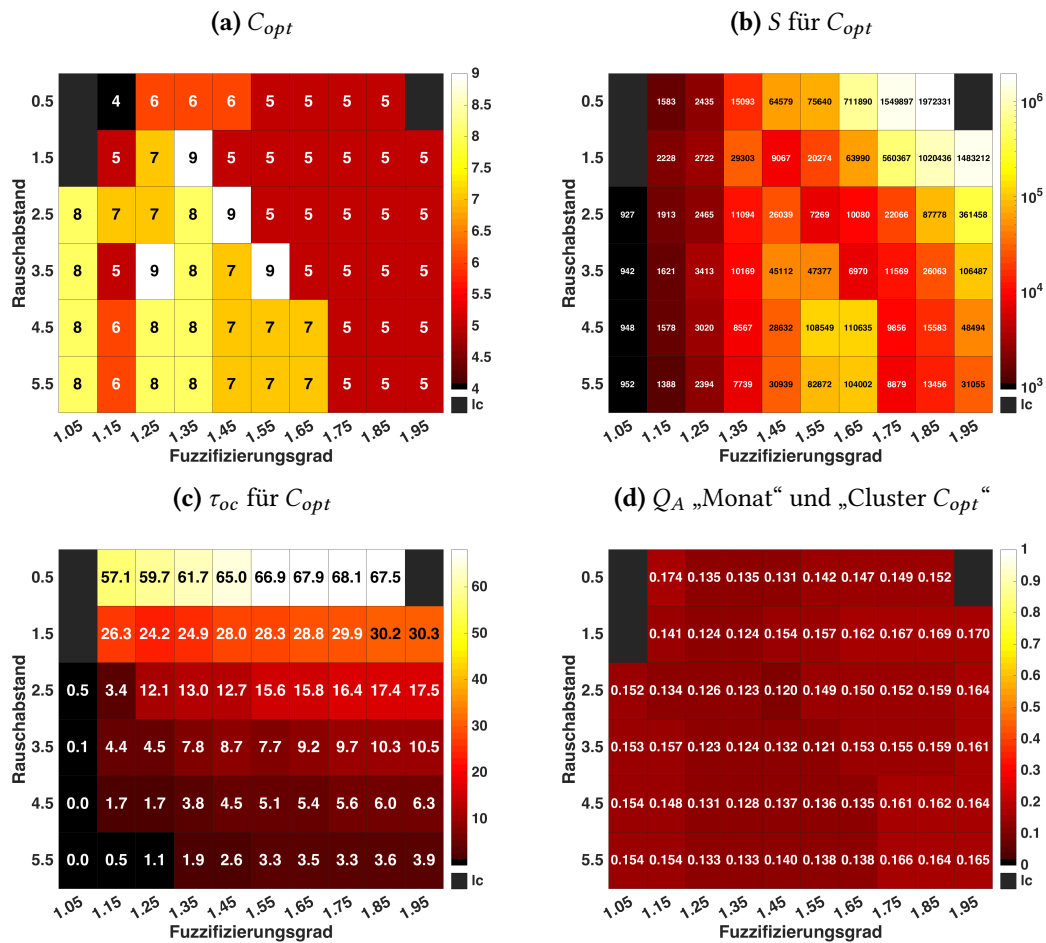


Abbildung B.23: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{TOU,Standard}^{wo}$ des OPP-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d) relative Transformation pro Ausgangsentropie). Für Fall $DH_{TOU,Standard,norm}^{wo}$ siehe Abb. B.24, für Fall $DH_{TOU,Standard,normmean}^{wo}$ siehe Abb. B.25. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.1.13 Bestimmung von optimalen Clusterparametern für Fall $DH_{TOU,Standard,norm}^{wo}$

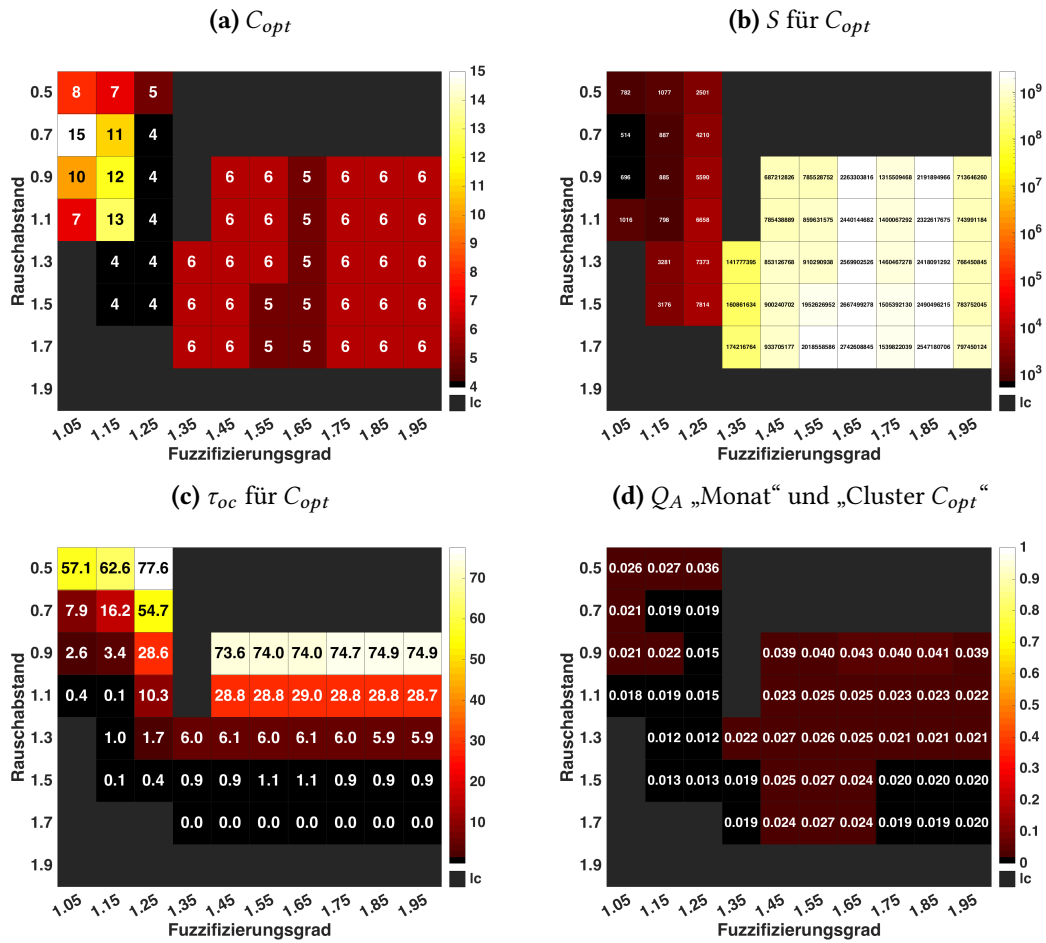


Abbildung B.24: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{TOU,Standard,norm}^{wo}$ des OPP-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d) relative Transformation pro Ausgangsentropie). Für Fall $DH_{TOU,Standard}^{wo}$ siehe Abb. B.23, für Fall $DH_{TOU,Standard,normmean}^{wo}$ siehe Abb. B.25. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.1.14 Bestimmung von optimalen Clusterparametern für Fall $DH_{TOU,Standard,normmean}^{wo}$

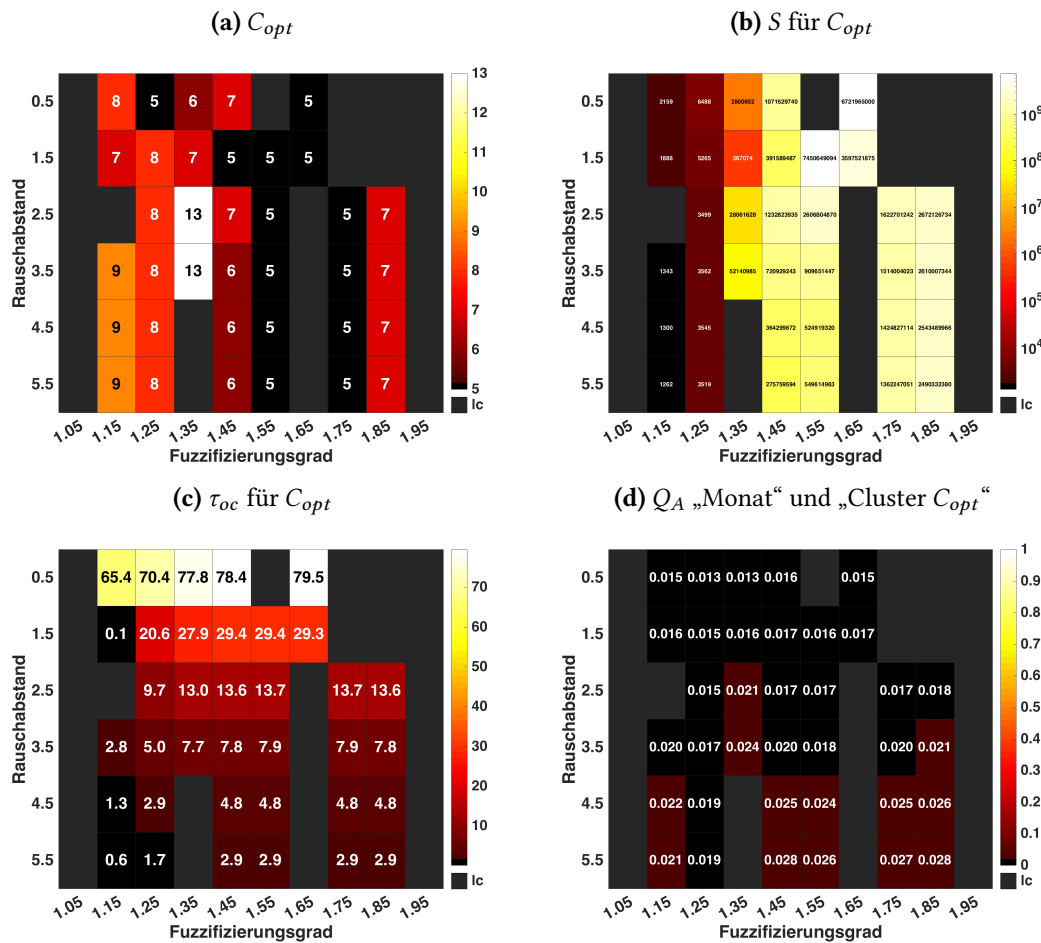


Abbildung B.25: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{TOU,Standard,normmean}^{wo}$ des OPP-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d) relative Transinformation pro Ausgangsentropie). Für Fall $DH_{TOU,Standard}^{wo}$ siehe Abb. B.23, für Fall $DH_{TOU,Standard,norm}^{wo}$ siehe Abb. B.24. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

Tabelle B.13: Ergebnis der Bestimmung von optimalen Clusterparametern für Fälle $DH_{TOU,Standard}^{wo}$, $DH_{TOU,Standard,norm}^{wo}$ und $DH_{TOU,Standard,normmean}^{wo}$ des OPP-Datensatzes; der Stern in der letzten Spalte kennzeichnet eine Parameterkombination, die nicht in der entsprechenden Abbildung auftaucht, sondern in weiteren Clusterdurchläufen ermittelt wurde

Fall	Tage	opt. Clusteranzahl	opt. Fuzzifizierungsgrad	opt. Rauschabstand	Verweis
$DH_{TOU,Standard}^{wo}$	nur Werktage	$C_{opt} = 6$	$q_{opt} = 1.15$	$\delta_{opt} = 5.5$	Abb. B.23
$DH_{TOU,Standard,norm}^{wo}$	nur Werktage	$C_{opt} = 7$	$q_{opt} = 1.05$	$\delta_{opt} = 1.1$	Abb. B.24
$DH_{TOU,Standard,normmean}^{wo}$	nur Werktage	$C_{opt} = 5$	$q_{opt} = 1.75$	$\delta_{opt} = 11$	Abb. B.25*

B.4.1.15 Aufgabenstellung 4: Prognose des preisbeeinflussten Verbrauchsverhaltens

Tabelle B.14: Vergleich zwischen verschiedenen VS-Modellen (siehe (b)-(e)) und ARX-Modellen (siehe (a)) auf Basis der DT-Fälle aus Tabelle 4.6 (RTP-Haushalte sind ausgeschlossen); als Modellbewertungsmaße dienen der Korrelationskoeffizient r und der mittlere absolute Fehler $Q_{C,MAE}$, die im Format $r(Q_{C,MAE})$ für jede Kombination aus Lern- und Testdatensatz angegeben sind. Farbliche Hervorhebung: $r \geq 0.8$ dunkelgrün, $0.8 > r \geq 0.5$ grün, $0.5 > r \geq 0.3$ hellgrün.

Lerndatensatz	Testdatensatz					$\bar{r}(Q_{C,MAE})$
	Fall DT ^{wo} _{Sommer}	Fall DT ^{we} _{Sommer}	Fall DT ^{wo} _{Standard}	Fall DT ^{we} _{Standard}	Fall DT ^{wo} _{CPP}	
(a) ARX-Modelle						
Fall DT ^{wo} _{Sommer}	0.5534 (0.0681)	0.3865 (0.0919)	0.0341 (0.1510)	-0.1142 (0.1456)	0.1677 (0.3003)	0.2055 (0.1514)
Fall DT ^{we} _{Sommer}	-0.0984 (0.1271)	0.3429 (0.0791)	-0.5141 (0.2910)	-0.5533 (0.2908)	-0.0385 (0.4076)	-0.1723 (0.2391)
Fall DT ^{wo} _{Standard}	0.1593 (0.1222)	0.0344 (0.1195)	0.7810 (0.0715)	0.7730 (0.0711)	0.4085 (0.2661)	0.4312 (0.1301)
Fall DT ^{we} _{Standard}	0.0841 (0.1257)	-0.0702 (0.1209)	0.7728 (0.0876)	0.8216 (0.0723)	0.4603 (0.2394)	0.4137 (0.1292)
Fall DT ^{wo} _{CPP}	0.0445 (0.2235)	-0.0650 (0.2145)	0.4704 (0.2116)	0.4370 (0.2067)	0.5403 (0.2527)	0.2854 (0.2218)
(b) VSA-Modelle						
Fall DT ^{wo} _{Sommer}	0.2313 (0.0912)	0.3587 (0.0821)	-0.1134 (0.1156)	0.0840 (0.0991)	0.1043 (0.2730)	0.1330 (0.1322)
Fall DT ^{we} _{Sommer}	0.1891 (0.0977)	0.4331 (0.0792)	-0.1865 (0.1302)	0.0161 (0.1103)	-0.0066 (0.2875)	0.0890 (0.1410)
Fall DT ^{wo} _{Standard}	0.0198 (0.1156)	-0.0275 (0.1101)	0.7017 (0.0801)	0.6731 (0.0724)	0.4817 (0.2327)	0.3698 (0.1222)
Fall DT ^{we} _{Standard}	0.0802 (0.1105)	0.1192 (0.0976)	0.6447 (0.0858)	0.7365 (0.0694)	0.4537 (0.2374)	0.4069 (0.1201)
Fall DT ^{wo} _{CPP}	0.0980 (0.1430)	0.0731 (0.1330)	0.6641 (0.1107)	0.6933 (0.1026)	0.5621 (0.2260)	0.4181 (0.1431)
(c) VSA*-Modelle						
Fall DT ^{wo} _{Sommer}	0.3143 (0.0915)	0.5488 (0.0770)	-0.2647 (0.1218)	-0.3006 (0.1126)	-0.2034 (0.2778)	0.0189 (0.1361)
Fall DT ^{we} _{Sommer}	0.3170 (0.0963)	0.5619 (0.0775)	0.3113 (0.1024)	0.2247 (0.0943)	0.1971 (0.2665)	0.3224 (0.1274)
Fall DT ^{wo} _{Standard}	-0.0386 (0.1039)	-0.0066 (0.0966)	0.6974 (0.0805)	0.6918 (0.0709)	0.4201 (0.2411)	0.3528 (0.1186)
Fall DT ^{we} _{Standard}	-0.0575 (0.1077)	0.0226 (0.0961)	0.6903 (0.0795)	0.7335 (0.0681)	0.5125 (0.2352)	0.3803 (0.1173)
Fall DT ^{wo} _{CPP}	-0.0162 (0.1244)	0.0252 (0.1138)	0.6837 (0.0985)	0.7168 (0.0830)	0.5710 (0.2184)	0.3961 (0.1276)
(d) VSB-Modelle						
Fall DT ^{wo} _{Sommer}	0.0720 (0.1402)	-0.0206 (0.1269)	-0.0429 (0.1832)	0.0048 (0.1688)	0.2121 (0.2947)	0.0451 (0.1828)
Fall DT ^{we} _{Sommer}	0.0699 (0.1790)	-0.0073 (0.1659)	-0.0230 (0.2620)	0.0308 (0.2386)	0.2128 (0.3479)	0.0566 (0.2387)
Fall DT ^{wo} _{Standard}	-0.1119 (0.1182)	-0.0535 (0.1073)	0.7300 (0.0744)	0.7650 (0.0643)	0.5104 (0.2362)	0.3680 (0.1201)
Fall DT ^{we} _{Standard}	-0.0588 (0.1150)	-0.0911 (0.1070)	0.6658 (0.0886)	0.7504 (0.0718)	0.4327 (0.2371)	0.3398 (0.1239)
Fall DT ^{wo} _{CPP}	-0.0398 (0.1180)	-0.1018 (0.1096)	0.6623 (0.0990)	0.7202 (0.0857)	0.6212 (0.2057)	0.3724 (0.1236)
(e) VSB*-Modelle						
Fall DT ^{wo} _{Sommer}	0.0182 (0.1295)	-0.0628 (0.1158)	0.0634 (0.1527)	0.0701 (0.1374)	0.0678 (0.3189)	0.0313 (0.1709)
Fall DT ^{we} _{Sommer}	0.0178 (0.1571)	-0.0640 (0.1412)	0.0751 (0.1971)	0.0801 (0.1800)	0.0813 (0.3985)	0.0381 (0.2148)
Fall DT ^{wo} _{Standard}	-0.0414 (0.1047)	-0.1108 (0.0949)	0.7247 (0.0793)	0.7236 (0.0654)	0.4663 (0.2313)	0.3525 (0.1151)
Fall DT ^{we} _{Standard}	0.0083 (0.1174)	-0.0823 (0.1121)	0.4779 (0.1164)	0.4751 (0.1034)	0.3737 (0.2576)	0.2505 (0.1414)
Fall DT ^{wo} _{CPP}	-0.0574 (0.0991)	-0.1012 (0.0902)	0.6902 (0.0937)	0.6778 (0.0802)	0.5009 (0.2421)	0.3421 (0.1211)

B.4.2 RESIDENS-Datensatz

Die Struktur des RESIDENS-Datensatzes ist in Tabelle B.15 zusammengefasst.

