



Thomas Nigitz, BSc

Modellierung und Prognose des Wärmebedarfs von Gebäuden und Wärmenetzen

MASTERARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

Masterstudium Elektrotechnik

eingereicht an der

Technischen Universität Graz

Betreuer

Em.Univ.-Prof. Dipl.-Ing. Dr.techn. Nicolaos Dourdoumas

Institut für Regelungs- und Automatisierungstechnik

Dipl.-Ing. Dr.techn. Markus Gölles
BIOENERGY 2020+ GmbH

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Masterarbeit identisch.

Datum

Unterschrift

Vorwort

Diese Arbeit wurde von August 2014 bis Mai 2015 am COMET-Kompetenzzentrum BIOENERGY 2020+ in Kooperation mit dem Institut für Regelungs- und Automatisierungstechnik (IRT) der Technischen Universität Graz verfasst. BIOENERGY 2020+ forscht anwendungsorientiert auf dem Gebiet der energetischen Nutzung von Biomasse. Die Regelungsgruppe unter der Leitung von Dr. Markus Göllles steht in langjährigem, engem Kontakt zum IRT. Die Lastprognose, das Thema dieser Arbeit, ist ein neues Forschungsgebiet beider Partner, weshalb das Institut für Statistik unter der Leitung von Prof. Ernst Stadlober im Zuge der Methodenauswahl unterstützend zu Rate gezogen wurde.

An dieser Stelle möchte ich mich bei meinem Betreuer in der Regelungsgruppe, Dr. Markus Göllles, bedanken, der durch wöchentliche Besprechungen die Arbeit auf Kurs gehalten und mit vielen nützlichen Hinweisen und Ratschlägen weitergeholfen hat. Zusätzlich möchte ich der gesamten Regelungsgruppe für das gute Arbeitsklima und die anregenden Diskussionen danken. Außerdem bedanke ich mich bei meinem Betreuer am IRT, Prof. Nicolaos Dourdoumas, und bei Prof. Ernst Stadlober für seine hilfreichen fachlichen Hinweise. Abschließend möchte ich mich bei meinen Eltern Manfred und Veronika bedanken, die mir ein fokussiertes und erfolgreiches Studium ermöglicht haben; sowie bei meiner Freundin Julia die für den nötigen Ausgleich gesorgt hat.

Die entwickelte Prognosemethode ist mit Hinsicht auf die Implementierung in der Praxis entwickelt worden, basiert auf anerkannten wissenschaftlichen Methoden und liefert, trotz ihrer Einfachheit, erstaunlich gute Ergebnisse. In diesem Sinne wünsche ich dem Leser / der Leserin viel Freude beim Lesen der Arbeit, wie auch mir das Verfassen bereitet hat.

Ort, Datum

Unterschrift

Kurzfassung

Die vorteilhafte, erneuerbare Wärmeerzeugung durch Biomasse-Feuerungsanlagen bzw. durch wetterabhängige Solarthermie kann Wärme nicht punktuell zur Verfügung stellen. Daher werden in der Praxis Wärmereserven in Pufferspeichern angelegt bzw. wird mit schnelleren, fossilen oder elektrischen Wärmeerzeugern auf Lastwechsel bzw. Lastspitzen reagiert. Die verschiedenen Einzelkomponenten sind bereits gut entwickelt, deren Zusammenspiel bietet aber noch deutliches Verbesserungspotential. Dieses Potential kann mit einer übergeordneten Regelung ausgeschöpft werden, welche aufgrund der nicht beliebig modulierbaren bzw. wetterabhängigen Wärmeerzeuger vorausplanend arbeiten muss. Damit diese Regelung den zukünftigen Wärmebedarf berücksichtigen kann, muss dieser prognostiziert werden.

Ziel dieser Arbeit war eine Methode für die Prognose des Wärmebedarfs von Verbrauchern zu entwickeln, wobei sieben unterschiedliche Verbraucher – vom Wohnhaus bis zum Fernwärmenetz – betrachtet werden sollten. Die entwickelte Prognosemethode entspricht einer Prädiktion, welche durch eine Korrektur und durch zusätzliches Expertenwissen verbessert wird. Der Prädiktor basiert auf einer ausführlichen Messdatenauswertung aller sieben Verbraucher, anhand derer aus den üblichen Prognosemethoden der Literatur eine leicht implementierbare Methode ausgewählt wurde. Der Korrektor reduziert systematische Fehler in der Prädiktion und verbessert somit die Prognose für die nahe Zukunft. Hierbei wurde die spätere praktische Anwendung auf Recheneinheiten mit begrenzten Ressourcen berücksichtigt und somit immer von einem einfachen Ansatz aus erweitert, bis eine Erweiterung keine weitere Verbesserung mehr brachte.