Tabelle B.15: Struktur des RESIDENS-Datensatzes gemäß Tabelle 2.3

Aggregations-niveau	Daten-tupel (N_D)	Zeitreihen-länge (K)	Zeitreihen (s_z), Einzelmerkmale (s)	Abtast-perioden-dauer (T_S) in Stunden
Fall GH	189	37248	$s_z = 2$ ($p[k]$ in ct/kWh, $P_n[k]$ in kW)	0.25
Fall GT	1	37248	$s_z = 2$ ($p[k]$ in ct/kWh, $P_n[k]$ in kW)	0.25
Fall DH	69169	96	$s_z = 2$ ($p[k]$ in ct/kWh, $P_n[k]$ in kW)	0.25
Fall DH ^{wo}	38918	96	$s_z = 2$ ($p[k]$ in ct/kWh, $P_n[k]$ in kW)	0.25
Fall DT	364	96	$s_z = 2$ ($p[k]$ in ct/kWh, $P_n[k]$ in kW)	0.25
Fall DT ^{wo}	207	96	$s_z = 2$ ($p[k]$ in ct/kWh, $P_n[k]$ in kW)	0.25
Fall SH	6640224	–	$s = 2$ (p in ct/kWh, P_n in kW)	–
Fall SH ^{wo}	3736128	–	$s = 2$ (p in ct/kWh, P_n in kW)	–
Fall ST	34944	–	$s = 2$ (p in ct/kWh, P_n in kW)	–
Fall ST ^{wo}	19872	–	$s = 2$ (p in ct/kWh, P_n in kW)	–

B.4.2.1 Bestimmung von optimalen Clusterparametern für Fall DH

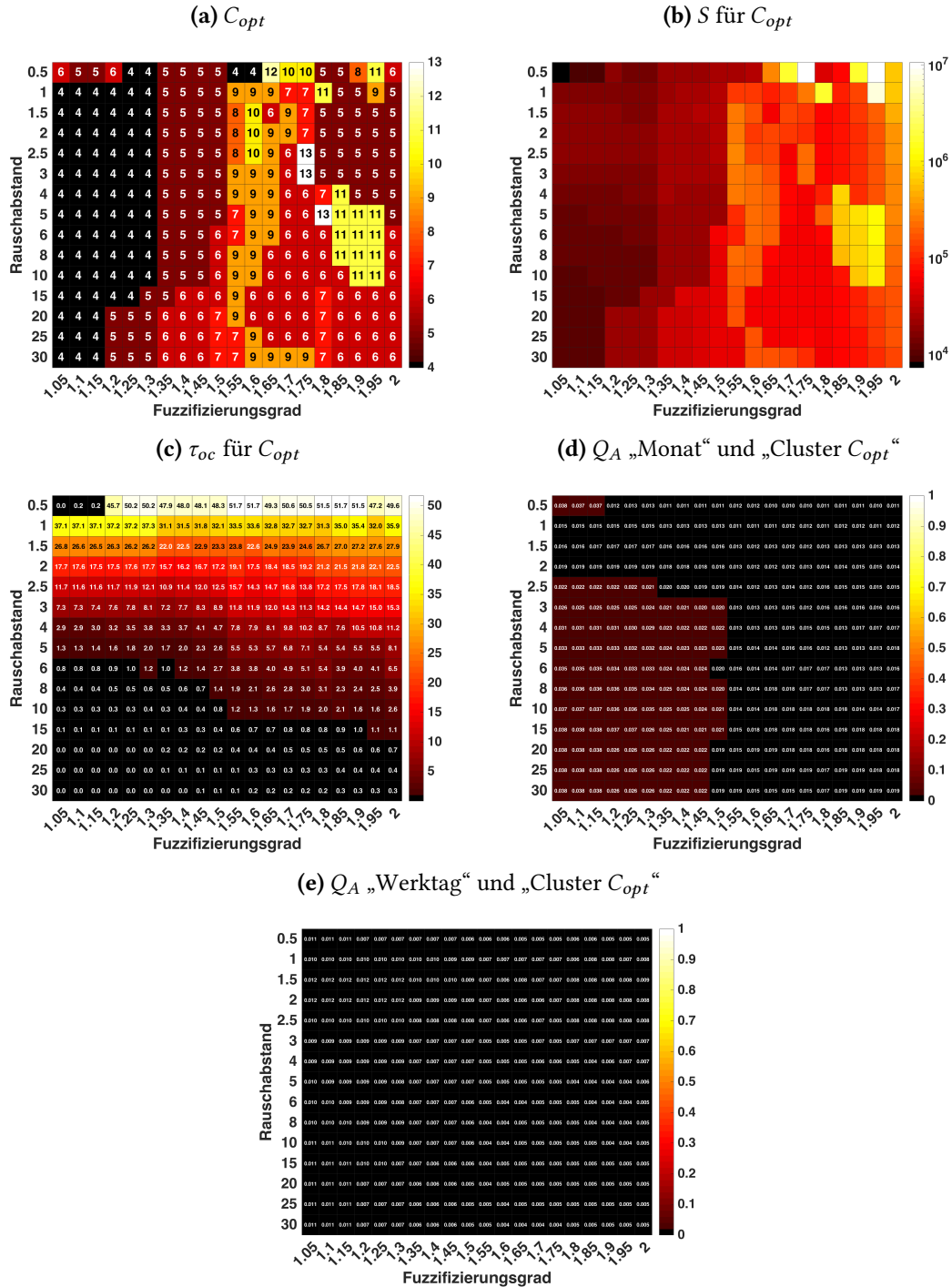


Abbildung B.26: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH des RESIDENS-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(e) relative Transinformation pro Ausgangsentropie). Für Fall DH_{norm} siehe Abb. B.27, für Fall $DH_{normmean}$ siehe Abb. B.28. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.2.2 Bestimmung von optimalen Clusterparametern für Fall DH_{norm}

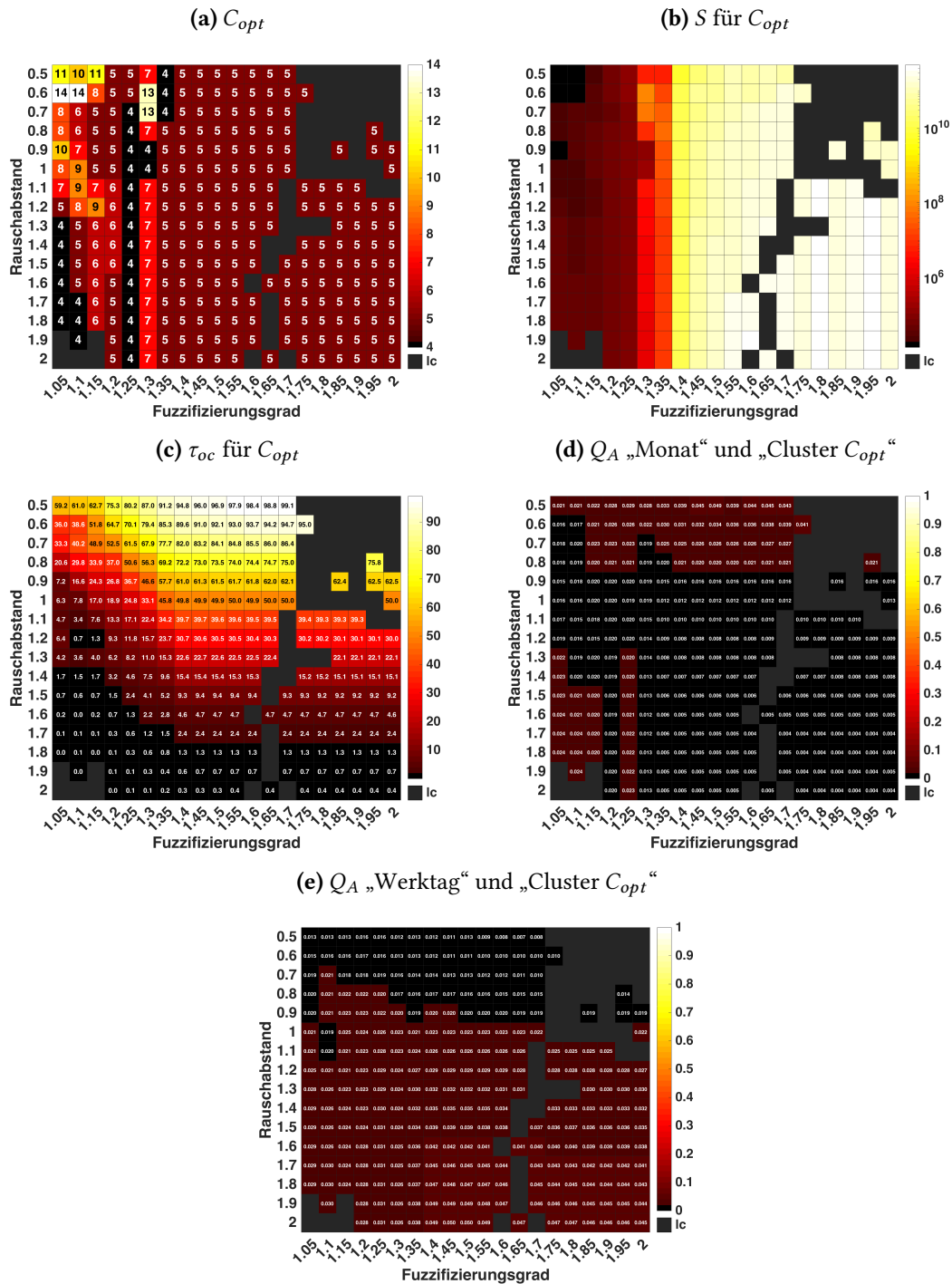


Abbildung B.27: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH_{norm} des RESIDENS-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(e) relative Transinformation pro Ausgangsentropie). Für Fall DH siehe Abb. B.26, für Fall $DH_{normmean}$ siehe Abb. B.28. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.2.3 Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}$

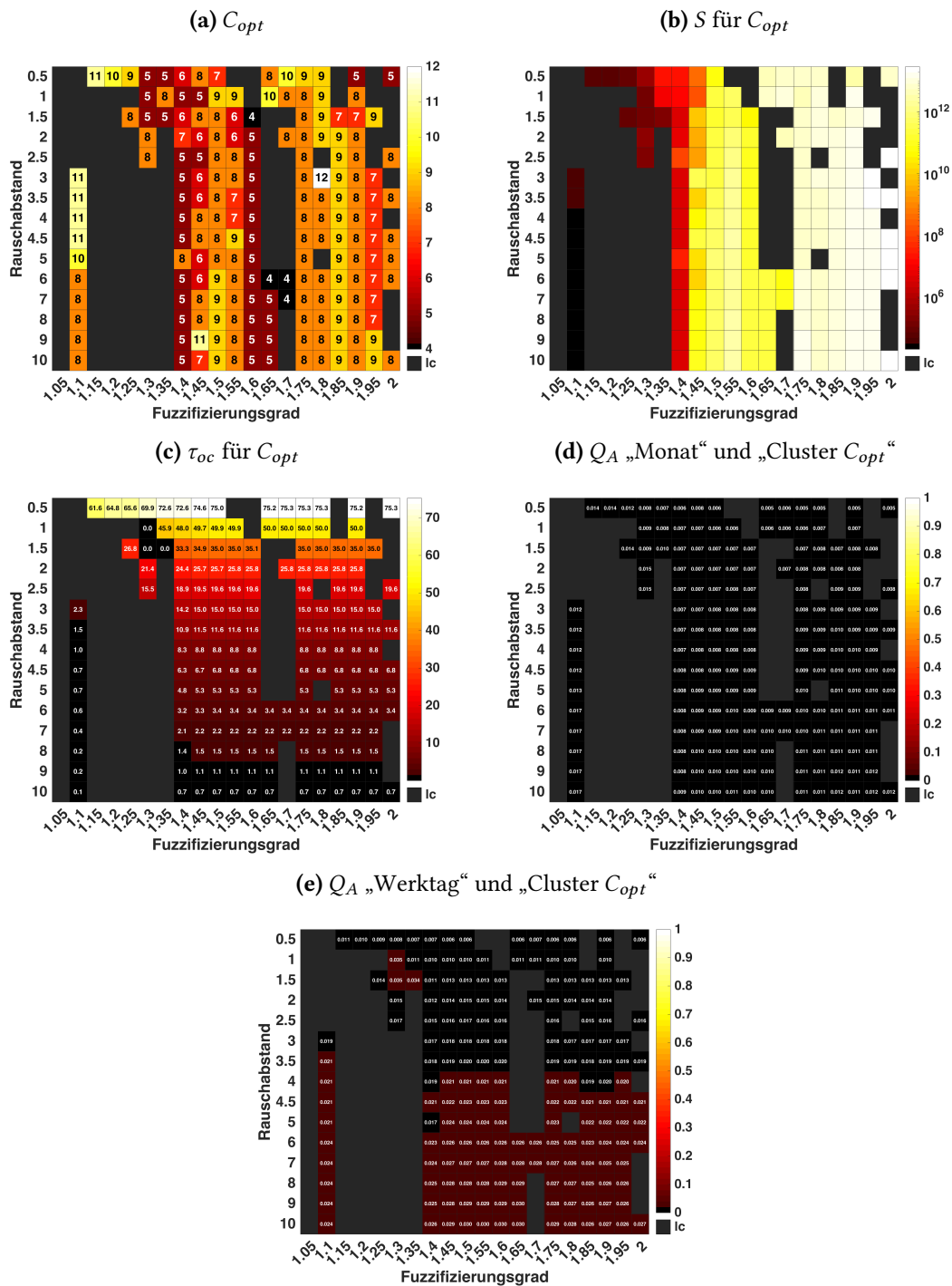


Abbildung B.28: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}$ des RESIDENS-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d)-(e) relative Transformation pro Ausgangsentropie). Für Fall DH siehe Abb. B.26, für Fall DH_{norm} siehe Abb. B.27. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.2.4 Bestimmung von optimalen Clusterparametern für Fall DH^{wo}

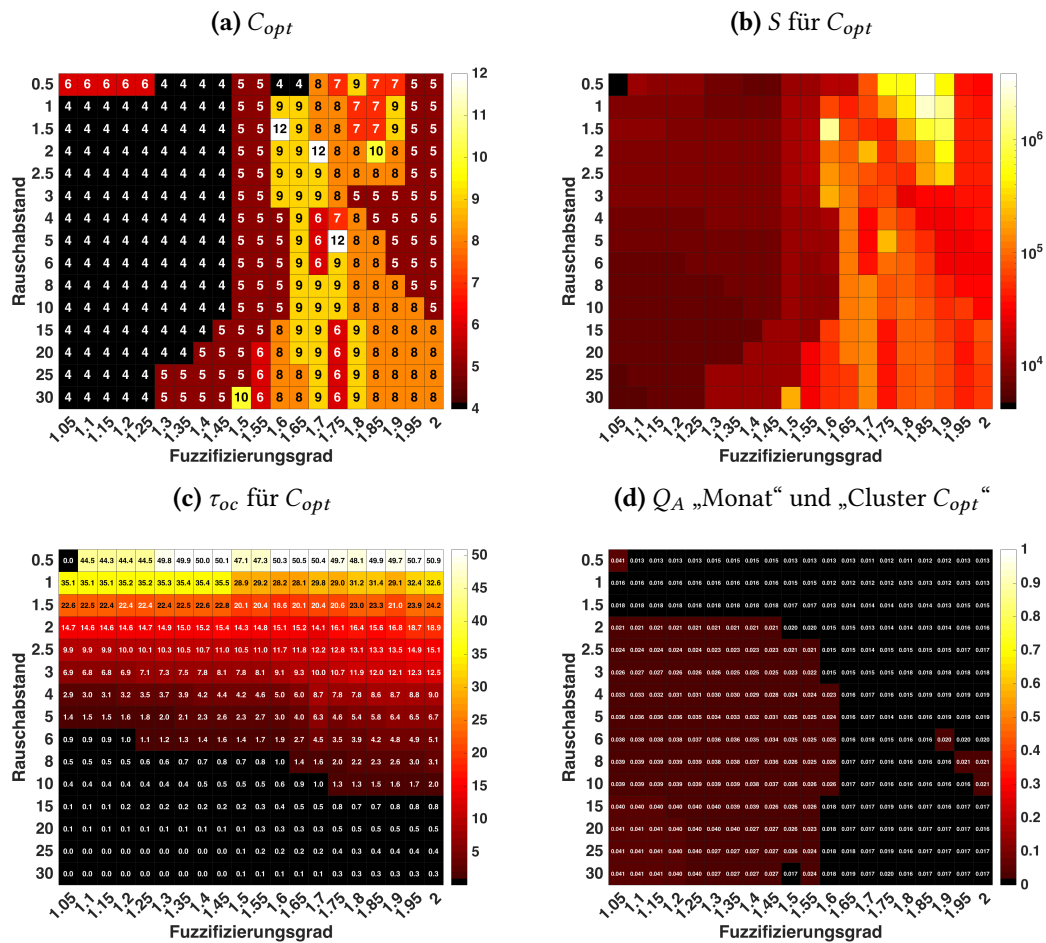


Abbildung B.29: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH^{wo} des RESIDENS-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d) relative Transinformation pro Ausgangsentropie). Für Fall DH^{wo}_{norm} siehe Abb. B.30, für Fall $DH^{wo}_{normmean}$ siehe Abb. B.31. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.2.5 Bestimmung von optimalen Clusterparametern für Fall DH_{norm}^{wo}

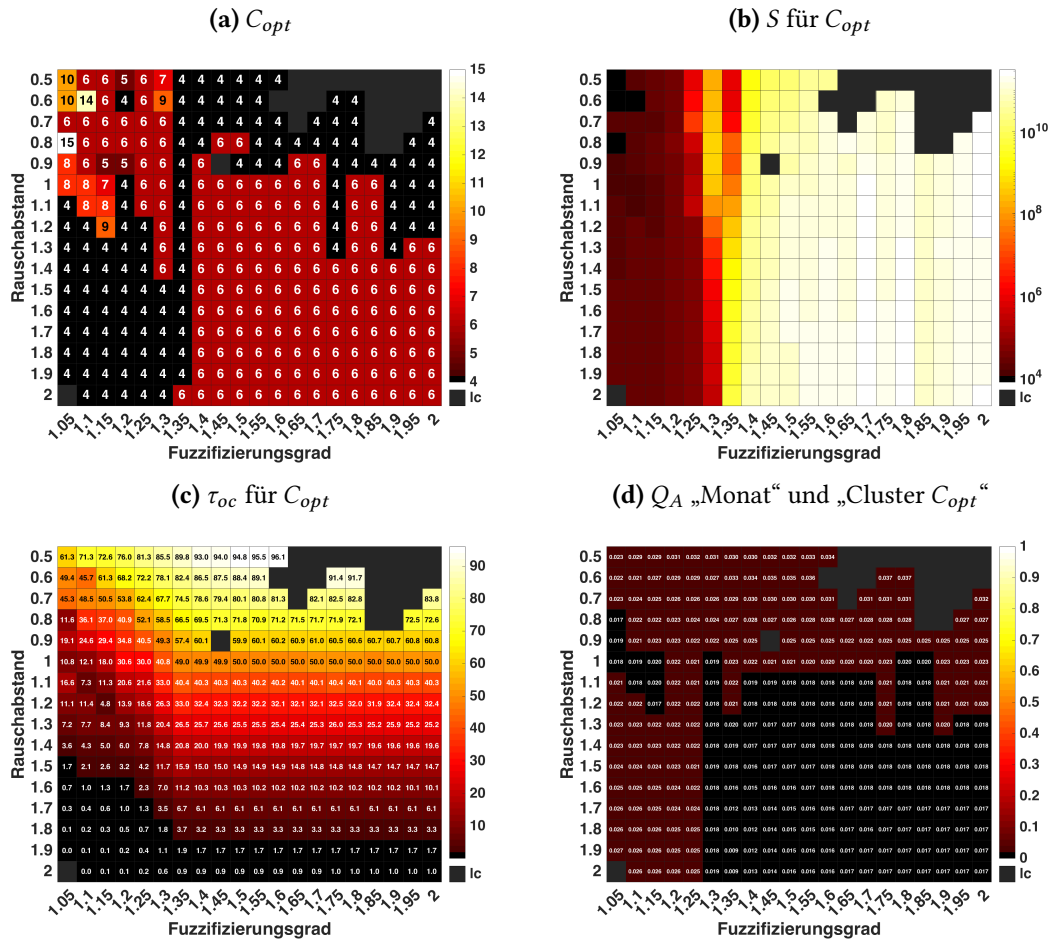


Abbildung B.30: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH_{norm}^{wo} des RESIDENS-Datensatzes ((a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d) relative Transformation pro Ausgangsentropie). Für Fall DH_{norm}^{wo} siehe Abb. B.29, für Fall $DH_{norm}^{wo,mean}$ siehe Abb. B.31. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

B.4.2.6 Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}^{wo}$

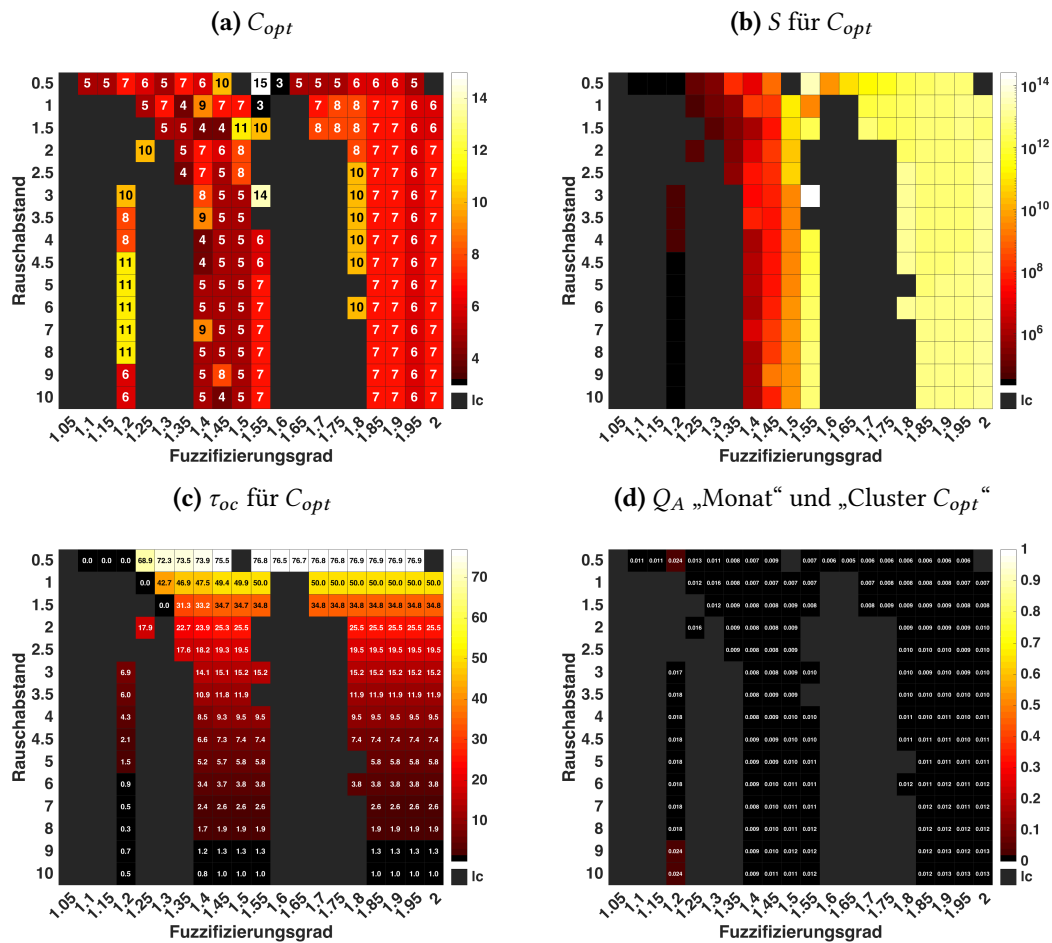


Abbildung B.31: Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}^{wo}$ des RESIDENS-Datensatzes (a) optimale Clusteranzahl, (b) Trennungsgrad, (c) prozentualer Anteil der Datentupel im Rauschcluster, (d) relative Transformation pro Ausgangsentropie). Für Fall DH^{wo} siehe Abb. B.29, für Fall DH_{norm}^{wo} siehe Abb. B.30. Die grauen Zellen kennzeichnen diejenigen δ, q -Parameterkombinationen, für die mindestens ein Cluster keine Datentupel (lc - leere Cluster) enthält.

Tabelle B.16: Verbundverteilung der Ausgangsgrößen „DH-Cluster“ und „Monat“ des RESIDENS-Datensatzes als Tabelle mit der Anzahl an Datentupeln und der prozentualen Verteilung der Datentupel je Zeile

Cluster	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Cluster 1	3420 (6.6%)	3269 (6.3%)	3826 (7.4%)	4386 (8.4%)	4861 (9.3%)	4878 (9.4%)	4867 (9.4%)	4847 (9.3%)	4549 (8.7%)	4330 (8.3%)	5718 (11.0%)	3050 (5.9%)
Cluster 2	1845 (10.9%)	1645 (9.7%)	1526 (9.0%)	1144 (6.7%)	864 (5.1%)	753 (4.4%)	950 (5.6%)	939 (5.5%)	1056 (6.2%)	1451 (8.6%)	2706 (16.0%)	2083 (12.3%)
Cluster 3	15 (45.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	18 (54.5%)
Rauschcluster	29 (16.8%)	11 (6.4%)	12 (6.9%)	10 (5.8%)	15 (8.7%)	6 (3.5%)	4 (2.3%)	9 (5.2%)	5 (2.9%)	16 (9.2%)	28 (16.2%)	28 (16.2%)

C Abbildungsverzeichnis

1.1	Erneuerbare Energien in den Bereichen Strom, Wärme/Kälte und Verkehr	2
1.2	Bruttostromerzeugung in Deutschland im Jahr 2016	3
1.3	Maßnahmen zur Synchronisation von Stromerzeugung und -verbrauch	4
1.4	Unterteilung und Kategorisierung von <i>Demand Side Management</i>	8
1.5	Typische Verläufe des Preissignals bei <i>Demand-Response</i> -Programmen	9
1.6	Zusammenhang zwischen Spitzenlastreduktion und Strompreisspreizung in experimentellen <i>Demand-Response</i> -Studien	10
2.1	Struktur der Methodenpipeline des neuen Konzeptes zur datengetriebenen Analyse und Modellierung des preisbeeinflussten Verbrauchsverhaltens	26
2.2	Systemtechnische Einordnung des <i>Demand-Response</i> -Prozesses	28
2.3	Datenstrukturen in einem <i>Smart-Meter</i> -Datensatz	37
2.4	Synthetische Strompreisprofile des <i>Demand-Response</i> -Simulators für die Generierung eines Benchmarkdatensatzes	40
2.5	Flussdiagramm für die Generierung eines Benchmarkdatensatzes	46
2.6	Vergleich verschiedener Imputationsverfahren bei fehlerbehafteten Lastzeitreihen für kein Bereich und fünf Bereiche mit fehlenden Werten	49
2.7	Problematik von ungleichmäßig verteilten fehlenden Werten bei der Datenvorverarbeitung und Zeitreihenzerlegung	54
2.8	Zuordnung der Tageskurven zu den Clustern	65
2.9	Mittlere Clusterkurven	65
2.10	Clusterzugehörigkeiten der Datentupel zu den Clustern	65
2.11	Wahl der Startcluster aus Fuzzy-Menge	67
2.12	Informationsgehalt einer Merkmalsausprägung in Abhängigkeit von der Auftrittswahrscheinlichkeit und Entropie einer Binärquelle	70
2.13	Informationsflussdiagramm und Venn-Diagramm zur Darstellung der Abhängigkeit von Eingangsentropie, Ausgangsentropie, Verbundentropie, Transinformation, Irrelevanz und Äquivokation	71
2.14	Mittlere Clusterkurven	72
2.15	Balkendiagramm und Informationsflussdiagramm für Ausgangsgrößen „Cluster 6“ und „Monat“	73
2.16	Balkendiagramm und Informationsflussdiagramm für Ausgangsgrößen „Cluster 6“ und „Wochentag“	73
2.17	Ablauf der Sensitivitätsanalysen	75
2.18	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH des Benchmarkdatensatzes „Settlement210H365D_70M20S“ (3D-Oberflächengrafiken)	79

2.19	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH des Benchmarkdatensatzes „Settlement210H365D_70M20S“ (Heatmaps) . . .	80
2.20	Sensitivität der Clusterergebnisse bezüglich der Lastsignalcharakteristik für DH-, DH_{norm} - und $DH_{normmean}$ -Fälle	82
2.21	Sensitivität der Clusterergebnisse bezüglich der Wahl der Startcluster für den DH-Fall des Benchmarkdatensatzes „Settlement210H365D_70M20S“	85
2.22	Lineare und exponentielle Speicherausgleichsfunktion	97
2.23	Modellverhalten verschiedener Virtueller-Speicher-Modelle als Reaktion auf synthetische Strompreise	99
2.24	Entwurfs- und -anwendungsphase bei der Modellierung des preisbeeinflussten Verbrauchsverhaltens	101
3.1	Exemplarischer Screenshot der Menü- und Bedienelemente in der SciXMiner-Erweiterung <i>Energy</i>	112
3.2	Exemplarischer Screenshot der Bedienelemente in der SciXMiner-Erweiterung <i>Energy</i> zur Auswahl und Spezifikation des VSA-Modells	115
4.1	Strompreise der Tarifgruppen des <i>Olympic Peninsula Project</i>	118
4.2	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH des OPP-Datensatzes (Heatmaps)	121
4.3	Verbundverteilungen zweier Ausgangsgrößen des OPP-Datensatzes als Balkendiagramme	123
4.4	Verbundverteilungen zweier Ausgangsgrößen des OPP-Datensatzes als Balkendiagramme	125
4.5	Gefundene DH-Cluster des OPP-Datensatzes	126
4.6	Gefundene DT-Cluster des OPP-Datensatzes	127
4.7	Verteilung der Typtage für Fall $DT_{normmean}^{wo}$ des OPP-Datensatzes mit farbkodiertem normierten Verbrauch	128
4.8	Verbundverteilung der Ausgangsgrößen „ $DT_{normmean}^{wo}$ -Cluster“ und „Tarifgruppe“ des OPP-Datensatzes	129
4.9	Gefundene $DH_{TOU,Standard}^{wo}$ -Cluster des OPP-Datensatzes	131
4.10	Modellvergleich verschiedener VS-Modelle mit ARX-Modellen auf Basis der VT-Fälle des OPP-Datensatzes	135
4.11	Modellvergleich verschiedener VS-Modelle mit ARX-Modellen auf Basis der DT-Fälle des OPP-Datensatzes	136
4.12	Strompreisstufen des RESIDENS-Projektes	138
4.13	Verbundverteilung der Ausgangsgrößen „ DH_{norm} -Cluster“ und „Werktag“ des RESIDENS-Datensatzes	140
4.14	Gefundene DH-Cluster des RESIDENS-Datensatzes	142
B.1	Einfluss des Preises auf die Änderung der Stromnachfrage	175
B.2	Wärmegestehungskosten verschiedener Heiztechnologien in Einfamilienhäusern	179
B.3	Vergleich verschiedener Imputationsverfahren bei fehlerbehafteten Lastzeitreihen für ein Bereich und zwei Bereiche mit fehlenden Werten	192

B.4	Vergleich verschiedener Imputationsverfahren bei fehlerbehafteten Lastzeitreihen für drei und vier Bereiche mit fehlenden Werten	193
B.5	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH_{norm} des Benchmarkdatensatzes „Settlement210H365D_70M20S“ (3D-Oberflächengrafiken)	198
B.6	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH_{norm} des Benchmarkdatensatzes „Settlement210H365D_70M20S“ (Heatmaps)	199
B.7	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}$ des Benchmarkdatensatzes „Settlement210H365D_70M20S“ (3D-Oberflächengrafiken)	200
B.8	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}$ des Benchmarkdatensatzes „Settlement210H365D_70M20S“ (Heatmaps)	201
B.9	Sensitivität der Clusterergebnisse bezüglich der Lastsignalcharakteristik für DT-, DT_{norm} - und $DT_{normmean}$ -Fälle	202
B.10	Sensitivität der Clusterergebnisse bezüglich der Wahl der Startcluster für den DH_{norm} -Fall des Benchmarkdatensatzes „Settlement210H365D_70M20S“	203
B.11	Sensitivität der Clusterergebnisse bezüglich der Wahl der Startcluster für den $DH_{normmean}$ -Fall des Benchmarkdatensatzes „Settlement210H365D_70M20S“	204
B.12	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH_{norm} des OPP-Datensatzes (Heatmaps)	210
B.13	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}$ des OPP-Datensatzes (Heatmaps)	211
B.14	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH^{wo} des OPP-Datensatzes (Heatmaps)	212
B.15	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH_{norm}^{wo} des OPP-Datensatzes (Heatmaps)	213
B.16	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}^{wo}$ des OPP-Datensatzes (Heatmaps)	214
B.17	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DT des OPP-Datensatzes (Heatmaps)	215
B.18	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DT_{norm} des OPP-Datensatzes (Heatmaps)	216
B.19	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DT_{normmean}$ des OPP-Datensatzes (Heatmaps)	217
B.20	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DT^{wo} des OPP-Datensatzes (Heatmaps)	218
B.21	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DT_{norm}^{wo} des OPP-Datensatzes (Heatmaps)	219
B.22	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DT_{normmean}^{wo}$ des OPP-Datensatzes (Heatmaps)	220
B.23	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{TOU,Standard}^{wo}$ des OPP-Datensatzes (Heatmaps)	221
B.24	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{TOU,Standard,norm}^{wo}$ des OPP-Datensatzes (Heatmaps)	222

B.25	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{TOU,Standard,normmean}^{wo}$ des OPP-Datensatzes (Heatmaps)	223
B.26	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH des RESIDENS-Datensatzes (Heatmaps)	226
B.27	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH_{norm} des RESIDENS-Datensatzes (Heatmaps)	227
B.28	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}$ des RESIDENS-Datensatzes (Heatmaps)	228
B.29	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH^{wo} des RESIDENS-Datensatzes (Heatmaps)	229
B.30	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall DH_{norm}^{wo} des RESIDENS-Datensatzes (Heatmaps)	230
B.31	Sensitivitätsanalyse zur Bestimmung von optimalen Clusterparametern für Fall $DH_{normmean}^{wo}$ des RESIDENS-Datensatzes (Heatmaps)	231

D Tabellenverzeichnis

1.1	Überblick über verschiedene anreizorientierte Programme	7
1.2	Überblick über verschiedene preisorientierte Programme	12
1.3	Übersicht über Imputationsverfahren	14
1.4	Übersicht zum Forschungsstand auf dem Gebiet der <i>Demand-Response-Model-</i> <i>lierung</i>	22
2.1	Checkliste für <i>Smart-Meter</i> -Datensätze (Teil 1)	33
2.2	Checkliste für <i>Smart-Meter</i> -Datensätze (Teil 2)	34
2.3	Strukturierung und Einteilung von <i>Smart-Meter</i> -Datensätzen	59
2.4	Konfiguration der Sensitivitätsanalyse zur Bestimmung von optimalen Cluster- parametern	76
2.5	Konfiguration der Sensitivitätsanalyse zur Bestimmung des Einflusses der Last- signalcharakteristik	81
2.6	Konfiguration der Sensitivitätsanalyse zur Bestimmung des Einflusses der in- itiales Startcluster	84
2.7	Fall $SH_{normmean}$: Auswertung der Preisklassen für den TOU2-Tarif gegenüber der Kontrollgruppe	87
2.8	Fall $SH_{normmean}$: Auswertung der Preisklassen für den TOU3-Tarif gegenüber der Kontrollgruppe	87
2.9	Bewertung der analyseorientierten Methoden der Pipeline	90
2.10	Berechnungsvorschriften für die vorwärts und rückwärts gerichtete Bewertung des Strompreises für Virtuelle-Speicher-Modelle	98
2.11	Zusammenfassung ausgewählter Modellbewertungsmaße	106
2.12	Realisierung von Modellannahmen und Erkenntnisse der datengetriebenen <i>Demand-Response-Analyse</i> als <i>a-priori</i> -Wissen im Virtuellen-Speicher-Modell .	107
3.1	Einzustellende Parameter des Benchmarkgenerators zur Erzeugung von Bench- markdatensätzen	111
4.1	Ergebnis der Bestimmung von optimalen Clusterparametern für den OPP- Datensatz	119
4.2	Untersuchte Fälle des OPP-Datensatzes: Auswertung der Clusteranalysen mit informationstheoretischen Maßen	123
4.3	Auswertung der Preisklassen der variablen Stromtarife gegenüber der Kontroll- gruppe für Fall $SH_{normmean}$ des OPP-Datensatzes	129
4.4	Einstufung der $DH_{TOU,Standard}^{wo}$ -Cluster in Responder-, Semi-Responder- und Non- Responder-Haushalte	132

4.5	Strukturierung und Einteilung des OPP-Datensatzes in VT-Fälle für die Aufgabenstellung A4	134
4.6	Strukturierung und Einteilung des OPP-Datensatzes in DT-Fälle für die Aufgabenstellung A4	135
4.7	Bewertung verschiedener <i>Demand-Response</i> -Modelltypen	137
4.8	Ergebnis der Bestimmung von optimalen Clusterparametern für den RESIDENS-Datensatz	139
4.9	Untersuchte Fälle des RESIDENS-Datensatzes: Auswertung der Clusteranalysen mit informationstheoretischen Maßen	141
B.1	Übersicht über <i>Smart-Grid</i> -Projekte in der Europäischen Union	163
B.2	Technisches Lastverlagerungspotenzial verschiedener Anwendungen im Haushaltsbereich	178
B.3	Fortsetzung der Checklisten (Tabellen 2.1 und 2.2) für <i>Smart-Meter</i> -Datensätze .	180
B.4	<i>Demand-Response</i> -Simulator: <i>Lookup</i> -Tabelle mit Preiselastizitäten	181
B.5	Parametereinstellungen des Benchmarkgenerators zur Erzeugung des Benchmarkdatensatzes zur systematischen Untersuchung verschiedener Imputationsverfahren	182
B.6	Parametereinstellungen des Benchmarkgenerators zur Erzeugung des Benchmarkdatensatzes zur Ermittlung von optimalen Hampel-Filterparametern	183
B.7	Parametereinstellungen des Benchmarkgenerators zur Erzeugung von Benchmarkdatensätzen zur systematischen Untersuchung der analyseorientierten Methoden der Pipeline	184
B.8	Ergebnis des Vergleichs verschiedener Imputationsverfahren bei fehlerbehafteten Lastzeitreihen	186
B.9	Ermittlung von optimalen Hampel-Filterparametern bei ausreißerbehafteten Lastzeitreihen	194
B.10	Normierungsmethoden bei Energiezeitreihen	196
B.11	Mögliche Ausgangsgrößen für unterschiedliche Aggregationen der Datensätze .	197
B.12	Struktur des OPP-Datensatzes	209
B.13	Ergebnis der Bestimmung von optimalen Clusterparametern für Fälle $DH_{TOU,Standard}^{wo}$, $DH_{TOU,Standard,norm}^{wo}$ und $DH_{TOU,Standard,normmean}^{wo}$ des OPP-Datensatzes	223
B.14	Vergleich zwischen verschiedenen VS-Modellen und ARX-Modellen auf Basis der DT-Fälle des OPP-Datensatzes	224
B.15	Struktur des RESIDENS-Datensatzes	225
B.16	Verbundverteilung der Ausgangsgrößen „DH-Cluster“ und „Monat“ des RESIDENS-Datensatzes	232

E Listings

3.1	Skript processDHClustering.m	113
3.2	Auszug aus der Funktion regression_en.m	114
B.1	Funktion load_benchmark_generator_settlement.m	205
B.2	Auszug aus dem Skript defineNames.m	206
B.3	Auszug aus der Funktion vsa_en.m	207
B.4	Auszug aus der Funktion vsa_an.m	208

F Literaturverzeichnis

- [1] BMU und BMWi: *Das Energiekonzept der Bundesregierung 2010 und die Energiewende 2011*. Technischer Bericht. Bundesministerium für Wirtschaft und Technologie (BMWi), Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit (BMU), 2011.
- [2] DENA: *dena-Netzstudie II - Integration erneuerbarer Energien in die deutsche Stromversorgung im Zeitraum 2015 - 2020 mit Ausblick 2025*. Technischer Bericht. Deutsche Energie-Agentur GmbH (dena), 2010.
- [3] BMWi: *Die Energie der Zukunft – Fünfter Monitoring-Bericht zur Energiewende, Berichtsjahr 2015*. Technischer Bericht. Bundesministerium für Wirtschaft und Energie (BMWi), 2016.
- [4] AGEE-STAT: *Zeitreihen zur Entwicklung der erneuerbaren Energien in Deutschland*. Technischer Bericht. Umweltbundesamt, Fachgebiet I 2.5 Energieversorgung und -daten, Geschäftsstelle der Arbeitsgruppe Erneuerbare Energien-Statistik (AGEE-Stat), 2017.
- [5] KLAUS, T.; VOLLMER, C.; WERNER, K.; LEHMANN, H. und MÜSCHEN, K.: *Energieziel 2050: 100% Strom aus erneuerbaren Quellen*. Technischer Bericht. Umweltbundesamt, Dessau, 2010.
- [6] BDEW: *Erneuerbare Energien und das EEG: Zahlen, Fakten, Grafiken (2016)*. Technischer Bericht. BDEW Bundesverband der Energie- und Wasserwirtschaft e.V., 2016.
- [7] BMWi: *Erneuerbare Energien im Jahr 2013*. Technischer Bericht. Bundesministerium für Wirtschaft und Technologie (BMWi), 2014.
- [8] REHTANZ, C. und WIETFELD, C.: „Das Internet der Energie. Trends in der Automatisierung von Energienetzen“. In: *at-Automatisierungstechnik*, Bd. 57, Nr. 10 (2009), S. 514–524.
- [9] ROBINIUS, M.; OTTO, A.; HEUSER, P.; WELDER, L.; SYRANIDIS, K.; RYBERG, S. D.; GRUBE, T.; MARKEWITZ, P.; PETERS, R. und STOLTEN, D.: „Linking the Power and Transport Sectors - Part 1: The Principle of Sector Coupling“. In: *Energies*, Bd. 10, Nr. 7 (2017), S. 956. DOI: [10.3390/en10070956](https://doi.org/10.3390/en10070956).
- [10] ROBINIUS, M.; OTTO, A.; SYRANIDIS, K.; RYBERG, S. D.; HEUSER, P.; WELDER, L.; GRUBE, T.; MARKEWITZ, P.; TIETZE, V. und STOLTEN, D.: „Linking the Power and Transport Sectors - Part 2: Modelling a Sector Coupling Scenario for Germany“. In: *Energies*, Bd. 10, Nr. 7 (2017), S. 957. DOI: [10.3390/en10070957](https://doi.org/10.3390/en10070957).

- [11] APPINO, R.; GONZÁLEZ ORDIANO, J.; MIKUT, R.; FAULWASSER, T. und HAGENMEYER, V.: „On the Use of Probabilistic Forecasts in Scheduling of Renewable Energy Sources Coupled to Storages“. In: *Applied Energy*, Bd. 210 (2018), S. 1207–1218. DOI: [10.1016/j.apenergy.2017.08.133](https://doi.org/10.1016/j.apenergy.2017.08.133).
- [12] KOTTONAU, D.; SHABAGIN, E.; NOE, M. und GROHMANN, S.: „Opportunities for High-Voltage AC Superconducting Cables as Part of New Long-Distance Transmission Lines“. In: *IEEE Transactions on Applied Superconductivity*, Bd. 27, Nr. 4 (2017), S. 1–5. DOI: [10.1109/TASC.2017.2652856](https://doi.org/10.1109/TASC.2017.2652856).
- [13] HAGENMEYER, V.; ÇAKMAK, H. K.; DÜPMEIER, C.; FAULWASSER, T.; ISELE, J.; KELLER, H. B.; KOHLHEPP, P.; KÜHNAPFEL, U.; STUCKY, U.; WACZOWICZ, S. und MIKUT, R.: „Information and Communication Technology in Energy Lab 2.0: Smart Energies System Simulation and Control Center with an Open-Street-Map-based Power Flow Simulation Example“. In: *Energy Technology*, Bd. 4 (2016), S. 145–162. DOI: [10.1002/ente.201500304](https://doi.org/10.1002/ente.201500304).
- [14] BACHEM, A. und BUCHAL, C.: „Energiewende - quo vadis?“ In: *Physik Journal*, Bd. 12, Nr. 12 (2013), S. 33–39. URL: <https://acadeuro-bergen.no/energy/201410-Buchal-et-al-PhysJ1.pdf> (besucht am 24.03.2014).
- [15] ENTELIOS. 2013. URL: <http://entelios.de/demand-response/> (besucht am 06.06.2013).
- [16] BUNDESNETZAGENTUR: *Monitoringbericht 2015*. Technischer Bericht. Bundesnetzagentur, Bundeskartellamt, 2015.
- [17] HAYN, M.: „Modellgestützte Analyse neuer Stromtarife für Haushalte unter Berücksichtigung bedarfsorientierter Versorgungssicherheitsniveaus“. Dissertation. Karlsruher Institut für Technologie, KIT Scientific Publishing, Karlsruhe, Germany, 2016.
- [18] SCHWAB, A. J.: *Elektroenergiesysteme - Erzeugung, Transport, Übertragung und Verteilung elektrischer Energie*. 3. Aufl. Springer, Berlin, Heidelberg, Germany, 2012. URL: <http://link.springer.com/book/10.1007%2F978-3-662-46856-2> (besucht am 03.12.2013).
- [19] MCKENNA, R.; BERTSCH, V.; JOCHEM, P.; GENOESE, M. und FICHTNER, W.: „Angebotsseitige Herausforderungen für die Energiewirtschaft: Erneuerbare und Konventionelle Stromerzeugung“. In: *BFuP - Betriebswirtschaftliche Forschung und Praxis*, Bd. 66, Nr. 3 (2014), S. 317–329.
- [20] APPELRATH, H.-J.; KAGERMANN, H. und MAYER, C.: *Future Energy Grid - Migrationspfade ins Internet der Energie*. Technischer Bericht. Acatech Studie, 2012.
- [21] DEUTSCHE KOMMISSION ELEKTROTECHNIK ELEKTRONIK UND INFORMATIONSTECHNIK IM DIN UND VDE (DKE): *E-Energy - Smart Grids made in Germany*. 2014. URL: <https://teamwork.dke.de/specials/7/Wiki-Seiten/Homepage.aspx> (besucht am 17.03.2014).