Die entwickelte Prädiktor-Korrektor-Prognosemethode ist keine Speziallösung, sondern für ein breites Verbraucherspektrum entwickelt und getestet. Außerdem enthält sie keine Parameter die von Hand gewählt werden müssen, und erfüllt einen guten Kompromiss aus Flexibilität und Robustheit.

Abstract

The advantageous, renewable heat production via biomass furnaces or weather-dependent solar collectors cannot track an occurring heat demand in real-time. Therefore they are typically combined with buffer storages and quick, fossil or electric heat producers to cover load peaks as well as quick load increases. The single components are already highly developed, but there is still large potential for improvement of the interaction of these components. To utilize this potential a superordinate controller is needed, which has to pre-plan the operation of the slow and weather-dependent renewable heat producers. For this pre-plan the future heat load needs to be forecasted.

The aim of this thesis was to develop a method for the forecast of the heat load required by different consumers. Seven different consumers—from a residential building to a district heating grid—have been investigated. The developed forecast method consists of a prediction that is improved by a correction based on the current prediction error and by additional expert knowledge. The prediction is based on an extensive investigation of measurement data from the different consumers and can predict far ahead. The correction reduces systematic prediction errors and improves the forecast for the near future. Both parts are chosen to be easy implementable on state-of-the-art control units with limited capacities.

This prediction-correction-forecast-method is no specific solution, but rather developed and tested for a wide range of consumers. Furthermore, it contains no parameters that need to be adjusted and is a good compromise between adaptivity and robustness.

Inhaltsverzeichnis

1. Einleitung	1
2. Statistische Begriffe und Grundlagen	5
2.1. Notation und Begriffe	5
2.1.1. Notation	5
2.1.2. Begriffe	6
2.2. Regressionsanalyse	7
2.3. Zeitreihenanalyse	11
2.4. Qualitätsmaße	12
3. Prognosemethoden in der Literatur	15
3.1. Prädiktor	15
3.1.1. Klassische Methoden	16
3.1.2. Künstliche-Intelligenz-Methoden	17
3.1.3. Diskussion	18
3.2. Korrektor	19
4. Messdaten	21
4.1. Wärmestrom als Messgröße	21
4.2. Klassifizierung der Verbraucher	22
4.3. Messdatenaufbereitung	22
4.4. Messdatenauswertung	23
4.4.1. Einfluss des Wetters	23
4.4.2. Einfluss zeitlicher Faktoren	28
4.4.3. Einfluss des vergangenen Lastverlaufs	35
4.4.4. Zusammenfassung	38
5. Prognosemethode	41
5.1. Grundlegende Überlegungen	41
5.2. Erster Prädiktor	43
5.2.1. Ansatz	43
5.2.2. Parameterbestimmung	44
5.2.3. Ergebnis	51
5.3. Erweiterung des Prädiktors	57
5.3.1. Ansatz	57
5.3.2. Parameterbestimmung	57
5.3.3. Ergebnis	60

Inhaltsverzeichnis

5.4.	Erster Korrektor	64
5.4.1.	Ansatz	64
5.4.2.	Parameterbestimmung	65
5.4.3.	Ergebnis	66
5.5.	Erweiterung des Korrektors	78
5.5.1.	Ansatz	79
5.5.2.	Parameterbestimmung	79
5.5.3.	Ergebnis	84
6.	Finale Methode und deren Implementierung	87
6.1.	Finale Methode	87
6.2.	Implementierung	89
6.2.1.	Wetterprognosedaten	90
6.2.2.	Ressourcenschonende Kostenfunktion	91
6.2.3.	Ergebnis	92
7.	Zusammenfassung und Ausblick	99
7.1.	Zusammenfassung	99
7.2.	Ausblick	100
A.	Zusatzmaterial	105
A.1.	Verbraucherdaten und Prognosebeispiele	105
A.2.	Lineare Programme in Matlab	116