-
- [22] FANG, X.; MISRA, S.; XUE, G. und YANG, D.: „Smart Grid - the New and Improved Power Grid: A Survey“. In: *IEEE Communications Surveys Tutorials*, Bd. 14, Nr. 4 (2012), S. 944–980. DOI: [10.1109/SURV.2011.101911.00087](https://doi.org/10.1109/SURV.2011.101911.00087).
- [23] IPAKCHI, A. und ALBUYEH, F.: „Grid of the Future“. In: *IEEE Power and Energy Magazine*, Bd. 7, Nr. 2 (2009), S. 52–62. DOI: [10.1109/MPE.2008.931384](https://doi.org/10.1109/MPE.2008.931384).
- [24] PUSTKA, K.: „Smart Energy Grids“. In: *Innovative Internet Technologies and Mobile Communications (IITM) and Aerospace Networks (AN)* (2012), S. 57–63. URL: <https://pdfs.semanticscholar.org/4c65/9ed0b24493279ee867e6f57d4b1c225f2f6c.pdf> (besucht am 03. 12. 2013).
- [25] BUCHHOLZ, B. M. und STYCZYNSKI, Z.: *Smart Grids - Fundamentals and Technologies in Electricity Networks*. Springer, Berlin, Heidelberg, Germany, 2014. DOI: [10.1007/978-3-642-45120-1](https://doi.org/10.1007/978-3-642-45120-1).
- [26] BUCHHOLZ, B. M. und STYCZYNSKI, Z.: *Smart Grids - Grundlagen und Technologien der elektrischen Netze der Zukunft*. VDE, Berlin, Germany, 2014. URL: <http://link.springer.com/book/10.1007%2F978-3-642-45120-1> (besucht am 14. 10. 2016).
- [27] SIANO, P.: „Demand Response and Smart Grids - a Survey“. In: *Renewable and Sustainable Energy Reviews*, Bd. 30 (2014), S. 461–478. DOI: [10.1016/j.rser.2013.10.022](https://doi.org/10.1016/j.rser.2013.10.022).
- [28] FARUQUI, A.; SERGICI, S. und SHARIF, A.: „The Impact of Informational Feedback on Energy Consumption - a Survey of the Experimental Evidence“. In: *Energy*, Bd. 35, Nr. 4 (2010), S. 1598–1608. DOI: [10.1016/j.energy.2009.07.042](https://doi.org/10.1016/j.energy.2009.07.042).
- [29] WESTERMANN, D.; DÖRING, N. und BRETSCHNEIDER, P., Hrsg.: *Smart Metering: Zwischen technischer Herausforderung und gesellschaftlicher Akzeptanz - Interdisziplinärer Status Quo*. Universitätsverlag Ilmenau, Ilmenau, Germany, 2013.
- [30] WILLE-HAUSSMANN, B.: „Einsatz der symbolischen Modellreduktion zur Untersuchung der Betriebsführung im Smart Grid“. Dissertation. Fakultät für Mathematik und Informatik, Fernuniversität Hagen, 2011.
- [31] GIORDANO, V.; GANGALE, F.; FULLI, G.; JIMÉNEZ, M. S.; ONYEJI, I.; COLTA, A.; PAPAIOANNOU, I.; MENGOLINI, A.; ALECU, C.; OJALA, T. u. a.: *Smart Grid Projects in Europe: Lessons Learned and Current Developments*. Technischer Bericht. European Commission, Joint Research Centre, Institute for Energy, Luxembourg, 2013.
- [32] LEHNHOFF, S.: *Dezentrales vernetztes Energiemanagement*. Vieweg+Teubner, Wiesbaden, Germany, 2010. DOI: [10.1007/978-3-8348-9658-2](https://doi.org/10.1007/978-3-8348-9658-2).
- [33] HEUCK, K.: *Elektrische Energieversorgung: Erzeugung, Übertragung und Verteilung elektrischer Energie für Studium und Praxis*. Hrsg. von DETTMANN, K.-D. und SCHULZ, D. 8. Aufl. Vieweg+Teubner, Wiesbaden, Germany, 2010. URL: <http://link.springer.com/book/10.1007%2F978-3-8348-2174-4> (besucht am 03. 12. 2013).

- [34] MÜLLER, D.; MONTI, A.; STINNER, S.; SCHLÖSSER, T.; SCHÜTZ, T.; MATTHES, P.; WOLISZ, H.; MOLITOR, C.; HARB, H. und STREBLOW, R.: „Demand Side Management for City Districts“. In: *Building and Environment*, Bd. 91 (2015), S. 283–293. DOI: [10.1016/j.buildenv.2015.03.026](https://doi.org/10.1016/j.buildenv.2015.03.026).
- [35] SONNENSCHNEIN, M.; RAPP, B. und BREMER, J.: „Demand Side Management und Demand Response“. In: *Handbuch Energiemanagement*. EW Medien und Kongresse GmbH, Frankfurt am Main, Germany, 2010.
- [36] ALBADI, M. und EL-SAADANY, E.: „A Summary of Demand Response in Electricity Markets“. In: *Electric Power Systems Research*, Bd. 78, Nr. 11 (2008), S. 1989–1996. DOI: [10.1016/j.epsr.2008.04.002](https://doi.org/10.1016/j.epsr.2008.04.002).
- [37] HAN, J. und PIETTE, M.: „Solutions for Summer Electric Power Shortages: Demand Response and Its Applications in Air Conditioning and Refrigerating Systems“. In: *Refrigeration, Air Conditioning, & Electric Power Machinery*, Bd. 29, Nr. 1 (2008), S. 1–4. URL: <https://drcc.lbl.gov/sites/default/files/63806.pdf> (besucht am 23.07.2014).
- [38] HINTERBERGER, R. und POLAK, S.: „Lastverschiebung in Industrie und Gewerbe in Österreich Chancen und Potentiale in zukünftigen Smart Grids“. In: *7. Internationale Energiewirtschaftstagung an der TU Wien*. (February 16-18, Vienna, Austria), 2011, S. 1–16.
- [39] PALENSKY, P. und DIETRICH, D.: „Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads“. In: *IEEE Transactions on Industrial Informatics*, Bd. 7, Nr. 3 (2011), S. 381–388. DOI: [10.1109/TII.2011.2158841](https://doi.org/10.1109/TII.2011.2158841).
- [40] ROON, S. VON und GOBMAIER, T.: *Demand Response in der Industrie–Status und Potenziale in Deutschland*. Technischer Bericht. Forschungsstelle für Energiewirtschaft eV (FfE), München, 2010.
- [41] ESTHER, B. und KUMAR, K.: „A Survey on Residential Demand Side Management Architecture, Approaches, Optimization Models and Methods“. In: *Renewable and Sustainable Energy Reviews*, Bd. 59 (2016), S. 342–351. DOI: [10.1016/j.rser.2015.12.282](https://doi.org/10.1016/j.rser.2015.12.282).
- [42] KHAJAVI, P.; ABNIKI, H. und ARANI, A. B.: „The Role of Incentive Based Demand Response Programs in Smart Grid“. In: *10th International Conference on Environment and Electrical Engineering (EEEIC)*. (May 8-11, Rome, Italy). IEEE, Piscataway, NJ, USA, 2011, S. 1–4. DOI: [10.1109/EEEIC.2011.5874702](https://doi.org/10.1109/EEEIC.2011.5874702).
- [43] CALLAWAY, D. und HISKENS, I.: „Achieving Controllability of Electric Loads“. In: *Proc., IEEE*, Bd. 99, Nr. 1 (2011), S. 184–199. DOI: [10.1109/JPROC.2010.2081652](https://doi.org/10.1109/JPROC.2010.2081652).
- [44] KOCH, S.; ZIMA, M. und ANDERSSON, G.: „Potentials and Applications of Coordinated Groups of Thermal Household Appliances for Power System Control Purposes“. In: *IEEE PES/IAS Conference on Sustainable Alternative Energy (SAE)*. (September 28-30, Valencia, Spain). IEEE, Piscataway, NJ, USA, 2009, S. 1–8. DOI: [10.1109/SAE.2009.5534841](https://doi.org/10.1109/SAE.2009.5534841).

-
- [45] KOCH, S.; GALUS, M. D.; CHATZIVASILEIADIS, S. und ANDERSSON, G.: „Emergency Control Concepts for Future Power Systems“. In: *Proc., 18th IFAC World Congress*. (August 28 - September 2, Milano, Italy). Elsevier, Amsterdam, Netherlands, 2011, S. 6121–6129. DOI: [10.3182/20110828-6-it-1002.02389](https://doi.org/10.3182/20110828-6-it-1002.02389).
- [46] KOCH, S.: „Demand Response Methods for Ancillary Services and Renewable Energy Integration in Electric Power Systems“. Dissertation. ETH Zürich, 2012.
- [47] VIVEKANANTHAN, C.; MISHRA, Y.; LEDWICH, G. und LI, F.: „Demand Response for Residential Appliances Via Customer Reward Scheme“. In: *IEEE Transactions on Smart Grid*, Bd. 5, Nr. 2 (2014), S. 809–820. DOI: [10.1109/TSG.2014.2298514](https://doi.org/10.1109/TSG.2014.2298514).
- [48] AALAMI, H.; MOGHADDAM, M. P. und YOUSEFI, G.: „Demand Response Modeling Considering Interruptible/Curtailable Loads and Capacity Market Programs“. In: *Applied Energy*, Bd. 87, Nr. 1 (2010), S. 243–250. DOI: [10.1016/j.apenergy.2009.05.041](https://doi.org/10.1016/j.apenergy.2009.05.041).
- [49] AALAMI, H.; YOUSEFI, G.-R. und MOGHADDAM, M. P.: „Demand Response Model Considering EDRP and TOU Programs“. In: *IEEE PES Transmission and Distribution Conference and Exposition T&D*. (April 21-24, Chicago, USA). IEEE, Piscataway, NJ, USA, 2008, S. 1–6. DOI: [10.1109/TDC.2008.4517059](https://doi.org/10.1109/TDC.2008.4517059).
- [50] FERNANDES, R.; DE ALMEIDA, S. A. B.; BARBOSA, F. und PESTANA, R.: „Load Shedding - Coordination Between the Portuguese Transmission Grid and the Distribution Grid with Minimization of Loss of Distributed Generation“. In: *Proc., IEEE PowerTech*. (June 28 - July 2, Bucharest, Hungary). IEEE, Piscataway, NJ, USA, 2009, S. 1–6. DOI: [10.1109/PTC.2009.5282171](https://doi.org/10.1109/PTC.2009.5282171).
- [51] MAQBOOL, S.; BABAR, M. und AL-AMMAR, E.: „Effects of Demand Elasticity and Price Variation on Load Profile“. In: *IEEE PES Conference on Innovative Smart Grid Technologies - Middle East*. (December 17-20, Jeddah, Saudi Arabia). IEEE, Piscataway, NJ, USA, 2011, S. 1–5. DOI: [10.1109/ISGT-MidEast.2011.6220793](https://doi.org/10.1109/ISGT-MidEast.2011.6220793).
- [52] WANG, J.; KENNEDY, S. und KIRTLEY, J.: „A New Wholesale Bidding Mechanism for Enhanced Demand Response in Smart Grids“. In: *Innovative Smart Grid Technologies*. (January 19-21, Gaithersburg, USA). IEEE, Piscataway, NJ, USA, 2010, S. 1–8. DOI: [10.1109/ISGT.2010.5434766](https://doi.org/10.1109/ISGT.2010.5434766).
- [53] CHRISTAKOU, K.; TOMOZEI, D.-C.; LE BOUDEC, J.-Y. und PAOLONE, M.: „GECN: Primary Voltage Control for Active Distribution Networks via Real-Time Demand-Response“. In: *IEEE Transactions on Smart Grid*, Bd. 5, Nr. 2 (2014), S. 622–631. DOI: [10.1109/TSG.2013.2275004](https://doi.org/10.1109/TSG.2013.2275004).
- [54] HEFFNER, G.; GOLDMAN, C.; KIRBY, B. und KINTNER-MEYER, M.: *Loads Providing Ancillary Services: Review of International Experience*. Technischer Bericht. Lawrence Berkeley National Laboratory, 2008.
- [55] POURMOUSAVI, S. und NEHRIR, M.: „Real-Time Central Demand Response for Primary Frequency Regulation in Microgrids“. In: *IEEE Transactions on Smart Grid*, Bd. 3, Nr. 4 (2012), S. 1988–1996. DOI: [10.1109/TSG.2012.2201964](https://doi.org/10.1109/TSG.2012.2201964).

- [56] KLAIBER, S.; BRETSCHEIDER, P.; WACZOWICZ, S.; MIKUT, R.; KONOTOP, I. und WESTERMANN, D.: „Intelligente Prognoseverfahren für beeinflusstes Verbrauchsverhalten in Energiesystemen“. In: *Proc., 23. Workshop Computational Intelligence*. (December 5-6, Dortmund, Germany). KIT Scientific Publishing, Karlsruhe, Germany, 2013, S. 93–108.
- [57] DE BRUYN, K.; KOLLMANN, A.; MOSER, S.; SCHMIDTHALER, M.; AMANN, C.; ELBE, C.; SCHMAUTZER, E.; KRAUSSLER, A.; REINHOFER-GUBISCH, M.; PUCKER, J. und FRANTES, B.: *LoadShift: Lastverschiebung in Haushalt, Industrie, Gewerbe und kommunaler Infrastruktur – Potenzialanalyse für Smart Grids*. Technischer Bericht. Eigentümer, Herausgeber und Medieninhaber: Bundesministerium für Verkehr, Innovation und Technologie, Wien, Austria, 2015.
- [58] GOBMAIER, T. H.: „Entwicklung und Anwendung einer Methodik zur Synthese zukünftiger Verbraucherlastgänge“. Dissertation. Universität München, 2013.
- [59] GYAMFI, S.; KRUMDIECK, S. und URMEE, T.: „Residential Peak Electricity Demand Response - Highlights of Some Behavioural Issues“. In: *Renewable and Sustainable Energy Reviews*, Bd. 25 (2013), S. 71–77. DOI: [10.1016/j.rser.2013.04.006](https://doi.org/10.1016/j.rser.2013.04.006).
- [60] CAVES, D. W.; CHRISTENSEN, L. R. und HERRIGES, J. A.: „Consistency of Residential Customer Response in Time-Of-Use Electricity Pricing Experiments“. In: *Journal of Econometrics*, Bd. 26 (1984), S. 179–203. DOI: [10.1016/0304-4076\(84\)90017-4](https://doi.org/10.1016/0304-4076(84)90017-4).
- [61] REISS, P. C. und WHITE, M. W.: „Household Electricity Demand, Revisited“. In: *The Review of Economic Studies*, Bd. 72, Nr. 3 (2005), S. 853–883. DOI: [10.1111/0034-6527.00354](https://doi.org/10.1111/0034-6527.00354).
- [62] TEWS, K.: *Stromeffizienttarife für Verbraucher in Deutschland? Vom Sinn, der Machbarkeit und den Alternativen einer progressiven Tarifsteuerung*. Technischer Bericht. Freie Universität Berlin, Forschungszentrum für Umweltpolitik, 2011.
- [63] ADEPETU, A.; REZAEI, E.; LIZOTTE, D. und KESHAV, S.: „Critiquing Time-of-Use pricing in Ontario“. In: *IEEE International Conference on Smart Grid Communications (SmartGridComm)*. (October 21-24, Vancouver, Canada). IEEE, Piscataway, NJ, USA, 2013, S. 223–228. DOI: [10.1109/SmartGridComm.2013.6687961](https://doi.org/10.1109/SmartGridComm.2013.6687961).
- [64] eTELLIGENCE: *Abschlussbericht eTelligence*. Technischer Bericht. EWE AG, 2012. URL: http://www.e-energy.de/documents/EWE_102189_EVE_eTelligence_Abschlussbericht_Inhalt_GB_Internet_sc_pdf.pdf (besucht am 03. 12. 2013).
- [65] INTELLIEKON: *Nachhaltiger Energiekonsum von Haushalten durch intelligente Zähler-, Kommunikations- und Tarifsysteme*. Technischer Bericht. Forschungsprojekt gefördert durch das deutsche Bundesministerium für Bildung und Forschung (BMBF), 2011. URL: <http://www.intelliekon.de/> (besucht am 03. 12. 2013).
- [66] KIESSLING, A.: *Modellstadt Mannheim – Evaluation der Feldtests und Simulationen, Endbericht*. Technischer Bericht. E-Energy, Bundesministerium für Wirtschaft und Technologie (BMWi), Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit (BMU), 2013.

-
- [67] KLOBASA, M.: „Dynamische Simulation eines Lastmanagements und Integration von Windenergie in ein Elektrizitätsnetz auf Landesebene unter regelungstechnischen und Kostengesichtspunkten“. Dissertation. ETH Zürich, 2007.
- [68] FARUQUI, A. und SERGICI, S.: „Household Response to Dynamic Pricing of Electricity: A Survey of 15 Experiments“. In: *Journal of Regulatory Economics*, Bd. 38, Nr. 2 (2010), S. 193–225. DOI: [10.1007/s11149-010-9127-y](https://doi.org/10.1007/s11149-010-9127-y).
- [69] HILLEMACHER, L.: „Lastmanagement mittels dynamischer Strompreissignale bei Haushaltskunden“. Dissertation. Karlsruher Institut für Technologie, 2014.
- [70] FARUQUI, A.; HLEDIK, R. und PALMER, J.: *Time-Varying and Dynamic Rate Design*. Technischer Bericht. The Regulatory Assistance Project, 2012.
- [71] PATERAKIS, N. G.; ERDIŇ, O. und CATALÃO, J. P.: „An Overview of Demand Response: Key-elements and International Experience“. In: *Renewable and Sustainable Energy Reviews*, Bd. 69 (2017), S. 871–891. DOI: [10.1016/j.rser.2016.11.167](https://doi.org/10.1016/j.rser.2016.11.167).
- [72] KLOBASA, M.; ERGE, T.; BUKVIC-SCHÄFER, A. S. und HOLLMANN, M.: „Demand Side Management in dezentral geführten Verteilnetzen (Erfahrungen und Perspektiven)“. In: *Elfte Kasseler Symposium Energie-Systemtechnik Informations- und Kommunikationstechnologien für die Energieversorgung*. (November 9-10, Kassel, Germany). Institut für Solare Energieversorgungstechnik (ISET), Verein an der Universität Kassel e.V., Kassel, Germany, 2006, S. 115–134.
- [73] HAMMERSTROM, D. J.; AMBROSIO, R.; BROUS, J.; CARLON, T. A.; CHASSIN, D. P.; DESTEESE, J. G.; GUTTROMSON, R. T.; HORST, G. R.; JÄRVEGREN, O. M.; KAJFASZ, R. u. a.: *Pacific Northwest GridWise Testbed Demonstration Projects, Part I. Olympic Peninsula Project*. Technischer Bericht. Pacific Northwest National Laboratory (PNNL), Richland, Washington, 2007. DOI: [10.2172/926113](https://doi.org/10.2172/926113).
- [74] HAMMERSTROM, D. J.; BROUS, J.; CHASSIN, D. P.; HORST, G. R.; KAJFASZ, R.; MICHIE, P.; OLIVER, T. V.; CARLON, T. A.; EUSTIS, C.; JÄRVEGREN, O. M. u. a.: *Pacific Northwest GridWise Testbed Demonstration Projects; Part II. Grid Friendly Appliance Project*. Technischer Bericht. Pacific Northwest National Laboratory (PNNL), Richland, Washington, 2007. DOI: [10.2172/926122](https://doi.org/10.2172/926122).
- [75] CHASSIN, D. P. und KIESLING, L.: „Decentralized Coordination Through Digital Technology, Dynamic Pricing, and Customer-Driven Control: The GridWise Testbed Demonstration Project“. In: *The Electricity Journal*, Bd. 21, Nr. 8 (2008), S. 51–59. DOI: [10.1016/j.tej.2008.09.002](https://doi.org/10.1016/j.tej.2008.09.002).
- [76] WACZOWICZ, S.; KLAIBER, S.; BRETSCHNEIDER, P.; KONOTOP, I.; WESTERMANN, D.; REISCHL, M. und MIKUT, R.: „Data Mining zur Analyse der Auswirkungen von Preissignalen auf Haushaltsstromkunden“. In: *at-Automatisierungstechnik*, Bd. 62, Nr. 10 (2014), S. 740–752. DOI: [10.1515/auto-2014-1108](https://doi.org/10.1515/auto-2014-1108).
- [77] ARNOLDT, A.; WARWEG, O.; FRENZEL, S.; WERNER, M. und BRETSCHNEIDER, P.: „Analyse des Einflusses vom Demand Response RESIDENS Feldversuch auf das Kundenlastverhalten“. In: *Proc., VDE-Kongress 2012 - Intelligente Energieversorgung der Zukunft*. (November 5-6, Stuttgart, Germany). VDE, Berlin, Germany, 2012.