Abkürzungsverzeichnis

AKF	Autokorrelationsfunktion
ARMA	Autoregressive-Moving-Average
DTFT	Discrete-Time-Fourier-Transform
KW	Kalenderwoche
LAD	Least-Absolute-Deviation
LAV	Least-Absolute-Value
LP	Lineares Programm
MAE	Mean-Absolute-Error
MAPE	Mean-Absolute-Percentage-Error
MPE	Mean-Percentage-Error
NN	Neuronales-Netz
OLS	Ordinary-Least-Squares
SARIMA	Seasonal-Autoregressive-Integrated-Moving-Average
VB	Verbraucher
WLS	Weighted-Least-Squares

1. Einleitung

Die erneuerbaren Wärmeerzeugungstechnologien Biomasse-Feuerung bzw. Solarthermie können, trotz ihres hohen Entwicklungsstandes, einen auftretenden Wärmebedarf häufig nicht ohne Unterstützung und zeitnah decken. Dies ist durch die systembedingte Trägheit der Biomasse-Feuerungsanlagen und insbesondere auch die Wetterabhängigkeit der Solaranlagen begründet. Daher werden in der Praxis Wärmereserven in Pufferspeichern angelegt bzw. wird mit schnelleren Technologien unterstützt. Als schnellere Technologien kommen hier die Verbrennung nicht-erneuerbarer fossiler Brennstoffe, wie Erdgas und Erdöl, oder die Umwandlung der ansonsten vielseitig einsetzbaren elektrischen Energie in Wärme zum Einsatz.

Alle diese Einzelkomponenten sind bereits gut entwickelt; das Zusammenspiel dieser Komponenten kann aber noch deutlich verbessert werden. Hierfür ist eine übergeordnete Regelung erforderlich, welche den Einsatz der erneuerbaren Wärmeerzeugung maximiert und den Einsatz von fossiler und elektrischer Wärmeerzeugung minimiert. Aufgrund der Trägheit bzw. Wetterabhängigkeit der erneuerbaren Wärmeerzeuger muss diese Regelung vorausplanend arbeiten. Dafür ist unter anderem eine Prognose des Wärmebedarfs des Verbrauchers notwendig. Schlussendlich kann dadurch das Zusammenspiel der Einzelkomponenten für die Wärmeerzeugung stark verbessert werden, wodurch der Einsatz elektrischer Energie für die Wärmeerzeugung, der Brennstoffverbrauch und auch die Schadstoffemissionen reduziert werden können. Beim alleinigen Einsatz von Biomasse-Feuerungsanlagen könnte dadurch auch die notwendige Größe des Pufferspeichers reduziert werden, da die Trägheit der Anlage durch eine vorausplanende Regelung zum Teil kompensiert werden könnte.

Zur besseren Veranschaulichung der Problemstellung soll folgendes einfache Beispiel betrachtet werden. Ein Wohnhaus wird im Sommer mit einer Kombination aus Solarthermie und elektrischen Heizstäben (Begleitheizung) mit Wärme versorgt. Natürlich muss für diese Kombination ein Pufferspeicher eingesetzt werden, um die Sonnenenergie des Tages für die restliche Zeit zu speichern. Dabei kommen häufig sehr banale Steuerstrategien für die Begleitheizung zum Einsatz, die aktiv wird wenn ein bestimmter Ladezustand des Pufferspeichers unterschritten wird und wieder deaktiviert wird wenn ein bestimmter Ladezustand überschritten wird. Nun kann es auftreten, dass am Abend (durch Duschen o. Ä.) die Temperatur im Pufferspeicher einen Schwellwert unterschreitet, wodurch die Begleitheizung einschaltet. Diese Begleitheizung heizt nun den Pufferspeicher bis zu einer bestimmten Temperatur auf. Über Nacht wird keine Wärme benötigt und am Morgen wird nur ein geringer Teil der gespeicherten Wärme entnommen. Somit ist im Pufferspeicher mehr Wärme als erforderlich gespeichert und die Sonnenenergie des Tages kann womöglich nur teilweise gespeichert werden. Wäre hier eine übergeordnete Regelung inklusive Wärmebedarfsprognose (und in weiterer Folge auch eine Prognose

1. Einleitung

des zu erwartenden Solareintrages) im Einsatz, so könnte mit der Begleitheizung nur die am Morgen benötigte Wärme erzeugt werden und die Sonnenenergie könnte voll genutzt werden.

An diesem Punkt setzt diese Arbeit an. Ziel ist die Entwicklung einer Methode, mit welcher der Wärmebedarf von Gebäuden und Wärmenetzen laufend prognostiziert werden kann.

Solche Prognosemethoden sind seit Jahrzehnten bei elektrischen Energieversorgern im Einsatz, um ein stabiles Netz sicherzustellen und die Stromerzeugungskosten unter Beachtung vieler Randbedingungen zu minimieren (siehe z.B. [1, S. 1559]).

In Abbildung 1.1 ist das Ziel dieser Arbeit graphisch veranschaulicht. Wärmeerzeuger speisen Wärmeenergie in einen Pufferspeicher ein, von dem aus die Wärmeverbraucher ihren Wärmebedarf decken. Dieser Wärmebedarf wird mit dem Wärmestrom \dot{Q} beschrieben, der in der Literatur auch als Last L bezeichnet wird. Die Aufgabe der Prognosemethode besteht nun darin, den Wärmestrom, zum Beispiel für die nächsten 24 Stunden, zu prognostizieren.

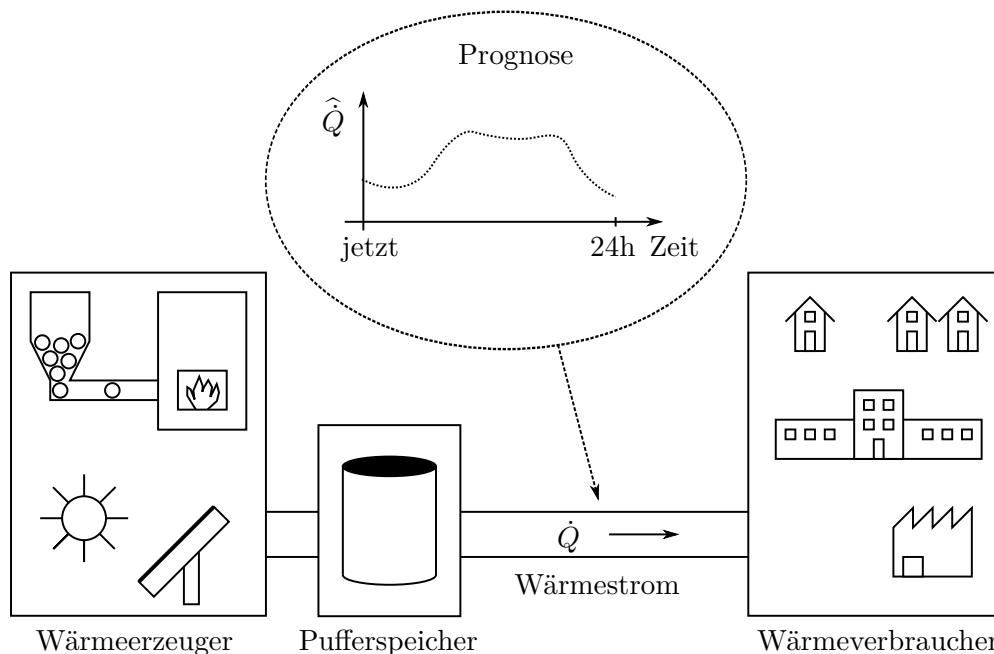


Abbildung 1.1.: Ziel der Arbeit ist die Prognose des Wärmestroms \dot{Q} der in den nächsten Stunden/Tagen von den Wärmeverbrauchern aus dem Pufferspeicher entnommen wird.

Diese Prognosemethode wird in eine Prädiktion (Prädiktor) und eine Korrektur dieser Prädiktion (Korrektor) aufgeteilt, wodurch diese modularen Komponenten getrennt entwickelt und eingesetzt werden können. Zusammen liefern diese eine Prognose für den Wärmestrom \hat{Q} , welche in jedem Zeitschritt (z.B. stündlich) aktualisiert wird.

Aufbau und Methodik

Zu Beginn wird in Kapitel 2 (Statistische Begriffe und Grundlagen) die Notation festgelegt und es werden grundlegende Begriffe zusammengefasst. Danach werden zwei Methoden der Statistik beschrieben, welche häufig für die Lastprognose eingesetzt werden. Im letzten Teil dieses Kapitels werden noch Qualitätsmaße behandelt, anhand derer die Güte der Prognose bestimmt werden kann.