- [78] IFLAND, M.; EXNER, N.; DÖRING, N. und WESTERMANN, D.: „Influencing Domestic Customers’ Market Behavior with Time Flexible Tariffs“. In: *3rd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies (ISGT Europe)*. (October 14-17, Berlin, Germany). IEEE, Piscataway, NJ, USA, 2012, S. 1–7. DOI: [10.1109/ISGTEurope.2012.6465638](https://doi.org/10.1109/ISGTEurope.2012.6465638).
- [79] DÜTSCHKE, E.; UNTERLÄNDER, M. und WIETSCHEL, M.: *Variable Stromtarife aus Kundensicht: Akzeptanzstudie auf Basis einer Conjoint-Analyse*. Working paper sustainability and innovation 1. Karlsruhe: Fraunhofer Institute for Systems und Innovation Research ISI, 2012. URL: <http://hdl.handle.net/10419/55218> (besucht am 03. 12. 2013).
- [80] HÜTTER, D.: „Auswirkungen von zeitvariablen Preisstrukturen auf den Spitzenleistungsbedarf in Österreich“. In: *7. Internationale Energiewirtschaftstagung an der TU Wien*. Bd. 2. (February 16-18, Vienna, Austria), 2011, S. 1–21.
- [81] NABE, C.; BEYER, C.; BRODERSEN, N.; SCHÄFFLER, H.; ADAM, D.; HEINEMANN, C.; TUSCH, T.; EDER, J.; WYL, C. DE; WEGE, J.-H. VOM u. a.: *Einführung von lastvariablen und zeitvariablen Tarifen*. Technischer Bericht. ECOFYS, EnCT & BBH im Auftrag der Bundesnetzagentur, Köln, 2010.
- [82] KOHLHEPP, P. und HAGENMEYER, V.: „Technical Potential of Buildings in Germany as Flexible Power-to-Heat Storages for Smart Grid Operation“. In: *Energy Technology*, Bd. 5, Nr. 7 (2017), S. 1084–1104. DOI: [10.1002/ente.201600655](https://doi.org/10.1002/ente.201600655).
- [83] SCHWARTZ, T.; STEVENS, G.; JAKOBI, T.; DENEFF, S.; RAMIREZ, L.; WULF, V. und RANDALL, D.: „What People Do with Consumption Feedback: A Long-Term Living Lab Study of a Home Energy Management System“. In: *Interacting with Computers* (2014), S. 1–26. DOI: [10.1093/iwc/iwu009](https://doi.org/10.1093/iwc/iwu009).
- [84] FAN, Z.; KULKARNI, P.; GORMUS, S.; EFTHYMIU, C.; KALOGRIDIS, G.; SOORIYABANDARA, M.; ZHU, Z.; LAMBOTHARAN, S. und CHIN, W. H.: „Smart Grid Communications: Overview of Research Challenges, Solutions, and Standardization Activities“. In: *IEEE Communications Surveys Tutorials*, Bd. 15, Nr. 1 (2013), S. 21–38. DOI: [10.1109/SURV.2011.122211.00021](https://doi.org/10.1109/SURV.2011.122211.00021).
- [85] MIKUT, R.: *Data Mining in der Medizin und Medizintechnik*. Universitätsverlag Karlsruhe, Karlsruhe, Germany, 2008. URL: <https://publikationen.bibliothek.kit.edu/1000008476> (besucht am 03. 12. 2013).
- [86] ALBERS, S.; KLAPPER, D.; KONRADT, U.; WALTER, A. und WOLF, J., Hrsg.: *Methodik der empirischen Forschung*. Gabler, Wiesbaden, Germany, 2009. DOI: [10.1007/978-3-322-96406-9](https://doi.org/10.1007/978-3-322-96406-9).
- [87] WACZOWICZ, S.: „Entwicklung einer Photovoltaik-Leistungsprognose für die Anwendung in einer dezentralen Energieversorgungsstruktur“. Diplomarbeit. Karlsruher Institut für Technologie, Karlsruhe, Germany, 2012.
- [88] GARCIA, D.: „Robust Smoothing of Gridded Data in One and Higher Dimensions with Missing Values“. In: *Computational Statistics & Data Analysis*, Bd. 54, Nr. 4 (2010), S. 1167–1178. DOI: [10.1016/j.csda.2009.09.020](https://doi.org/10.1016/j.csda.2009.09.020).

-
- [89] TAK, S.; WOO, S. und YEO, H.: „Data-Driven Imputation Method for Traffic Data in Sectional Units of Road Links“. In: *IEEE Transactions on Intelligent Transportation Systems*, Bd. 17, Nr. 6 (2016), S. 1762–1771. DOI: [10.1109/TITS.2016.2530312](https://doi.org/10.1109/TITS.2016.2530312).
- [90] MORITZ, S.; SARDÁ, A.; BARTZ-BEIELSTEIN, T.; ZAEFFERER, M. und STORK, J.: „Comparison of Different Methods for Univariate Time Series Imputation in R“. In: *ArXiv e-prints* (2015). URL: [arXiv:1510.03924](https://arxiv.org/abs/1510.03924) (besucht am 02. 02. 2018).
- [91] FREUND, R. W. und HOPPE, R. H. W.: *Stoer/Bulirsch: Numerische Mathematik 1*. Springer, Berlin, Heidelberg, Germany, 2007. DOI: [10.1007/978-3-540-45390-1](https://doi.org/10.1007/978-3-540-45390-1).
- [92] WANG, G.; GARCIA, D.; LIU, Y.; JEU, R. DE und DOLMAN, A. J.: „A Three-Dimensional Gap Filling Method for Large Geophysical Datasets: Application to Global Satellite Soil Moisture Observations“. In: *Environmental Modelling & Software*, Bd. 30 (2012), S. 139–142. DOI: [10.1016/j.envsoft.2011.10.015](https://doi.org/10.1016/j.envsoft.2011.10.015).
- [93] HARVEY, A. C. und SHEPHARD, N.: „10 Structural Time Series Models“. In: *Handbook of statistics*, Bd. 11 (1993), S. 261–302. DOI: [10.1016/s0169-7161\(05\)80045-8](https://doi.org/10.1016/s0169-7161(05)80045-8).
- [94] VAN BUUREN, S.: *Flexible Imputation of Missing Data*. Chapman und Hall/CRC Interdisciplinary Statistics, Boca Raton, Florida, USA, 2012. DOI: [10.1201/b11826](https://doi.org/10.1201/b11826).
- [95] GRAHAM, J. W.: „Missing Data Analysis: Making It Work in the Real World“. In: *Annual Review of Psychology*, Bd. 60 (2009), S. 549–576. DOI: [10.1146/annurev.psych.58.110405.085530](https://doi.org/10.1146/annurev.psych.58.110405.085530).
- [96] AKOUEMO, H. N. und POVINELLI, R. J.: „Time Series Outlier Detection and Imputation“. In: *IEEE PES General Meeting and Conference Exposition*. (July 27-31, National Harbor, USA). IEEE, Piscataway, NJ, USA, 2014, S. 1–5. DOI: [10.1109/PESGM.2014.6939802](https://doi.org/10.1109/PESGM.2014.6939802).
- [97] GONZÁLEZ ORDIANO, J. Á.; DONEIT, W.; WACZOWICZ, S.; GRÖLL, L.; MIKUT, R. und HAGENMEYER, V.: „Nearest-Neighbor Based Non-Parametric Probabilistic Forecasting with Applications in Photovoltaic Systems“. In: *Proc., 26. Workshop Computational Intelligence*. (November 24-25, Dortmund, Germany). KIT Scientific Publishing, Karlsruhe, Germany, 2016, S. 9–30.
- [98] LJUNG, G. M.: „On Outlier Detection in Time Series“. English. In: *Journal of the Royal Statistical Society. Series B (Methodological)*, Bd. 55, Nr. 2 (1993), S. 559–567. URL: <http://www.jstor.org/stable/2346215> (besucht am 26. 04. 2015).
- [99] STIER, W.: *Methoden der Zeitreihenanalyse*. Springer, Berlin, Heidelberg, Germany, 2001. DOI: [10.1007/978-3-642-56709-4](https://doi.org/10.1007/978-3-642-56709-4).
- [100] BARNETT, V.: „The Study of Outliers: Purpose and Model“. In: *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Bd. 27, Nr. 3 (1978), S. 242–250. DOI: [10.2307/2347159](https://doi.org/10.2307/2347159).
- [101] BEN-GAL, I.: „Outlier Detection“. In: *Data Mining and Knowledge Discovery Handbook*. Springer, Boston, MA, USA, 2005, S. 131–146. DOI: [10.1007/978-0-387-09823-4_7](https://doi.org/10.1007/978-0-387-09823-4_7).
- [102] CHROMIŃSKI, K. und TKACZ, M.: „Comparison of Outlier Detection Methods in Biomedical Data“. In: *Journal of Medical Informatics & Technologies*, Bd. 16 (2010), S. 89–94.

- [103] POSIO, J.; LEIVISKÄ, K.; RUUSKA, J. und RUHA, P.: „Outlier Detection for 2D Temperature Data“. In: *Proc., 17th World Congress The International Federation of Automatic Control (IFAC)*. (July 6-11, Seoul, Korea). Elsevier, Amsterdam, Netherlands, 2008. DOI: [10.3182/20080706-5-kr-1001.00333](https://doi.org/10.3182/20080706-5-kr-1001.00333).
- [104] ALFARES, H. K. und NAZEERUDDIN, M.: „Electric Load Forecasting: Literature Survey and Classification of Methods“. In: *International Journal of Systems Science*, Bd. 33, Nr. 1 (2002), S. 23–34. DOI: [10.1080/00207720110067421](https://doi.org/10.1080/00207720110067421).
- [105] CHAN, S. C.; TSUI, K. M.; WU, H. C.; HOU, Y.; WU, Y. C. und WU, F. F.: „Load/Price Forecasting and Managing Demand Response for Smart Grids: Methodologies and Challenges“. In: *IEEE Signal Processing Magazine*, Bd. 29, Nr. 5 (2012), S. 68–85. DOI: [10.1109/MSP.2012.2186531](https://doi.org/10.1109/MSP.2012.2186531).
- [106] DEB, C.; ZHANG, F.; YANG, J.; LEE, S. E. und SHAH, K. W.: „A Review on Time Series Forecasting Techniques for Building Energy Consumption“. In: *Renewable and Sustainable Energy Reviews*, Bd. 74 (2017), S. 902–924. DOI: [10.1016/j.rser.2017.02.085](https://doi.org/10.1016/j.rser.2017.02.085).
- [107] GONZÁLEZ ORDIANO, J.; WACZOWICZ, S.; HAGENMEYER, V. und MIKUT, R.: „Energy Forecasting Tools and Services“. In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* (2017). DOI: [10.1002/widm.1235](https://doi.org/10.1002/widm.1235).
- [108] SUGANTHI, L. und SAMUEL, A. A.: „Energy Models for Demand Forecasting - A Review“. In: *Renewable and Sustainable Energy Reviews*, Bd. 16, Nr. 2 (2012), S. 1223–1240.
- [109] SWAN, L. G. und UGURSAL, V. I.: „Modeling of End-Use Energy Consumption in the Residential Sector: A Review of Modeling Techniques“. In: *Renewable and Sustainable Energy Reviews*, Bd. 13, Nr. 8 (2009), S. 1819–1835. DOI: [10.1016/j.rser.2008.09.033](https://doi.org/10.1016/j.rser.2008.09.033).
- [110] COUGHLIN, K.; PIETTE, M. A.; GOLDMAN, C. und KILICCOTE, S.: *Estimating Demand Response Load Impacts: Evaluation of Baseline Load Models for Non-Residential Buildings in California*. Technischer Bericht. Lawrence Berkeley National Laboratory, 2008. DOI: [10.2172/928452](https://doi.org/10.2172/928452).
- [111] ENERNOC: *The Demand Response Baseline*. Technischer Bericht. EnerNOC, Inc., 2011. URL: https://library.cee1.org/sites/default/files/library/10774/CEE%5C_EvalDRBaseline%5C_2011.pdf (besucht am 19.07.2017).
- [112] BARTHOLOMEW, P.; CALLENDER, W.; HINDES, C.; GRIMM, C.; JOHNSON, K.; STRAUB, M.; WILLIAMS, D.; WILLIAMSON, M.; HAYES, D.; JOHNSON, W.; NIX, B.; LYNCH, J. und ROMER, S.: *Demand Response Measurement & Verification*. Technischer Bericht. AEIC Load Research Committee, 2009. URL: https://www.smartgrid.gov/files/demand%5C_response.pdf (besucht am 19.07.2017).
- [113] MOHAJERYAMI, S.; DOOSTAN, M. und SCHWARZ, P.: „The Impact of Customer Baseline Load (CBL) Calculation Methods on Peak Time Rebate Program Offered to Residential Customers“. In: *Electric Power Systems Research*, Bd. 137 (2016), S. 59–65. DOI: [10.1016/j.epsr.2016.03.050](https://doi.org/10.1016/j.epsr.2016.03.050).

-
- [114] MOHAJERYAMI, S.; DOOSTAN, M.; ASADINEJAD, A. und SCHWARZ, P.: „Error Analysis of Customer Baseline Load (CBL) Calculation Methods for Residential Customers“. In: *IEEE Transactions on Industry Applications*, Bd. 53, Nr. 1 (2017), S. 5–14. DOI: [10.1109/TIA.2016.2613985](https://doi.org/10.1109/TIA.2016.2613985).
- [115] ASARE-BEDIAKO, B.; KLING, W. und RIBEIRO, P.: „Future Residential Load Profiles: Scenario-Based Analysis of High Penetration of Heavy Loads and Distributed Generation“. In: *Energy and Buildings*, Bd. 75 (2014), S. 228–238. DOI: [10.1016/j.enbuild.2014.02.025](https://doi.org/10.1016/j.enbuild.2014.02.025).
- [116] SOCOLOW, R. H.: „Saving Energy in the Home: Princeton’s Experiments at Twin Rivers“. In: *Ballinger Publishing Company, Cambridge, MA* (1978), S. 231–254.
- [117] HAYN, M.; BERTSCH, V. und FICHTNER, W.: „Electricity Load Profiles in Europe: The Importance of Household Segmentation“. In: *Energy Research & Social Science*, Bd. 3 (2014), S. 30–45. DOI: [10.1016/j.erss.2014.07.002](https://doi.org/10.1016/j.erss.2014.07.002).
- [118] PARKER, D. S.; MAZZARA, M. D. und SHERWIN, J. R.: „Monitored Energy Use Patterns in Low-Income Housing in a Hot and Humid Climate“. In: *Proc., Tenth Symposium on Improving Building Systems in Hot Humid Climates*. Bd. 316. (May 13-14, Fort Worth, TX), 1996.
- [119] LUTZENHISER, L.: „Social and Behavioral Aspects of Energy Use“. In: *Annual Review of Energy and the Environment*, Bd. 18, Nr. 1 (1993), S. 247–289. DOI: [10.1146/annurev.eg.18.110193.001335](https://doi.org/10.1146/annurev.eg.18.110193.001335).
- [120] TWANABASU, S. R. und BREMDAL, B. A.: „Load Forecasting in a Smart Grid Oriented Building“. In: *22nd International Conference and Exhibition on Electricity Distribution (CIRED)*. (June 10-13, Stockholm, Sweden). Institution of Engineering und Technology (IET), London, UK, 2013, S. 1–4. DOI: [10.1049/cp.2013.0997](https://doi.org/10.1049/cp.2013.0997).
- [121] CHUJAI, P.; KERDPRASOP, N. und KERDPRASOP, K.: „Time Series Analysis of Household Electric Consumption with ARIMA and ARMA Models“. In: *Proc., International MultiConference of Engineers and Computer Scientists (IMECS)*. Hrsg. von AO, S. I.; CASTILLO, O.; DOUGLAS, C.; DAGAN FENG, D. und LEE, J.-A. Bd. 1. (March 13-15, Hong Kong). Newswood Limited, Hong Kong, 2013, S. 295–300. URL: http://www.iaeng.org/publication/IMECS2013/IMECS2013_pp295-300.pdf (besucht am 13.02.2018).
- [122] DING, Y.; NEUMANN, M. A.; STAMM, E.; BEIGL, M.; INOUE, S. und PAN, X.: „A Personalized Load Forecasting Enhanced by Activity Information“. In: *Proc., IEEE First International Smart Cities Conference (ISC2)*. (October 25-28, Guadalajara, Mexico). IEEE, Piscataway, NJ, USA, 2015, S. 1–6. DOI: [10.1109/ISC2.2015.7366172](https://doi.org/10.1109/ISC2.2015.7366172).
- [123] JAVED, F.; ARSHAD, N.; WALLIN, F.; VASSILEVA, I. und DAHLQUIST, E.: „Forecasting for Demand Response in Smart Grids: An analysis on Use of Anthropologic and Structural Data and Short Term Multiple Loads Forecasting“. In: *Applied Energy*, Bd. 96 (2012), S. 150–160. DOI: [10.1016/j.apenergy.2012.02.027](https://doi.org/10.1016/j.apenergy.2012.02.027).
- [124] GHIASSI, M.; ZIMBRA, D. K. und SAIDANE, H.: „Medium Term System Load Forecasting with a Dynamic Artificial Neural Network Model“. In: *Electric Power Systems Research*, Bd. 76, Nr. 5 (2006), S. 302–316.

- [125] HIPPERT, H. S.; PEDREIRA, C. E. und SOUZA, R. C.: „Neural Networks for Short-Term Load Forecasting: A Review and Evaluation“. In: *IEEE Transactions on Power Systems*, Bd. 16, Nr. 1 (2001), S. 44–55.
- [126] SENJYU, T.; TAKARA, H.; UEZATO, K. und FUNABASHI, T.: „One-Hour-Ahead Load Forecasting Using Neural Network“. In: *IEEE Transactions on Power Systems*, Bd. 17, Nr. 1 (2002), S. 113–118. DOI: [10.1109/59.982201](https://doi.org/10.1109/59.982201).
- [127] TAYLOR, J. W. und BUIZZA, R.: „Neural Network Load Forecasting with Weather Ensemble Predictions“. In: *IEEE Transactions on Power Systems*, Bd. 17, Nr. 3 (2002), S. 626–632.
- [128] YU, F. und XU, X.: „A Short-Term Load Forecasting Model of Natural Gas Based on Optimized Genetic Algorithm and Improved BP Neural Network“. In: *Applied Energy*, Bd. 134 (2014), S. 102–113. DOI: [10.1016/j.apenergy.2014.07.104](https://doi.org/10.1016/j.apenergy.2014.07.104).
- [129] PARK, D. C.; EL-SHARKAWI, M.; MARKS, R.; ATLAS, L. und DAMBORG, M.: „Electric Load Forecasting using an Artificial Neural Network“. In: *IEEE Transactions on Power Systems*, Bd. 6, Nr. 2 (1991), S. 442–449. DOI: [10.1109/59.76685](https://doi.org/10.1109/59.76685).
- [130] LLANOS, J.; SAEZ, D.; PALMA-BEHNKE, R.; NUNEZ, A. und JIMENEZ-ESTEVEZ, G.: „Load Profile Generator and Load Forecasting for a Renewable Based Microgrid Using Self Organizing Maps and Neural Networks“. In: *Proc., Int. Joint Conf. on Neural Networks (IJCNN)*. (June 10-15, Brisbane, Australia). IEEE, Piscataway, NJ, USA, 2012, S. 1–8. DOI: [10.1109/IJCNN.2012.6252648](https://doi.org/10.1109/IJCNN.2012.6252648).
- [131] CHEN, M. Y. und LINKENS, D. A.: „Rule-Base Self-Generation and Simplification for Data-Driven Fuzzy Models“. In: *Fuzzy Sets and Systems*, Bd. 142, Nr. 2 (2004), S. 243–265. DOI: [10.1016/S0165-0114\(03\)00160-X](https://doi.org/10.1016/S0165-0114(03)00160-X).
- [132] LIU, H.; LIU, D.; LIANG, Y.-M. und ZHENG, G.: „Research on Natural Gas Load Forecasting Based on Least Squares Support Vector Machine“. In: *Proc., International Conference on Machine Learning and Cybernetics*. Bd. 5. (August 26-29, Shanghai, China). IEEE, Piscataway, NJ, USA, 2004, S. 3124–3128.
- [133] CEPERIC, E.; CEPERIC, V. und BARIC, A.: „A Strategy for Short-term Load Forecasting by Support Vector Regression Machines“. In: *IEEE Transactions on Power Systems*, Bd. 28, Nr. 4 (2013), S. 4356–4364. DOI: [10.1109/tpwrs.2013.2269803](https://doi.org/10.1109/tpwrs.2013.2269803).
- [134] HUMEAU, S.; WIJAYA, T. K.; VASIRANI, M. und ABERER, K.: „Electricity Load Forecasting for Residential Customers: Exploiting Aggregation and Correlation between Households“. In: *Sustainable Internet and ICT for Sustainability (SustainIT)*. (October 30-31, Palermo, Italy). IEEE, Piscataway, NJ, USA, 2013, S. 1–6. DOI: [10.1109/SustainIT.2013.6685208](https://doi.org/10.1109/SustainIT.2013.6685208).
- [135] CHRISTIAANSE, W. R.: „Short-Term Load Forecasting Using General Exponential Smoothing“. In: *IEEE Transactions on Power Apparatus and Systems*, Bd. PAS-90, Nr. 2 (1971), S. 900–911. DOI: [10.1109/TPAS.1971.293123](https://doi.org/10.1109/TPAS.1971.293123).
- [136] KARIM, S. A. A. und ALWI, S. A.: „Electricity Load Forecasting in UTP using Moving Averages and Exponential Smoothing Techniques“. In: *Applied Mathematical Sciences*, Bd. 7, Nr. 80 (2013), S. 4003–4014. DOI: [10.12988/ams.2013.33149](https://doi.org/10.12988/ams.2013.33149).

-
- [137] TAYLOR, J. W.; MENEZES, L. M. DE und McSHARRY, P. E.: „A Comparison of Univariate Methods for Forecasting Electricity Demand up to a Day Ahead“. In: *International Journal of Forecasting*, Bd. 22, Nr. 1 (2006), S. 1–16. DOI: [10.1016/j.ijforecast.2005.06.006](https://doi.org/10.1016/j.ijforecast.2005.06.006).
- [138] TAYLOR, J. W.: „An Evaluation of Methods for Very Short-Term Load Forecasting Using Minute-By-Minute British Data“. In: *International Journal of Forecasting*, Bd. 24, Nr. 4 (2008), S. 645–658. DOI: [10.1016/j.ijforecast.2008.07.007](https://doi.org/10.1016/j.ijforecast.2008.07.007).
- [139] TAYLOR, J. W.: „Triple Seasonal Methods for Short-Term Electricity Demand Forecasting“. In: *European Journal of Operational Research*, Bd. 204, Nr. 1 (2010), S. 139–152. DOI: [10.1016/j.ejor.2009.10.003](https://doi.org/10.1016/j.ejor.2009.10.003).
- [140] BRETSCHNEIDER, P.: „Ein Beitrag zur Vorhersage musterbasierter nichtlinearer stochastischer Signale“. Dissertation. Fakultät für Information und Automatisierung, Technische Universität Ilmenau, 2002.
- [141] BRETSCHNEIDER, P. und RAUSCHENBACH, T.: „Ein Beitrag zur Vorhersage nichtlinearer stochastischer Signale (A Contribution to Forecast of Nonlinear Stochastic Signals)“. In: *at-Automatisierungstechnik Methoden und Anwendungen der Steuerungs-, Regelungs- und Informationstechnik*, Bd. 50, Nr. 7 (2002), S. 334. DOI: [10.1524/auto.2002.50.7.334](https://doi.org/10.1524/auto.2002.50.7.334).
- [142] WERNSTEDT, J.: *Experimentelle Prozeßanalyse*. Oldenbourg, München, Germany, 1989.
- [143] GRIMM, C.: „Evaluating Baselines for Demand Response Programs“. In: *AEIC Load Research Workshop*. (February 25-27, San Antonio, USA), 2008. URL: http://publications.aeic.org/lrc/12_Evaluating_Baselines_for_Demand_Response_Programs_PAPER.pdf (besucht am 13.02.2018).
- [144] WIJAYA, T.; VASIRANI, M. und ABERER, K.: „When Bias Matters: An Economic Assessment of Demand Response Baselines for Residential Customers“. In: *IEEE Transactions on Smart Grid*, Bd. 5, Nr. 4 (2014), S. 1755–1763. DOI: [10.1109/TSG.2014.2309053](https://doi.org/10.1109/TSG.2014.2309053).
- [145] KEMA: *PJM Empirical Analysis of Demand Response Baseline Methods*. Technischer Bericht. KEMA, Inc., 2011. URL: <https://www.pjm.com/%5C~/media/markets-ops/dsr/pjm-analysis-of-dr-baseline-methods-full-report.ashx> (besucht am 20.07.2017).
- [146] COUGHLIN, K.; PIETTE, M. A.; GOLDMAN, C. und KILICCOTE, S.: „Statistical Analysis of Baseline Load Models for Non-Residential Buildings“. In: *Energy and Buildings*, Bd. 41, Nr. 4 (2009), S. 374–381. DOI: [10.1016/j.enbuild.2008.11.002](https://doi.org/10.1016/j.enbuild.2008.11.002).
- [147] WACZOWICZ, S.; KLAIBER, S.; BRETSCHNEIDER, P.; KONOTOP, I.; WESTERMANN, D. und MIKUT, R.: „Virtuelle Speicher als adaptierbare Verbrauchermodelle zur Lastprognose und Betriebsführung in Verteilnetzen“. In: *Proc., 23. Workshop Computational Intelligence*. (December 5-6, Dortmund, Germany). KIT Scientific Publishing, Karlsruhe, Germany, 2013, S. 109–128.

- [148] SIMMONS-SÜER, B.; ATUKEREN, E. und BUSCH, C.: *Elastizitäten und Substitutionsmöglichkeiten der Elektrizitätsnachfrage: Literaturübersicht mit besonderem Fokus auf den Schweizer Strommarkt*. KOF Studien 26. Studie im Auftrag der Economiesuisse, 2011. URL: <http://hdl.handle.net/10419/54689> (besucht am 22. 05. 2014).
- [149] AALAMI, H.; MOGHADDAM, M. P. und YOUSEFI, G.: „Modeling and Prioritizing Demand Response Programs in Power Markets“. In: *Electric Power Systems Research*, Bd. 80, Nr. 4 (2010), S. 426–435. DOI: [10.1016/j.epsr.2009.10.007](https://doi.org/10.1016/j.epsr.2009.10.007).
- [150] MOGHADDAM, M. P.; ABDOLLAHI, A. und RASHIDINEJAD, M.: „Flexible Demand Response Programs Modeling in Competitive Electricity Markets“. In: *Applied Energy*, Bd. 88, Nr. 9 (2011), S. 3257–3269. DOI: [10.1016/j.apenergy.2011.02.039](https://doi.org/10.1016/j.apenergy.2011.02.039).
- [151] RUPPERT, M.; HAYN, M.; BERTSCH, V. und FICHTNER, W.: „Impact of Residential Electricity Tariffs with Variable Energy Prices on Low Voltage Grids with Photovoltaic Generation“. In: *International Journal of Electrical Power & Energy Systems*, Bd. 79 (2016), S. 161–171. DOI: [10.1016/j.ijepes.2016.01.017](https://doi.org/10.1016/j.ijepes.2016.01.017).
- [152] DIETRICH, K.; LATORRE, J.; OLMOS, L. und RAMOS, A.: „Demand Response in an Isolated System with High Wind Integration“. In: *IEEE Transactions on Power Systems*, Bd. 27, Nr. 1 (2012), S. 20–29. DOI: [10.1109/TPWRS.2011.2159252](https://doi.org/10.1109/TPWRS.2011.2159252).
- [153] EGHBAL, M.; SAHA, T. und MAHMOUDI-KOHAN, N.: „Utilizing Demand Response Programs in Day Ahead Generation Scheduling for Micro-Grids with Renewable Sources“. In: *IEEE PES Innovative Smart Grid Technologies - Asia*. (November 13-16, Perth, Australia). IEEE, Piscataway, NJ, USA, 2011, S. 1–6. DOI: [10.1109/ISGT-Asia.2011.6167148](https://doi.org/10.1109/ISGT-Asia.2011.6167148).
- [154] HAYN, M.; RINGLER, P.; RUPPERT, M.; SCHMID, J.; BERTSCH, V.; KELES, D. und FICHTNER, W.: „A Residential Price Elasticity Model for Agent-Based Simulation of Decentralised Energy Systems“. In: *Proc., SMARTER-Europe, E-World Energy & Water*. (February 13, Essen, Germany), 2014.
- [155] TUMULURU, V.; HUANG, Z. und TSANG, D.: „Integrating Price Responsive Demand Into the Unit Commitment Problem“. In: *IEEE Transactions on Smart Grid*, Bd. 5, Nr. 6 (2014), S. 2757–2765. DOI: [10.1109/TSG.2014.2331357](https://doi.org/10.1109/TSG.2014.2331357).
- [156] CRA: *Impact Evaluation of the California Statewide Pricing Pilot*. Technischer Bericht. Charles River Associates, 2005. URL: https://www.smartgrid.gov/files/Impact_Evaluation_California_Statewide_Pricing_Pilot_200501.pdf (besucht am 19. 09. 2017).
- [157] BABOLI, P.; EGHBAL, M.; MOGHADDAM, M. und AALAMI, H.: „Customer Behavior Based Demand Response Model“. In: *IEEE Power and Energy Society General Meeting*. (July 22-26, San Diego, USA). IEEE, Piscataway, NJ, USA, 2012, S. 1–7. DOI: [10.1109/pesgm.2012.6345101](https://doi.org/10.1109/pesgm.2012.6345101).
- [158] DORINI, G.; PINSON, P. und MADSEN, H.: „Chance-Constrained Optimization of Demand Response to Price Signals“. In: *IEEE Transactions on Smart Grid*, Bd. 4, Nr. 4 (2013), S. 2072–2080. DOI: [10.1109/TSG.2013.2258412](https://doi.org/10.1109/TSG.2013.2258412).

-
- [159] CORRADI, O.; OCHSENFELD, H.; MADSEN, H. und PINSON, P.: „Controlling Electricity Consumption by Forecasting Its Response to Varying Prices“. In: *IEEE Transactions on Power Systems*, Bd. 28, Nr. 1 (2013), S. 421–429. DOI: [10.1109/TPWRS.2012.2197027](https://doi.org/10.1109/TPWRS.2012.2197027).
- [160] CONEJO, A.; MORALES, J. und BARINGO, L.: „Real-Time Demand Response Model“. In: *IEEE Transactions on Smart Grid*, Bd. 1, Nr. 3 (2010), S. 236–242. DOI: [10.1109/TSG.2010.2078843](https://doi.org/10.1109/TSG.2010.2078843).
- [161] CHEN, Z. und WU, L.: „Residential Appliance DR Energy Management with Electric Privacy Protection by Online Stochastic Optimization“. In: *IEEE Transactions on Smart Grid*, Bd. 4, Nr. 4 (2013), S. 1861–1869. DOI: [10.1109/TSG.2013.2256803](https://doi.org/10.1109/TSG.2013.2256803).
- [162] TSUI, K. M. und CHAN, S.: „Demand Response Optimization for Smart Home Scheduling Under Real-Time Pricing“. In: *IEEE Transactions on Smart Grid*, Bd. 3, Nr. 4 (2012), S. 1812–1821. DOI: [10.1109/TSG.2012.2218835](https://doi.org/10.1109/TSG.2012.2218835).
- [163] HOLTSCHEIDER, T. und ERLICH, I.: „Modeling Demand Response of Consumers to Incentives Using Fuzzy Systems“. In: *IEEE Power and Energy Society General Meeting*. (July 22-26, San Diego, USA). IEEE, Piscataway, NJ, USA, 2012, S. 1–8. DOI: [10.1109/pesgm.2012.6345280](https://doi.org/10.1109/pesgm.2012.6345280).
- [164] CHEN, Y.; LIN, W.; HAN, F.; YANG, Y.-H.; SAFAR, Z. und LIU, K.: „A Cheat-Proof Game Theoretic Demand Response Scheme for Smart Grids“. In: *IEEE International Conference on Communications (ICC)*. (June 10-15, Ottawa, Canada). IEEE, Piscataway, NJ, USA, 2012, S. 3362–3366. DOI: [10.1109/ICC.2012.6364397](https://doi.org/10.1109/ICC.2012.6364397).
- [165] DENG, R.; YANG, Z.; CHEN, J.; ASR, N. und CHOW, M.-Y.: „Residential Energy Consumption Scheduling: A Coupled-Constraint Game Approach“. In: *IEEE Transactions on Smart Grid*, Bd. 5, Nr. 3 (2014), S. 1340–1350. DOI: [10.1109/TSG.2013.2287494](https://doi.org/10.1109/TSG.2013.2287494).
- [166] MAHARJAN, S.; ZHU, Q.; ZHANG, Y.; GJESSING, S. und BASAR, T.: „Dependable Demand Response Management in the Smart Grid: A Stackelberg Game Approach“. In: *IEEE Transactions on Smart Grid*, Bd. 4, Nr. 1 (2013), S. 120–132. DOI: [10.1109/TSG.2012.2223766](https://doi.org/10.1109/TSG.2012.2223766).
- [167] LOESCH, M.; HUFNAGEL, D.; STEUER, S.; FABNACHT, T. und SCHMECK, H.: „Demand Side Management in Smart Buildings by Intelligent Scheduling of Heat Pumps“. In: *IEEE International Conference on Intelligent Energy and Power Systems (IEPS)*. (June 2-6, Kiev, Ukraine). IEEE, Piscataway, NJ, USA, 2014, S. 1–6. DOI: [10.1109/ieps.2014.6874181](https://doi.org/10.1109/ieps.2014.6874181).
- [168] MAUSER, I.; MÜLLER, J.; ALLERDING, F. und SCHMECK, H.: „Adaptive Building Energy Management with Multiple Commodities and Flexible Evolutionary Optimization“. In: *Renewable Energy*, Bd. 87, Nr. 2 (2016), S. 911–921. DOI: [10.1016/j.renene.2015.09.003](https://doi.org/10.1016/j.renene.2015.09.003).
- [169] VARDAKAS, J. S.; ZORBA, N. und VERIKOUKIS, C. V.: „A Survey on Demand Response Programs in Smart Grids: Pricing Methods and Optimization Algorithms“. In: *IEEE Communications Surveys Tutorials*, Bd. 17, Nr. 1 (2015), S. 152–178. DOI: [10.1109/COMST.2014.2341586](https://doi.org/10.1109/COMST.2014.2341586).

- [170] O'NEILL, D.; LEVORATO, M.; GOLDSMITH, A. und MITRA, U.: „Residential Demand Response Using Reinforcement Learning“. In: *IEEE International Conference on Smart Grid Communications (SmartGridComm)*. (October 4-6, Gaithersburg, USA). IEEE, Piscataway, NJ, USA, 2010, S. 409–414. DOI: [10.1109/SMARTGRID.2010.5622078](https://doi.org/10.1109/SMARTGRID.2010.5622078).
- [171] SONG, M.; ALVEHAG, K.; WIDÉN, J. und PARISIO, A.: „Estimating the Impacts of Demand Response by Simulating Household Behaviours under Price and CO2 Signals“. In: *Electric Power Systems Research*, Bd. 111 (2014), S. 103–114. DOI: [10.1016/j.epsr.2014.02.016](https://doi.org/10.1016/j.epsr.2014.02.016).
- [172] SHI, W. und WONG, V.: „Real-Time Vehicle-To-Grid Control Algorithm Under Price Uncertainty“. In: *IEEE International Conference on Smart Grid Communications (SmartGridComm)*. (October 17-20, Brussels, Belgium). IEEE, Piscataway, NJ, USA, 2011, S. 261–266. DOI: [10.1109/SmartGridComm.2011.6102330](https://doi.org/10.1109/SmartGridComm.2011.6102330).
- [173] SHAO, S.; PIPATTANASOMPORN, M. und RAHMAN, S.: „Development of Physical-Based Demand Response-Enabled Residential Load Models“. In: *IEEE Transactions on Power Systems*, Bd. 28, Nr. 2 (2013), S. 607–614. DOI: [10.1109/TPWRS.2012.2208232](https://doi.org/10.1109/TPWRS.2012.2208232).
- [174] MOLINA-GARCÍA, A.; KESSLER, M.; FUENTES, J. und GÓMEZ-LÁZARO, E.: „Probabilistic Characterization of Thermostatically Controlled Loads to Model the Impact of Demand Response Programs“. In: *IEEE Transactions on Power Systems*, Bd. 26, Nr. 1 (2011), S. 241–251. DOI: [10.1109/TPWRS.2010.2047659](https://doi.org/10.1109/TPWRS.2010.2047659).
- [175] PALENSKY, P.; KUPZOG, F.; ZAIDI, A. und ZHOU, K.: „Modeling Domestic Housing Loads for Demand Response“. In: *34th Annual Conference of IEEE Industrial Electronics*. (November 10-13, Orlando, USA). IEEE, Piscataway, NJ, USA, 2008, S. 2742–2747. DOI: [10.1109/IECON.2008.4758392](https://doi.org/10.1109/IECON.2008.4758392).
- [176] JAIN, A. K.: „Data Clustering: 50 Years Beyond K-means“. In: *Pattern Recognition Letters*, Bd. 31, Nr. 8 (2010), S. 651–666. DOI: [10.1016/j.patrec.2009.09.011](https://doi.org/10.1016/j.patrec.2009.09.011).
- [177] LIAO WARREN, T.: „Clustering of Time Series Data - A Survey“. In: *Pattern Recognition*, Bd. 38, Nr. 11 (2005), S. 1857–1874.
- [178] AGHABOZORGI, S.; SHIRKHORSHIDI, A. und WAH, T.: „Time-Series Clustering - A Decade Review“. In: *Information Systems*, Bd. 53 (2015), S. 16–38. DOI: [10.1016/j.is.2015.04.007](https://doi.org/10.1016/j.is.2015.04.007).
- [179] ZHOU, D.; BALANDAT, M. und TOMLIN, C.: „Residential Demand Response Targeting Using Machine Learning with Observational Data“. In: *CoRR*, Bd. abs/1607.00595 (2016). DOI: [10.1109/cdc.2016.7799295](https://doi.org/10.1109/cdc.2016.7799295).
- [180] HABEN, S.; SINGLETON, C. und GRINDROD, P.: „Analysis and Clustering of Residential Customers Energy Behavioral Demand Using Smart Meter Data“. In: *IEEE Transactions on Smart Grid*, Bd. 7, Nr. 1 (2016), S. 136–144. DOI: [10.1109/TSG.2015.2409786](https://doi.org/10.1109/TSG.2015.2409786).
- [181] MACEDO, M.; GALO, J.; ALMEIDA, L. DE und C. LIMA, A. DE: „Demand Side Management using Artificial Neural Networks in a Smart Grid Environment“. In: *Renewable and Sustainable Energy Reviews*, Bd. 41 (2015), S. 128–133. DOI: [10.1016/j.rser.2014.08.035](https://doi.org/10.1016/j.rser.2014.08.035).

-
- [182] KWAC, J.; FLORA, J. und RAJAGOPAL, R.: „Household Energy Consumption Segmentation Using Hourly Data“. In: *IEEE Transactions on Smart Grid*, Bd. 5, Nr. 1 (2014), S. 420–430. DOI: [10.1109/TSG.2013.2278477](https://doi.org/10.1109/TSG.2013.2278477).
- [183] LABEEUW, W. und DECONINCK, G.: „Residential Electrical Load Model Based on Mixture Model Clustering and Markov Models“. In: *IEEE Transactions on Industrial Informatics*, Bd. 9, Nr. 3 (2013), S. 1561–1569. DOI: [10.1109/TII.2013.2240309](https://doi.org/10.1109/TII.2013.2240309).
- [184] WIJAYA, T. K.; GANU, T.; CHAKRABORTY, D.; ABERER, K. und SEETHARAM, D. P.: „Consumer Segmentation and Knowledge Extraction from Smart Meter and Survey Data“. In: *Proc., SIAM International Conference on Data Mining*. (April 24–26, Philadelphia, USA). SIAM Publications Online, Philadelphia, PA, USA, 2014, S. 226–234. DOI: [10.1137/1.9781611973440.26](https://doi.org/10.1137/1.9781611973440.26).
- [185] BENÍTEZ, I.; QUIJANO, A.; DíEZ, J.-L. und DELGADO, I.: „Dynamic Clustering Segmentation Applied to Load Profiles of Energy Consumption from Spanish Customers“. In: *International Journal of Electrical Power & Energy Systems*, Bd. 55 (2014), S. 437–448. DOI: [10.1016/j.ijepes.2013.09.022](https://doi.org/10.1016/j.ijepes.2013.09.022).
- [186] CHICCO, G. und ILIE, I.-S.: „Support Vector Clustering of Electrical Load Pattern Data“. In: *IEEE Transactions on Power Systems*, Bd. 24, Nr. 3 (2009), S. 1619–1628. DOI: [10.1109/tpwrs.2009.2023009](https://doi.org/10.1109/tpwrs.2009.2023009).
- [187] RHODES, J. D.; COLE, W. J.; UPSHAW, C. R.; EDGAR, T. F. und WEBBER, M. E.: „Clustering Analysis of Residential Electricity Demand Profiles“. In: *Applied Energy*, Bd. 135 (2014), S. 461–471. DOI: [10.1016/j.apenergy.2014.08.111](https://doi.org/10.1016/j.apenergy.2014.08.111).
- [188] McLOUGHLIN, F.; DUFFY, A. und CONLON, M.: „A Clustering Approach to Domestic Electricity Load Profile Characterisation using Smart Metering Data“. In: *Applied Energy*, Bd. 141 (2015), S. 190–199. DOI: [10.1016/j.apenergy.2014.12.039](https://doi.org/10.1016/j.apenergy.2014.12.039).
- [189] DIAO, L.; SUN, Y.; CHEN, Z. und CHEN, J.: „Modeling Energy Consumption in Residential Buildings: A Bottom-Up Analysis Based on Occupant Behavior Pattern Clustering and Stochastic Simulation“. In: *Energy and Buildings*, Bd. 147 (2017), S. 47–66. DOI: [10.1016/j.enbuild.2017.04.072](https://doi.org/10.1016/j.enbuild.2017.04.072).
- [190] BOCK, C.: „Clustering-Ansatz zur Erstellung von Lastprofilen zur Vorhersage des Stromverbrauchs“. In: *Proc., 28th GI-Workshop on Foundations of Databases (Grundlagen von Datenbanken)*. (May 24–27, Nörten-Hardenberg, Germany). CEUR-WS.org, 2016, S. 21–26. URL: <http://ceur-ws.org/Vol-1594/paper5.pdf> (besucht am 13.02.2018).
- [191] BOCK, C.: „Generating Load Profiles Using Smart Metering Time Series“. In: *Advances in Fuzzy Logic and Technology 2017: Proceedings of: EUSFLAT-2017 – The 10th Conference of the European Society for Fuzzy Logic and Technology; IWIFSGN’2017 – The Sixteenth International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets*. Hrsg. von KACPRZYK, J.; SZMIDT, E.; ZADROŻNY, S.; ATANASSOV, K. T. und KRAWCZAK, M. Bd. 1. (September 11–15, Warsaw, Poland). Springer International Publishing, Cham, Switzerland, 2018, S. 211–223. DOI: [10.1007/978-3-319-66830-7_20](https://doi.org/10.1007/978-3-319-66830-7_20).
- [192] FLATH, C.; NICOLAY, D.; CONTE, T.; DINTHER, C. VAN und FILIPOVA-NEUMANN, L.: „Cluster Analysis of Smart Metering Data“. In: *Business & Information Systems Engineering*, Bd. 4, Nr. 1 (2012), S. 31–39. DOI: [10.1007/s12599-011-0201-5](https://doi.org/10.1007/s12599-011-0201-5).