In Kapitel 3 (Prognosemethoden in der Literatur) wird ein umfassender Überblick über übliche Methoden zur Lastprognose in der Literatur gegeben. Hierbei wird grob in Prädiktor und Korrektor unterteilt, da dies dem Aufbau dieser Arbeit entspricht. In der Literatur wird in der Regel jedoch nicht explizit zwischen Prädiktor und Korrektor unterschieden, sondern eine geschlossene Prognosemethode beschrieben. Daher ist der überwiegende Teil der Prognosemethoden im Prädiktor-Teil des Kapitels zu finden, wo auch versucht wird einen Trend in den wissenschaftlichen Artikeln zu identifizieren. Im Korrektor-Teil werden nur einzelne Ansätze zur Reduktion systematischer Prognosefehler beschrieben, die jedoch aus anderen Disziplinen als der Lastprognose kommen.

Darauf folgt in Kapitel 4 (Messdaten) eine ausführliche Messdatenauswertung sieben stark unterschiedlicher Verbraucher, womit zuverlässige Aussagen über das Verhalten des zu prognostizierenden Wärmestroms gemacht werden können. Zusätzlich wird auf die Messung des Wärmestroms, dessen Aufbereitung und auf die betrachteten Verbraucher eingegangen.

Die Ergebnisse der Messdatenauswertung und der Literaturrecherche sind die Basis für die in Kapitel 5 (Prognosemethode) entwickelte Prognosemethode, wobei diese klar in Prädiktor und Korrektor aufgespalten ist. Bei der Entwicklung der Prognosemethode wird immer von einem sehr einfachen Ansatz ausgegangen und dieser schrittweise erweitert, bis eine Erweiterung keine Verbesserung mehr bringt. Somit wird zunächst der Prädiktor einfach angesetzt, erweitert und vollständig evaluiert. Danach wird der Korrektor entwickelt und ebenfalls für alle Verbraucher evaluiert.

Da diese Prognosemethode schlussendlich auf einer Recheneinheit mit begrenzten Ressourcen eingesetzt werden soll, wird in Kapitel 6 (Finale Methode und deren Implementierung) die finale Prognosemethode beschrieben und auf Auswirkungen möglicher Implementierungsvarianten eingegangen. Für eine reine Implementierung, ohne Kenntnis der zugrundeliegenden Überlegungen, ist es ausreichend nur dieses Kapitel zu lesen.

Den Abschluss bildet Kapitel 7 (Zusammenfassung und Ausblick), wo die Kernpunkte der Arbeit zusammengefasst werden und auf mögliche zukünftige Erweiterungen eingegangen wird. Hierbei wird auch auf mögliche Speziallösungen eingegangen, die in dieser Arbeit nicht betrachtet werden, da eine allgemeine Prognosemethode gesucht wurde.

Im Anhang A (Zusatzmaterial) sind die Zeiträume der Verbraucherdaten tabellarisch aufbereitet und Prognosebeispiele zu allen Verbrauchern zu finden. Zusätzlich wird ein Spezialfall Linearer Programme in Matlab¹ behandelt, da dieser bei der Parameterbestimmung des Prädiktors zur Anwendung kommt.

¹Matlab ist eine kommerzielle Software zur numerischen Lösung mathematischer Probleme und zur grafischen Darstellung der Ergebnisse. <http://www.mathworks.com/products/matlab>

2. Statistische Begriffe und Grundlagen

Zu Beginn dieses Kapitels wird die Notation festgelegt und es werden einige Begriffe aus dem Gebiet der Statistik und Signalverarbeitung zusammengefasst. Dann werden die wichtigen Methoden der Regressions- und der Zeitreihenanalyse vorgestellt. Die Regressionsanalyse wird in der Messdatenauswertung verwendet und dient als Grundlage für den Prädiktor. Die Zeitreihenanalyse könnte gemäß Messdatenauswertung und Literaturrecherche ebenfalls für den Prädiktor eingesetzt werden und wird daher kurz beschrieben. Zum Schluss werden die in dieser Arbeit verwendeten Qualitätsmaße definiert und diskutiert. Somit kann ein Begriff bzw. eine Definition schnell nachgeschlagen werden und für den interessierten Leser / die interessierte Leserin werden entsprechende Literaturhinweise gegeben.

2.1. Notation und Begriffe

In dieser Arbeit kommen Methoden der Statistik, der Signalverarbeitung und der Regelungstechnik zum Einsatz. Alle diese Disziplinen haben ihre eigene Notation und zum Teil auch verschiedene Definitionen für den gleichen Begriff. Damit keine Unklarheiten bei der später verwendeten Notation auftreten, wird diese hier festgelegt.

2.1.1. Notation

Die Notation wird anhand einer allgemeinen Variable x beschrieben, wobei diese hier noch nicht der späteren Bezeichnung des Parametervektors entspricht. Wird ein Laufindex, wie bei x_i , angegeben, so handelt es sich um Elemente eines Spaltenvektors,

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_N \end{bmatrix}^T. \quad (2.1)$$

Wird hingegen eine Laufvariable in eckigen Klammern, wie bei $x[n]$, angegeben so handelt es sich um eine diskrete Abhängigkeit der Zeit. Diese Schreibweise der Signalverarbeitung entspricht einer Erzeugung von $x[n]$ durch Abtastung eines analogen Signals $x_a(t)$, mit der Abtastzeit (Diskretisierungszeit) T_d

$$x[n] = x_a(nT_d). \quad (2.2)$$

2. Statistische Begriffe und Grundlagen

So könnte zum Beispiel ein Vektor \mathbf{x}_{BSP} die letzten N Werte in einem Abstand von 24 Zeitschritten enthalten:

$$\mathbf{x}_{\text{BSP}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N-1} \\ x_N \end{bmatrix} = \begin{bmatrix} x[n - N \cdot 24] \\ x[n - (N - 1) \cdot 24] \\ \vdots \\ x[n - 48] \\ x[n - 24] \end{bmatrix}. \quad (2.3)$$

Eine weitere wichtige Unterscheidung betrifft die Ableitung der Zeit. Wird die Ableitung der Zeit für eine Differentialgleichung, wie bei einem dynamischen System, verwendet, so wird dies mit $\frac{dx}{dt}$ gekennzeichnet. Eine Bezeichnung mit einem Punkt, wie \dot{x} , entspricht nur dem Namen einer Flussgröße. So bezeichnet zum Beispiel $\dot{x} = f(u)$ keine Differentialgleichung sondern eine allgemeine Funktion $f : u \mapsto \dot{x}$.

2.1.2. Begriffe

Die folgenden Definitionen sind sehr grundlegend, und werden nur zur Wiederholung und Klarstellung der Notation angeführt. Sofern keine andere Quelle angegeben ist, sind sowohl Definitionen als auch Notation [2] entnommen.

Die in der Praxis sicher am häufigsten verwendete statistische Maßzahl ist der **arithmetische Mittelwert** (kurz: Mittelwert)

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i. \quad (2.4)$$

Wie in [2, Kap. 3.3.1] beschrieben minimiert der Mittelwert die Summe der Abweichungsquadrate $\sum_i (x_i - \bar{x})^2$ und ist somit empfindlich gegenüber Ausreißern. Daher wird der Mittelwert auch als ein empfindliches Lagemaß bezeichnet.

Ein robusteres Lagemaß ist der **Medianwert** (kurz: Median) \tilde{x} , da er für die Summe der absoluten Abweichungen $\sum_i |x_i - \tilde{x}|$ ein Minimum liefert [2, Kap. 3.3.1]. Zu Bestimmung des Medians werden zuerst die Werte aufsteigend geordnet und mit Indizes in Klammern versehen, um sie von der ursprünglichen Reihenfolge zu unterscheiden:

$$x_{(1)} \leq x_{(2)} \leq x_{(3)} \cdots \leq x_{(N-1)} \leq x_{(N)}. \quad (2.5)$$

Der Medianwert entspricht nun der Zahl, welche diese geordnete Liste in zwei gleich große Hälften teilt. Ist N gerade, so gibt es zwei Pseudomedianwerte \tilde{x}_1 und \tilde{x}_2 und der Medianwert wird als Mittelwert $\tilde{x} = \frac{1}{2}(\tilde{x}_1 + \tilde{x}_2)$ definiert. Wird die geordnete Liste (2.5) in vier gleiche Teile zerlegt, so spricht man von **Quartilen**, wobei das zweite Quartil Q_2 dem Median entspricht.

Eine häufig verwendete Funktion für die Datenanalyse ist die **Autokorrelationsfunktion** (AKF) $r_{xx}[m]$, da sie eine Auskunft über statistische Abhängigkeiten bzw.