- [193] KIM, Y.-I.; KO, J.-M. und CHOI, S.-H.: „Methods for Generating TLPs (Typical Load Profiles) for Smart Grid-Based Energy Programs“. In: *IEEE Symposium on Computational Intelligence Applications in Smart Grid (CIASG)*. (April 11-15, Paris, France). IEEE, Piscataway, NJ, USA, 2011, S. 1–6. DOI: [10.1109/CIASG.2011.5953331](https://doi.org/10.1109/CIASG.2011.5953331).
- [194] KIM, Y.-I.; KANG, S.-J.; KO, J.-M. und CHOI, S.-H.: „A Study for Clustering Method to Generate Typical Load Profiles for Smart Grid“. In: *IEEE 8th International Conference on Power Electronics and ECCE Asia (ICPE ECCE)*. (May 30 - June 3, Jeju, South Korea). IEEE, Piscataway, NJ, USA, 2011, S. 1102–1109. DOI: [10.1109/ICPE.2011.5944675](https://doi.org/10.1109/ICPE.2011.5944675).
- [195] CHICCO, G.: „Overview and Performance Assessment of the Clustering Methods for Electrical Load Pattern Grouping“. In: *Energy*, Bd. 42, Nr. 1 (2012), S. 68–80. DOI: [10.1016/j.energy.2011.12.031](https://doi.org/10.1016/j.energy.2011.12.031).
- [196] BANDA, E. und LE, T.: „Modeling of Demand Response in Electricity Markets: Effects of Price Elasticity“. In: *Proc., 9th IASTED International Conference on Power and Energy Systems*. (January 3-5, Clearwater, USA). ACTA Press, Calgary, Canada, 2007.
- [197] DATCHANAMOORTHY, S.; KUMAR, S.; OZTURK, Y. und LEE, G.: „Optimal Time-Of-Use Pricing for Residential Load Control“. In: *IEEE International Conference on Smart Grid Communications (SmartGridComm)*. (October 17-20, Brussels, Belgium). IEEE, Piscataway, NJ, USA, 2011, S. 375–380. DOI: [10.1109/SmartGridComm.2011.6102350](https://doi.org/10.1109/SmartGridComm.2011.6102350).
- [198] DIETRICH, K.; LATORRE, J. M.; OLMOS, L. und RAMOS, A.: „Demand Response and Its Sensitivity to Participation Rates and Elasticities“. In: *8th International Conference on the European Energy Market (EEM)*. (May 25-27, Zagreb, Croatia). IEEE, Piscataway, NJ, USA, 2011, S. 717–716. DOI: [10.1109/eem.2011.5953103](https://doi.org/10.1109/eem.2011.5953103).
- [199] FILIPPINI, M.: „Swiss Residential Demand for Electricity by Time-Of-Use: An Application of the Almost Ideal Demand System“. In: *Energy Journal*, Bd. 16, Nr. 1 (1995), S. 27–39. DOI: [10.5547/issn0195-6574-ej-vol16-no1-2](https://doi.org/10.5547/issn0195-6574-ej-vol16-no1-2).
- [200] FILIPPINI, M. und PACHAURI, S.: „Elasticities of Electricity Demand in Urban Indian Households“. In: *Energy Policy*, Bd. 32, Nr. 3 (2004), S. 429–436. DOI: [10.1016/S0301-4215\(02\)00314-2](https://doi.org/10.1016/S0301-4215(02)00314-2).
- [201] YUSTA, J.; KHODR, H. und URDANETA, A.: „Optimal Pricing of Default Customers in Electrical Distribution Systems: Effect Behavior Performance of Demand Response Models“. In: *Electric Power Systems Research*, Bd. 77 (2007), S. 548–558. DOI: [10.1016/j.epsr.2006.05.001](https://doi.org/10.1016/j.epsr.2006.05.001).
- [202] CORRADI, O. und OCHSENFELD, H.: „Integration of Fluctuating Energy by Electricity Price Control“. Masterarbeit. Technical University Denmark, Lyngby, Denmark, 2011.
- [203] LI, S.; ZHANG, D.; ROGET, A. und O’NEILL, Z.: „Integrating Home Energy Simulation and Dynamic Electricity Price for Demand Response Study“. In: *IEEE Transactions on Smart Grid*, Bd. 5, Nr. 2 (2014), S. 779–788. DOI: [10.1109/TSG.2013.2279110](https://doi.org/10.1109/TSG.2013.2279110).
- [204] OZTURK, Y.; SENTHILKUMAR, D.; KUMAR, S. und LEE, G.: „An Intelligent Home Energy Management System to Improve Demand Response“. In: *IEEE Transactions on Smart Grid*, Bd. 4, Nr. 2 (2013), S. 694–701. DOI: [10.1109/TSG.2012.2235088](https://doi.org/10.1109/TSG.2012.2235088).

-
- [205] SAEZ-GALLEGO, J.; MORALES, J. M.; ZUGNO, M. und MADSEN, H.: „A Data-Driven Bidding Model for a Cluster of Price-responsive Consumers of Electricity“. In: *IEEE Transactions on Power Systems* (2016). DOI: [10.1109/tpwrs.2016.2530843](https://doi.org/10.1109/tpwrs.2016.2530843).
- [206] WU, Z.; ZHOU, S.; LI, J. und ZHANG, X.-P.: „Real-Time Scheduling of Residential Appliances Via Conditional Risk-At-Value“. In: *IEEE Transactions on Smart Grid*, Bd. 5, Nr. 3 (2014), S. 1282–1291. DOI: [10.1109/TSG.2014.2304961](https://doi.org/10.1109/TSG.2014.2304961).
- [207] YANG, L.; DONG, C.; WAN, C. J. und NG, C. T.: „Electricity Time-Of-Use Tariff with Consumer Behavior Consideration“. In: *International Journal of Production Economics*, Bd. 146, Nr. 2 (2013), S. 402–410. DOI: [10.1016/j.ijpe.2013.03.006](https://doi.org/10.1016/j.ijpe.2013.03.006).
- [208] GHOSH, S.; SUN, X. und ZHANG, X.: „Consumer Profiling for Demand Response Programs in Smart Grids“. In: *IEEE Innovative Smart Grid Technologies - Asia*. (May 21-24, Tianjin, China). IEEE, Piscataway, NJ, USA, 2012, S. 1–6. DOI: [10.1109/ISGT-Asia.2012.6303309](https://doi.org/10.1109/ISGT-Asia.2012.6303309).
- [209] WACZOWICZ, S.; REISCHL, M.; HAGENMEYER, V.; MIKUT, R.; KLAIBER, S.; BRETSCHNEIDER, P.; KONOTOP, I. und WESTERMANN, D.: „Demand Response Clustering - How do Dynamic Prices Affect Household Electricity Consumption?“ In: *Proc., IEEE PowerTech*. (June 29 - July 2, Eindhoven, Netherlands). IEEE, Piscataway, NJ, USA, 2015, S. 1–6. DOI: [10.1109/ptc.2015.7232493](https://doi.org/10.1109/ptc.2015.7232493).
- [210] KLAIBER, S.; WACZOWICZ, S.; MIKUT, R.; KONOTOP, I.; WESTERMANN, D. und BRETSCHNEIDER, P.: „A Contribution to the Load Forecast of Price Elastic Consumption Behavior“. In: *Proc., IEEE PowerTech*. (June 29 - July 2, Eindhoven, Netherlands). IEEE, Piscataway, NJ, USA, 2015.
- [211] BRETSCHNEIDER, P.: „Forschungs- und Entwicklungsarbeiten für die Planung und Betriebsführung von Energieversorgungssystemen unter Berücksichtigung Erneuerbarer Energien, Fraunhofer IOSB, Institutsteil Angewandte Systemtechnik IOSB-AST“. (April 22, 2016), Präsentation in Karlsruhe, Germany. 2016.
- [212] SCHUFFT, W.: *Taschenbuch der elektrischen Energietechnik*. Hanser, München, Germany, 2007.
- [213] KONOTOP, I.; WACZOWICZ, S.; KLAIBER, S.; BRETSCHNEIDER, P.; MIKUT, R. und WESTERMANN, D.: „Investigation and Mathematical Modelling of the Impact of Incentive Signals to Consumers on their Consumption, Load Forecast and Network Operation“. In: *IET Generation, Transmission & Distribution*, Bd. 10, Nr. 9 (2016), S. 2138–2146. DOI: [10.1049/iet-gtd.2015.1175](https://doi.org/10.1049/iet-gtd.2015.1175).
- [214] SCHNORR, F. und HINZE, H.: *Erstellung von Lastprognosen für den elektrischen Strombedarf von Einfamilienhäusern*. Technischer Bericht. Projektarbeit im Studiengang Regenerative Energien an der Hochschule für Technik und Wirtschaft Berlin, 2014.
- [215] KLAIBER, S.; WACZOWICZ, S.; KONOTOP, I.; WESTERMANN, D.; MIKUT, R. und BRETSCHNEIDER, P.: „Prognose für preisbeeinflusstes Verbrauchsverhalten“. In: *at-Automatisierungstechnik*, Bd. 65, Nr. 3 (2017), S. 179–188. DOI: [10.1515/auto-2016-0125](https://doi.org/10.1515/auto-2016-0125).

- [216] BARANSKI, M. und VOSS, J.: „Genetic Algorithm for Pattern Detection in NIALM Systems“. In: *IEEE International Conference on Systems, Man and Cybernetics*. Bd. 4. (October 10-13, The Hague, Netherlands). IEEE, Piscataway, NJ, USA, 2004, S. 3462–3468. DOI: [10.1109/ICSMC.2004.1400878](https://doi.org/10.1109/ICSMC.2004.1400878).
- [217] HART, G.: „Nonintrusive Appliance Load Monitoring“. In: Bd. 80, Nr. 12 (1992), S. 1870–1891. DOI: [10.1109/5.192069](https://doi.org/10.1109/5.192069).
- [218] ZEIFMAN, M. und ROTH, K.: „Nonintrusive Appliance Load Monitoring: Review and Outlook“. In: *IEEE Transactions on Consumer Electronics*, Bd. 57, Nr. 1 (2011), S. 76–84. DOI: [10.1109/TCE.2011.5735484](https://doi.org/10.1109/TCE.2011.5735484).
- [219] BERGMAN, D.; JIN, D.; JUEN, J.; TANAKA, N.; GUNTER, C. und WRIGHT, A.: „Distributed Non-Intrusive Load Monitoring“. In: *IEEE PES Innovative Smart Grid Technologies*. (January 17-19, Anaheim, USA). IEEE, Piscataway, NJ, USA, 2011, S. 1–8. DOI: [10.1109/ISGT.2011.5759180](https://doi.org/10.1109/ISGT.2011.5759180).
- [220] VOGIATZIS, E.; KALOGRIDIS, G. und DENIC, S.: „Real-Time and Low Cost Energy Disaggregation of Coarse Meter Data“. In: *4th IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)*. (October 6-9, Lyngby, Denmark). IEEE, Piscataway, NJ, USA, 2013, S. 1–5. DOI: [10.1109/ISGTEurope.2013.6695417](https://doi.org/10.1109/ISGTEurope.2013.6695417).
- [221] KIM, H.; MARWAH, M.; ARLITT, M. F.; LYON, G. und HAN, J.: „Unsupervised Disaggregation of Low Frequency Power Measurements.“ In: *Proc., SIAM International Conference on Data Mining*. Bd. 11. (April 28-30, Phoenix, USA). SIAM Publications Online, Philadelphia, PA, USA, 2011, S. 747–758. DOI: [10.1137/1.9781611972818.64](https://doi.org/10.1137/1.9781611972818.64).
- [222] KOLTER, J. Z. und JAAKKOLA, T.: „Approximate Inference in Additive Factorial HMMs with Application to Energy Disaggregation“. In: *Proc., Fifteenth International Conference on Artificial Intelligence and Statistics*. Hrsg. von LAWRENCE, N. D. und GIROLAMI, M. Bd. 22. Proceedings of Machine Learning Research. (April 21-23, La Palma, Canary Islands). PMLR, 2012, S. 1472–1482. URL: <http://proceedings.mlr.press/v22/zico12.html> (besucht am 13.02.2018).
- [223] PARSON, O.; GHOSH, S.; WEAL, M. und ROGERS, A.: „Non-Intrusive Load Monitoring using Prior Models of General Appliance Types“. In: *Proc., 26th Conference on Artificial Intelligence (AAAI)*. (July 22-26, Toronto, Canada). AAAI press, Palo Alto, CA, USA, 2012, S. 356–362. URL: <http://eprints.soton.ac.uk/336812/> (besucht am 13.02.2018).
- [224] EIBL, G. und ENGEL, D.: „Influence of Data Granularity on Smart Meter Privacy“. In: *IEEE Transactions on Smart Grid*, Bd. 6, Nr. 2 (2015), S. 930–939. DOI: [10.1109/TSG.2014.2376613](https://doi.org/10.1109/TSG.2014.2376613).
- [225] BDEW: *Energie-Info: Stromverbrauch im Haushalt*. Technischer Bericht. BDEW Bundesverband der Energie- und Wasserwirtschaft e.V., 2014.
- [226] KLAIBER, S.: „Prognoseverfahren für beeinflusstes Verbrauchsverhalten in Energiesystemen“. Dissertation. (voraussichtliche Einreichung in 2018 an der Technischen Universität Ilmenau), 2018.
- [227] PFLUGRADT, N. und PLATZER, B.: „Behavior Based Load Profile Generator for Domestic Hot Water and Electricity Use“. In: (May 16-18, Lleida, Spain), 2012.

-
- [228] PFLUGRADT, N. und PLATZER, B.: „Verhaltensbasierter Lastprofilgenerator für Strom- und Warmwasser-Profile“. In: *22. Symposium "Thermische Solarenergie", Ostbayerisches Technologie Transfer Institut e.V. (OTTI)* (2012). Hrsg. von OSTBAYERISCHES TECHNOLOGIE TRANSFER INSTITUT E.V. (OTTI), R., S. 250–251.
- [229] PFLUGRADT, N.: „Modellierung von Wasser und Energieverbräuchen in Haushalten“. Dissertation. Technische Universität Chemnitz, 2016.
- [230] PFLUGRADT, N.: *LoadProfileGenerator*. 2016. URL: <http://www.loadprofilegenerator.de/> (besucht am 22. 12. 2016).
- [231] TJADEN, T.; BERGNER, J.; WENIGER, J.; QUASCHNING, V. und SOLARSPEICHERSYSTEME, F.: *Repräsentative elektrische Lastprofile für Wohngebäude in Deutschland auf 1-sekündiger Datenbasis*. Technischer Bericht. 2015, S. 8.
- [232] LIU, H.; SHAH, S. und JIANG, W.: „On-Line Outlier Detection and Data Cleaning“. In: *Computers & chemical engineering*, Bd. 28, Nr. 9 (2004), S. 1635–1647. DOI: [10.1016/j.compchemeng.2004.01.009](https://doi.org/10.1016/j.compchemeng.2004.01.009).
- [233] RATNAM, E. L.; WELLER, S. R.; KELLETT, C. M. und MURRAY, A. T.: „Residential Load and Rooftop PV Generation: an Australian Distribution Network Dataset“. In: *International Journal of Sustainable Energy* (2015), S. 1–20. DOI: [10.1080/14786451.2015.1100196](https://doi.org/10.1080/14786451.2015.1100196).
- [234] ARNOLDT, A.; KÖNIG, S.; MIKUT, R. und BRETSCHNEIDER, P.: „Application of Data Mining Methods for Power Forecast of Wind Power Plants“. In: *Proc., 9th International Workshop on Large-scale Integration of Wind Power and Transmission Networks for Offshore Wind Farms*. (October 18-19, Quebec, Canada), 2010.
- [235] UNIEJEWSKI, B.; WERON, R. und ZIEL, F.: *Variance Stabilizing Transformations for Electricity Spot Price Forecasting*. HSC Research Reports HSC/17/01. Hugo Steinhaus Center, Wroclaw University of Technology, 2017. DOI: [10.1109/tpwrs.2017.2734563](https://doi.org/10.1109/tpwrs.2017.2734563).
- [236] KEIM, D.; ANDRIENKO, G.; FEKETE, J.-D.; GÖRG, C.; KOHLHAMMER, J. und MELANÇON, G.: „Visual Analytics: Definition, Process, and Challenges“. In: *Information Visualization*. Hrsg. von KERREN, A.; STASKO, J.; FEKETE, J.-D. und NORTH, C. Bd. 4950. Lecture Notes on Computer Science. Springer, Berlin, Heidelberg, Germany, 2008. DOI: [10.1007/978-3-540-70956-5](https://doi.org/10.1007/978-3-540-70956-5).
- [237] COOK, K. A. und THOMAS, J. J.: *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society, Los Alamitos, CA, United States, 2005.
- [238] MUNSHI, A. A. und MOHAMED, Y. A. I.: „Cloud-Based Visual Analytics for Smart Grids Big Data“. In: *IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. (September 6-9, Minneapolis, USA). IEEE, Piscataway, NJ, USA, 2016, S. 1–5. DOI: [10.1109/ISGT.2016.7781163](https://doi.org/10.1109/ISGT.2016.7781163).
- [239] JANETZKO, H.; STOFFEL, F.; MITTELSTÄDT, S. und KEIM, D. A.: „Anomaly Detection for Visual Analytics of Power Consumption Data“. In: *Computers & Graphics*, Bd. 38, Nr. Supplement C (2014), S. 27–37. DOI: [10.1016/j.cag.2013.10.006](https://doi.org/10.1016/j.cag.2013.10.006).

- [240] ESSLINGER, P. und WITZMANN, R.: „Entwicklung und Verifikation eines stochastischen Verbraucherlastmodells für Haushalte“. In: *12. Symposium Energieinnovation*. (February 15-17, Graz, Austria). Verlag der Technischen Universität Graz, Graz, Austria, 2012, S. 1–11.
- [241] RUNKLER, T. A.: *Data Analytics: Models and Algorithms for Intelligent Data Analysis*. Springer Vieweg, Wiesbaden, Germany, 2016.
- [242] BEZDEK, J. C.: *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York, USA, 1981.
- [243] HÖPPNER, F.; KLAWONN, F. und KRUSE, R.: *Fuzzy Cluster Analysis*. John Wiley, New York, USA, 1999.
- [244] FLETLING, R.: „Fuzzy-Clusterverfahren zur Datenanalyse“. In: *Theorie und Anwendung lernender Algorithmen in den Ingenieurs- und Naturwissenschaften an der TU Braunschweig, Geodätische Schriftenreihe Heft Nr. 25*. Hrsg. von HEINERT, M. und RIEDEL, B. 2010.
- [245] KROLL, A.: *Computational Intelligence: Eine Einführung in Probleme, Methoden und technische Anwendungen*. Oldenbourg, München, Germany, 2013.
- [246] DAVÉ, R. und SEN, S.: „Robust Fuzzy Clustering of Relational Data“. In: *IEEE Transactions on Fuzzy Systems*, Bd. 10, Nr. 6 (2002), S. 713–727. DOI: [10.1109/TFUZZ.2002.805899](https://doi.org/10.1109/TFUZZ.2002.805899).
- [247] REHM, F.; KLAWONN, F. und KRUSE, R.: „A Novel Approach to Noise Clustering for Outlier Detection“. In: *Soft Computing*, Bd. 11, Nr. 5 (2007), S. 489–494. DOI: [10.1007/s00500-006-0112-4](https://doi.org/10.1007/s00500-006-0112-4).
- [248] CIMINO, M. G. C. A.; FROSINI, G.; LAZZERINI, B. und MARCELLONI, F.: „On the Noise Distance in Robust Fuzzy C-Means“. In: *International Journal of Computer Science and Engineering*, Bd. 1, Nr. 1 (2007), S. 124–127.
- [249] DAVÉ, R. N.: „Characterization and Detection of Noise in Clustering“. In: *Pattern Recognition Letters*, Bd. 12, Nr. 11 (1991), S. 657–664. DOI: [10.1016/0167-8655\(91\)90002-4](https://doi.org/10.1016/0167-8655(91)90002-4).
- [250] KLAWONN, F.: „Noise Clustering with a Fixed Fraction of Noise“. In: *Applications and Science in Soft Computing*. Hrsg. von LOTFI, A. und GARIBALDI, J. M. 24 Bde. Springer, Berlin, Heidelberg, Germany, 2004, S. 133–138. DOI: [10.1007/978-3-540-45240-9_19](https://doi.org/10.1007/978-3-540-45240-9_19).
- [251] LIN, J.; VLACHOS, M.; KEOGH, E. und GUNOPULOS, D.: „Iterative Incremental Clustering of Time Series“. In: *Advances in Database Technology-EDBT (2004)*, S. 521–522. DOI: [10.1007/978-3-540-24741-8_8](https://doi.org/10.1007/978-3-540-24741-8_8).
- [252] ZHANG, H.; HO, T. B.; ZHANG, Y. und LIN, M.-S.: „Unsupervised Feature Extraction for Time Series Clustering Using Orthogonal Wavelet Transform“. In: *Informatica*, Bd. 30, Nr. 3 (2006).