Regelmäßigkeiten in den Daten gibt. Diese Funktion ist in der mathematischen Statistik anders als in der ingenieurwissenschaftlichen Signalverarbeitung definiert. In dieser Arbeit wird die Definition der Signalverarbeitung entsprechend [3, S. 67] gewählt

$$r_{xx}[m] = \sum_{n=-\infty}^{\infty} x[n]x[n+m]. \quad (2.6)$$

Man beachte den Unterschied zu Faltung, welche über $x[n]x[n-m]$ definiert ist.

Eine verwandte Größe ist das **Leistungsdichtespektrum** $S_{xx}(e^{j\omega})$, welches der Discrete-Time-Fourier-Transform (DTFT) der Autokorrelationsfunktion entspricht

$$S_{xx}(e^{j\omega}) = \text{DTFT}(r_{xx}[m]) = \sum_{m=-\infty}^{\infty} r_{xx}[m] e^{-j\omega m}. \quad (2.7)$$

Mit dem Leistungsdichtespektrum kann die AKF im Frequenzbereich analysiert werden. Hierfür wird in der Regel nur eine Schätzung des normierten Betrags von $S_{xx}(e^{j\omega})$ betrachtet, was als **Periodogramm** $I(\omega)$ bezeichnet wird. Das Argument von $I(\omega)$ ist die Kreisfrequenz, die gemäß $\omega = 2\pi f$ mit der Frequenz f zusammenhängt. Für weitere Ausführungen wird auf einschlägige Literatur wie zum Beispiel [3, Kap. 10.7] verwiesen.

2.2. Regressionsanalyse

Die Regressionsanalyse ist eine übliche Methode um funktionale Zusammenhänge zwischen einer abhängigen Variable y und einer/mehreren unabhängiger Variablen u zu untersuchen. Dabei wird häufig zwischen linearer Regression (z.B. $y(u) = ku$) und nicht-linearer Regression (z.B. $y(u) = e^{ku}$) unterschieden. Die lineare Regression wird gerne verwendet, da zum Beispiel die Berechnung der Parameter deutlich einfacher ist. Allerdings sind viele Zusammenhänge in der Praxis nichtlinear und können daher durch eine nichtlineare Regression besser beschrieben werden. Man beachte, dass $y(u) = ke^u$ einer linearen Regression entspricht, da der Parameter k linear eingeht.

Die Ausführungen in diesem Abschnitt beschränken sich auf **lineare Regression** und auf eine einzige unabhängige Variable. Die Methode kann aber leicht auf mehrere unabhängige Variablen erweitert werden, was als *multiple linear Regression* bezeichnet wird.

Angenommen es liegen jeweils N Daten einer unabhängigen Variable u und einer (vermutlich) davon abhängigen Variable y vor,

$$\mathbf{u} = [u_1 \quad u_2 \quad \cdots \quad u_N]^\top, \quad \mathbf{y} = [y_1 \quad y_2 \quad \cdots \quad y_N]^\top, \quad (2.8)$$

so kann für jeden Index (häufig Zeitschritt) i ein linearer Zusammenhang gemäß folgender Form angesetzt werden:

$$y_i = \alpha + \beta u_i + e_i = \hat{y}_i + e_i, \quad \text{mit } i = 1, 2, \dots, N, \quad (2.9)$$

2. Statistische Begriffe und Grundlagen

wobei mit \hat{y}_i die Schätzung bezeichnet wird. Man beachte, dass die (noch unbekannt) Parameter β und α für alle i konstant sind. Der Fehler (oder das Residuum) e_i entspricht laut (2.9) der Abweichung zwischen wahren Wert y_i und der Schätzung \hat{y}_i

$$e_i = y_i - \hat{y}_i, \quad (2.10)$$

und enthält alles, was durch den linearen Zusammenhang allein nicht beschrieben wird.

Für die weiteren Überlegungen wird auf eine Matrixschreibweise übergegangen. Somit kann (2.9) geschrieben werden als

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} + \beta \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix}, \quad (2.11)$$

$$\mathbf{y} = \alpha \mathbf{1} + \beta \mathbf{u} + \mathbf{e}, \quad (2.12)$$

$$\mathbf{y} = \underbrace{\begin{bmatrix} \mathbf{1} & \mathbf{u} \end{bmatrix}}_{=\mathbf{A}} \underbrace{\begin{bmatrix} \alpha \\ \beta \end{bmatrix}}_{=\mathbf{x}} + \mathbf{e}, \quad (2.13)$$

$$\mathbf{y} = \mathbf{Ax} + \mathbf{e} = \hat{\mathbf{y}} + \mathbf{e}. \quad (2.14)$$