-
- [253] RATANAMAHAATANA, C.; KEOGH, E.; BAGNALL, A. J. und LONARDI, S.: „A Novel Bit Level Time Series Representation with Implication of Similarity Search and Clustering“. In: *Proc., Advances in Knowledge Discovery and Data Mining: 9th Pacific-Asia Conference (PAKDD)*. Hrsg. von HO, T. B.; CHEUNG, D. und LIU, H. (May 18-20, Hanoi, Vietnam). Springer, Berlin, Heidelberg, Germany, 2005, S. 771–777. DOI: [10.1007/11430919_90](https://doi.org/10.1007/11430919_90).
- [254] KAMEDA, S. und YAMAMURA, M.: „Spider Algorithm for Clustering Time Series“. In: *Proc., 5th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases*. (February 15-17, Madrid, Spain). World Scientific, Engineering Academy und Society (WSEAS), Stevens Point, Wisconsin, USA, 2006, S. 378–383.
- [255] CHIŞ, M.; BANERJEE, S. und HASSANIEN, A.: „Clustering Time Series Data: An Evolutionary Approach“. In: *Foundations of Computational Intelligence*, Bd. 6 (2009), S. 193–207. DOI: [10.1007/978-3-642-01091-0_9](https://doi.org/10.1007/978-3-642-01091-0_9).
- [256] CALIŃSKI, T. und HARABASZ, J.: „A Dendrite Method for Cluster Analysis“. In: *Communications in Statistics*, Bd. 3, Nr. 1 (1974), S. 1–27. DOI: [10.1080/03610927408827101](https://doi.org/10.1080/03610927408827101).
- [257] ROUSSEEUW, P.: „Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis“. In: *Journal of Computational and Applied Mathematics*, Bd. 20 (1987), S. 53–65. DOI: [10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).
- [258] DUNN, J. C.: „Well-Separated Clusters and Optimal Fuzzy Partitions“. In: *Journal of Cybernetics*, Bd. 4, Nr. 1 (1974), S. 95–104. DOI: [10.1080/01969727408546059](https://doi.org/10.1080/01969727408546059).
- [259] HALKIDI, M. und VAZIRGIANNIS, M.: „Clustering Validity Assessment: Finding the Optimal Partitioning of a Data Set“. In: *Proc., IEEE International Conference on Data Mining*. (November 29 - December 2, San Jose, USA). IEEE, Piscataway, NJ, USA, 2001, S. 187–194. DOI: [10.1109/ICDM.2001.989517](https://doi.org/10.1109/ICDM.2001.989517).
- [260] XIE, X. und BENI, G.: „A Validity Measure for Fuzzy Clustering“. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Bd. 13, Nr. 8 (1991), S. 841–847. DOI: [10.1109/34.85677](https://doi.org/10.1109/34.85677).
- [261] DAVIES, D. L. und BOULDIN, D. W.: „A Cluster Separation Measure“. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Bd. PAMI-1, Nr. 2 (1979), S. 224–227. DOI: [10.1109/TPAMI.1979.4766909](https://doi.org/10.1109/TPAMI.1979.4766909).
- [262] SMAOUI, T.; MÜLLER, S. und MÜLLER-SCHLOER, C.: „CSBIterKmeans: A New Clustering Algorithm Based on Quantitative Assessment of the Clustering Quality“. In: *Machine Learning and Data Mining in Pattern Recognition Proceedings of 6th International Conference (MLDM)*. Hrsg. von PERNER, P. (June 10-13, Leipzig, Germany). Springer, Berlin, Heidelberg, Germany, 2009, S. 337–346. DOI: [10.1007/978-3-642-03070-3_25](https://doi.org/10.1007/978-3-642-03070-3_25).
- [263] WU, W.; XIONG, H. und SHEKHAR, S.: *Clustering and Information Retrieval*. Network Theory and Applications. Springer US, New York, USA, 2013.

- [264] KRUSE, R.; BORGELT, C.; KLAWONN, F.; MOEWES, C.; RUSS, G. und STEINBRECHER, M.: „Fuzzy-Clusteranalyse“. In: *Computational Intelligence*. Computational Intelligence. Vieweg+Teubner, Wiesbaden, Germany, 2011, S. 339–346. DOI: [10.1007/978-3-8348-8299-8_21](https://doi.org/10.1007/978-3-8348-8299-8_21).
- [265] HALKIDI, M.; BATISTAKIS, Y. und VAZIRGIANNIS, M.: „On Clustering Validation Techniques“. In: *Journal of intelligent information systems*, Bd. 17, Nr. 2-3 (2001), S. 107–145. DOI: [10.1023/A:1012801612483](https://doi.org/10.1023/A:1012801612483).
- [266] LIU, Y.; LI, Z.; XIONG, H.; GAO, X. und WU, J.: „Understanding of Internal Clustering Validation Measures“. In: *IEEE International Conference on Data Mining*. (December 13-17, Sydney, Australia). IEEE, Piscataway, NJ, USA, 2010, S. 911–916. DOI: [10.1109/ICDM.2010.35](https://doi.org/10.1109/ICDM.2010.35).
- [267] AMIGÓ, E.; GONZALO, J.; ARTELES, J. und VERDEJO, F.: „A Comparison of Extrinsic Clustering Evaluation Metrics Based on Formal Constraints“. In: *Information Retrieval*, Bd. 12, Nr. 4 (2009), S. 461–486. DOI: [10.1007/s10791-008-9066-8](https://doi.org/10.1007/s10791-008-9066-8).
- [268] MEILĀ, M.: „Comparing Clusterings by the Variation of Information“. In: *Proc., Learning Theory and Kernel Machines: 16th Annual Conference on Learning Theory and 7th Kernel Workshop*. Hrsg. von SCHÖLKOPF, B. und WARMUTH, M. K. (August 24-27, Washington, DC, USA). Springer, Berlin, Heidelberg, Germany, 2003, S. 173–187. DOI: [10.1007/978-3-540-45167-9_14](https://doi.org/10.1007/978-3-540-45167-9_14).
- [269] ROSENBERG, A. und HIRSCHBERG, J.: „V-Measure: A Conditional Entropy-Based External Cluster Evaluation Measure“. In: *Proc., Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. Bd. 7. (June 28-30, Prague, Czech Republic). Association for Computational Linguistics, Stroudsburg, PA, USA, 2007, S. 410–420.
- [270] KREMER, H.; KRANEN, P.; JANSEN, T.; SEIDL, T.; BIFET, A.; HOLMES, G. und PFAHRINGER, B.: „An Effective Evaluation Measure for Clustering on Evolving Data Streams“. In: *Proc., 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD 2011. (August 21-24, San Diego, USA). ACM, New York, NY, USA, 2011, S. 868–876. DOI: [10.1145/2020408.2020555](https://doi.org/10.1145/2020408.2020555).
- [271] SCHNEIDER-OBERMANN, H.: *Basiswissen der Elektro-, Digital- und Informationstechnik: Für Informatiker, Elektrotechniker und Maschinenbauer*. Hrsg. von MILDENBERGER, O. Vieweg+Teubner, Wiesbaden, Germany, 2006. DOI: [10.1007/978-3-8348-9122-8](https://doi.org/10.1007/978-3-8348-9122-8).
- [272] WERNER, M.: *Information und Codierung - Grundlagen und Anwendungen*. Hrsg. von WERNER, M. 2., vollständig überarbeitete und erweiterte Auflage. Vieweg+Teubner, Wiesbaden, Germany, 2009.
- [273] SHANNON, C. E.: „A Mathematical Theory of Communication“. In: *The Bell System Technical Journal*, Bd. 27 (1948), S. 379–423.
- [274] BRELL, C.: *Statistik von Null auf Hundert: Mit Kochrezepten schnell zum Statistik-Grundwissen*. Hrsg. von BRELL, J. und KIRSCH, S. 2. Aufl. Springer, Berlin, Heidelberg, Germany, 2017.

-
- [275] CRAMER, E.: *Grundlagen der Wahrscheinlichkeitsrechnung und Statistik: Eine Einführung für Studierende der Informatik, der Ingenieur- und Wirtschaftswissenschaften*. Hrsg. von KAMPS, U. 4. Aufl. Springer, Berlin, Heidelberg, Germany, 2017.
- [276] LANGE, T.: *Statistik kompakt: Basiswissen für Ökonomen und Ingenieure*. Hrsg. von MOSLER, K. Springer, Berlin, Heidelberg, Germany: Berlin, Heidelberg, 2017.
- [277] HARTUNG, J.; ELPELT, B. und KLÖSENER, K.-H.: *Statistik: Lehr- und Handbuch der angewandten Statistik*. 15. Aufl. Oldenbourg, München, Germany, 2009. DOI: [10.1524/9783486710540](https://doi.org/10.1524/9783486710540).
- [278] ERNST & YOUNG: *Kosten-Nutzen-Analyse für einen flächendeckenden Einsatz intelligenter Zähler*. Technischer Bericht. Ernst & Young GmbH, 2013.
- [279] WACZOWICZ, S.; REISCHL, M.; KLAIBER, S.; BRETSCHNEIDER, P.; KONOTOP, I.; WESTERMANN, D.; HAGENMEYER, V. und MIKUT, R.: „Virtual Storages as Theoretically Motivated Demand Response Models for Enhanced Smart Grid Operations“. In: *Energy Technology*, Bd. 4 (2016), S. 163–176.
- [280] NYENG, P. und OSTERGAARD, J.: „Information and Communications Systems for Control-By-Price of Distributed Energy Resources and Flexible Demand“. In: *IEEE Transactions on Smart Grid*, Bd. 2, Nr. 2 (2011), S. 334–341. DOI: [10.1109/TSG.2011.2116811](https://doi.org/10.1109/TSG.2011.2116811).
- [281] HUNTER, J. S.: „The Exponentially Weighted Moving Average“. In: *Journal of Quality Technology*, Bd. 18, Nr. 4 (1986), S. 203–210.
- [282] BYRD, R. H.; SCHNABEL, R. B. und SHULTZ, G. A.: „Approximate Solution of the Trust Region Problem by Minimization over Two-Dimensional Subspaces“. In: *Mathematical Programming*, Bd. 40, Nr. 1-3 (1988), S. 247–263. DOI: [10.1007/BF01580735](https://doi.org/10.1007/BF01580735).
- [283] STEihaug, T.: „The Conjugate Gradient Method and Trust Regions in Large Scale Optimization“. In: *SIAM Journal on Numerical Analysis*, Bd. 20, Nr. 3 (1983), S. 626–637. DOI: [10.1137/0720042](https://doi.org/10.1137/0720042).
- [284] LJUNG, L.: *System Identification: Theory for the User*. Prentice Hall, Englewood Cliffs, NJ, USA, 1987.
- [285] TOTH, R.; SANANAJI, B.; POOLLA, K. und VINCENT, T.: „Compressive System Identification in the Linear Time-Invariant framework“. In: *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*. (December 12-15, Orlando, USA). IEEE, Piscataway, NJ, USA, 2011, S. 783–790. DOI: [10.1109/CDC.2011.6160383](https://doi.org/10.1109/CDC.2011.6160383).
- [286] KÜSTERS, U.: „Evaluation, Kombination und Auswahl betriebswirtschaftlicher Prognoseverfahren“. In: *Prognoserechnung*. Hrsg. von MERTENS, P. und RÄSSLER, S. Physica, Heidelberg, Germany, 2012, S. 423–467. DOI: [10.1007/978-3-7908-2797-2_18](https://doi.org/10.1007/978-3-7908-2797-2_18).
- [287] MIKUT, R. und REISCHL, M.: „Data Mining Tools“. In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Bd. 1, Nr. 5 (2011), S. 431–443.
- [288] MIKUT, R.; BARTSCHAT, A.; DONEIT, W.; GONZÁLEZ ORDIANO, J. Á.; SCHOTT, B.; STEGMAIER, J.; WACZOWICZ, S. und REISCHL, M.: *The MATLAB Toolbox SciXMiner: User’s Manual and Programmer’s Guide*. Technischer Bericht. arXiv:1704.03298, 2017.

- [289] KIRSCHEN, D. S.; STRBAC, G.; CUMPERAYOT, P. und PAIVA MENDES, D. DE: „Factoring the elasticity of demand in electricity prices“. In: *IEEE Transactions on Power Systems*, Bd. 15, Nr. 2 (2000), S. 612–617. DOI: [10.1109/59.867149](https://doi.org/10.1109/59.867149).
- [290] STERNER, M. und STADLER, I., Hrsg.: *Energiespeicher – Bedarf, Technologien, Integration*. 2. Aufl. Springer, Berlin, Heidelberg, Germany, 2017. DOI: [10.1007/978-3-662-48893-5_2](https://doi.org/10.1007/978-3-662-48893-5_2).
- [291] WALDSCHMIDT, W.: *ABC der Wärme in Wohngebäuden*. VWEW Energieverlag, Frankfurt, Berlin, Heidelberg, Germany, 2006.
- [292] FREY, G.; SCHULZ, W.; HORST, J. und LEPRICH, U.: *Studie zu den Energieeffizienzpotenzialen durch Einsatz von elektrischem Strom im Raumwärmebereich*. Technischer Bericht. Institut für ZukunftsEnergieSysteme gGmbH, bremer energie institut, 2007. URL: http://www.bmub.bund.de/fileadmin/bmu-import/files/pdfs/allgemein/application/pdf/studie_stromheizungen.pdf (besucht am 10.03.2015).
- [293] QUASCHNING, V.: *Regenerative Energiesysteme: Technologie-Berechnung-Simulation*. 8. Aufl. Hanser, München, Germany, 2013. URL: <http://www.hanser-elibrary.com/doi/book/10.3139/9783446443334> (besucht am 22.10.2015).
- [294] PLATT, M.; EXNER, S. und BRACKE, R.: *Analyse des deutschen Wärmepumpenmarktes, Bestandsaufnahme und Trends*. Technischer Bericht. GeothermieZentrum Bochum, Hochschule Bochum, 2010. URL: http://www.geothermiezentrum.de/fileadmin/media/geothermiezentrum/Projekte/WP-Studie/Abschlussbericht%5C_WP-Marktstudie%5C_Mar2010.pdf (besucht am 10.03.2015).
- [295] BWP: *BWP-Branchenstudie 2013 - Szenarien und politische Handlungsempfehlungen*. Technischer Bericht. Bundesverband Wärmepumpe (BWP) e.V., 2013.
- [296] STADLER, I.: „Ein gigantisches Speicherpotenzial“. In: *Solar zeitalter*, Bd. 1 (2008), S. 60–64.
- [297] SCHLOMANN, B.; GRUBER, E.; EICHHAMMER, W.; KLING, N.; DIEKMANN, J.; ZIESING, H.-J.; RIEKE, H.; WITTKE, F.; HERZOG, T.; BARBOSA, M. u. a.: *Energieverbrauch der privaten Haushalte und des Sektors Gewerbe, Handel, Dienstleistungen (GHD)*. Technischer Bericht. 2004.
- [298] STATISTISCHES BUNDESAMT: *Einkommens- und Verbrauchsstichprobe - Ausstattung privater Haushalte mit ausgewählten Gebrauchsgütern*. Technischer Bericht. Statistisches Bundesamt, Wiesbaden, Fachserie 15, Heft 1, 2013.
- [299] DENA: *Der dena-Gebäudereport 2012 – Statistiken und Analysen zur Energieeffizienz im Gebäudebestand*. Technischer Bericht. dena - Deutsche Energie-Agentur, 2012.

-
- [300] STUIBLE, A.; ZECH, D.; WÜLBECK, H.-F.; SPERBER, E.; NAST, M.; HARTMANN, H.; REISINGER, K.; BUDIG, C.; OROZALIEV, J.; PAG, F.; VAJEN, K.; ERLER, R.; JANCZIK, S.; KALTSCHMITT, M. und NIEDERBERGER, M.: *Evaluierung von Einzelmaßnahmen zur Nutzung erneuerbarer Energien im Wärmemarkt (Marktanreizprogramm) für den Zeitraum 2012 bis 2014*. Technischer Bericht. Ausarbeitung im Auftrag des Bundesministerium für Wirtschaft und Energie. Fichtner, DLR, TZF, IdE, DBI-Gas und Umwelttechnik, TUHH-IUE, COMPARE Consulting, 2016. URL: http://elib.dlr.de/107618/1/FICHT-%5C%2317477735-v1-Revidierter%5C_finaler%5C_Bericht%5C_mit%5C_Appendices.pdf (besucht am 16. 08. 2017).
- [301] ZECH, D.; JENSEN, T. und ELTROP, L.: *Technologien, Emissionen, Kosten – ein Überblick über Möglichkeiten der Wärmeversorgung von Wohngebäuden mit Erneuerbaren Energien*. Technischer Bericht. Diskussionspapier. Institut für Energiewirtschaft und Rationelle Energieanwendung, Abteilung Systemanalyse und Erneuerbare Energien, Universität Stuttgart, 2010.
- [302] BDEW: *BDEW-Steckbriefe zu Heiztechniken und additiven Erdgasanwendungen mit fachlichem Hintergrundwissen für eine gezielte Kommunikation und Kundenberatung*. Technischer Bericht. Bundesverband der Energie- und Wasserwirtschaft e.V., 2007.
- [303] HILLEMACHER, L.; HUFENDIEK, K.; BERTSCH, V.; WIECHMANN, H.; GRATENAU, J.; JOCHEM, P. und FICHTNER, W.: „Ein Rollenmodell zur Einbindung der Endkunden in eine smarte Energiewelt“. In: *Zeitschrift für Energiewirtschaft*, Bd. 37, Nr. 3 (2013), S. 195–210. DOI: [10.1007/s12398-013-0110-z](https://doi.org/10.1007/s12398-013-0110-z).
- [304] BANKHOFER, U. und VOGEL, J.: *Datenanalyse und Statistik - Eine Einführung für Ökonomen im Bachelor*. Gabler, Wiesbaden, Germany, 2008.

G Stichwortverzeichnis

- A**
Ancillary Service Markets 7
 Ausreißerbehandlung 13
 Automatisches *Demand Response* 12
- B**
 Beeinflusster Verbrauchsprozess 29
 Benchmarkdatensatz 44
 Benchmarkgenerator 37
Bottom-Up-Modell 20
- C**
Capacity Market 7
 Clusterverfahren 21, 62
 Fuzzy-C-Means-Clustering 63
Critical Peak Pricing 11, 117, 152
Customer Baseline Load 16
- D**
Demand Bidding 7
Demand Response 8
Demand Side Management 7
Demand-Response-Modelle
 Clustermodell 21
 dynamisches Regressionsmodell . . 19, 133
 ökonomisches DR-Modell 18
 Optimierungsmodell 20
 physikalisch motiviertes Modell 20
 stochastisches Modell 20
 Virtuelle-Speicher-Modell 93, 133
Demand-Response-Prozess 30
Demand-Response-Simulator 40
Direct load control 7
- E**
 Eigenpreiselastizität 18
Emergency Demand Response 7
 Entropie 70
- Äquivokation *siehe auch* Entropie,
 Verlustentropie
 Ausgangsentropie 70
 Eingangsentropie 70
 Fehlinformation 71
 Irrelevanz *siehe auch* Entropie,
 Fehlinformation
 Verlustentropie 71
 Erneuerbare Energien 2
- H**
 Hampel-Filter 50
- I**
 Imputationsverfahren 13, 48
Interruptible load 7
- K**
 Komponentenmodell 16
 Kreuzpreiselastizität 18
- L**
 Lastprognose 15
 Lastsignalgenerator 38
- M**
 Markov-Kette *siehe auch Demand-Response*-
 Modelle, stochastisches Modell
 Messfehlergenerator 41
 Modellierung
 Modellanwendung 101
 Modellentwurf 101
 Modellselektion 103
- N**
 NIALM 33
- P**
Peak-Time Rebate 12
 Preiselastizität 18
 Preissignal 28, 41, 118, 138

R

Real-Time Pricing 11, 118, 154

S

SAIDI 5

Smart Grid 6

Standardlastprofil 15

Startcluster 66, 84

T

Time-of-Use 10, 117, 138, 154

Transinformation 70

relative Transinformation 70, 72, 123, 140

U

Unbeeinflusster Verbrauchsprozess 29

V

Verbundverteilung 69, 73, 123, 125, 140

Versorgungssicherheit 5

Z

Zeitreihenaggregation 57

Zeitreihennormierung 54

Zeitreihenzerlegung 53