Falls $N = 2$ ist, entspricht die Anzahl der Gleichungen der Anzahl der unbekannt Parameter und (2.14) kann als *bestimmtes*, lineares Gleichungssystem eindeutig gelöst werden. In diesem Fall ist der Fehler $\mathbf{e} = \mathbf{0}$. In der Praxis werden aber mehr Daten verwendet, womit das Gleichungssystem *überbestimmt* ist und im Allgemeinen nicht eindeutig gelöst werden kann. Da keine eindeutige Lösung existiert, wird für das Gleichungssystem ein Optimierungsproblem formuliert, dass den Fehler \mathbf{e} geeignet minimiert

$$\min_{\mathbf{x}} J(\mathbf{x}) = \min_{\mathbf{x}} J(\mathbf{e}(\mathbf{x})). \quad (2.15)$$

Hierbei entspricht $J(\mathbf{x})$ der Kostenfunktion, die über den Fehler $\mathbf{e} = \mathbf{y} - \mathbf{Ax}$ von dem gesuchten Parametervektor \mathbf{x} abhängt. In dieser Arbeit werden folgende gängige Kostenfunktionen betrachtet

$$\begin{aligned} \text{Ordinary-Least-Squares (OLS)} \quad J(\mathbf{x}) &= \sum_i e_i^2 = \mathbf{e}^\top \mathbf{e}, \\ \text{Weighted-Least-Squares (WLS)} \quad J(\mathbf{x}) &= \sum_i (w_i e_i)^2 = \mathbf{e}^\top \mathbf{W} \mathbf{e} \quad \text{mit} \quad \mathbf{W} = \text{diag}(w_i^2), \\ \text{Least-Absolute-Value (LAV)} \quad J(\mathbf{x}) &= \sum_i |e_i|. \end{aligned} \quad (2.16)$$

Bei WLS wird jeder Fehler e_i mit dem Gewicht w_i multipliziert, womit zum Beispiel der Einfluss aktuellerer Messwerten auf die Kostenfunktion erhöht werden kann.

Der optimale Parametervektor kann für OLS und WLS mittels Pseudo-Inverser berechnet werden. LAV wird auch als Least-Absolute-Deviation (LAD) bezeichnet, und

kann als Lineares Programm angeschrieben und gelöst werden. Wenn ein Optimierungsproblem als Lineares Programm geschrieben werden kann, so hat es keine Nebenminima und das globale Minimum wird gefunden, sofern eines existiert (siehe z.B. [4, Kap. 1.3]). In Anhang A.2 sind drei konkrete Möglichkeiten für die Lösung des überbestimmten Gleichungssystems zur Parameterbestimmung mittels Linearem Programm in Matlab angegeben.

Nachdem der optimale Parametervektor $\mathbf{x}^* = [\alpha \ \beta]^\top$ bestimmt wurde, kann gemäß (2.9) der Zusammenhang zwischen u und \hat{y} als Regressionsmodell angeschrieben werden:

$$\hat{y}(u) = \alpha + \beta u. \quad (2.17)$$

Zum einen kann mit einem Regressionsmodell die Struktur des Zusammenhangs $y(u)$ genauer untersucht werden (Regressionsanalyse). Wurde zum Beispiel ein lineares Regressionsmodell (2.17) bestimmt und das Residuum e entspricht weißem Rauschen mit vergleichsweise geringer Standardabweichung, so kann der Zusammenhang $y(u)$ durch $\hat{y}(u) = \alpha + \beta u$ gut wiedergegeben werden. Andernfalls kann ein nicht-linearer Ansatz probiert und wieder das Residuum analysiert werden, bis eine Funktion den Zusammenhang $y(u)$ ausreichend genau beschreibt. Zum anderen kann mit einem Regressionsmodell auch eine **Prognose** \hat{y} bestimmt werden, sofern die unabhängige Variable u für die Zukunft bekannt ist. Somit kommt es auf den Kontext an, ob mit \hat{y} eine Schätzung für die Messdaten oder eine Prognose für die Zukunft bezeichnet wird.

Beispiel 2.1. Die Theorie zur linearen Regression soll nun für die OLS- und LAV-Kostenfunktion anhand eines Beispiels veranschaulicht werden. Zusätzlich soll die Wahl der Kostenfunktion diskutiert und der Fehler untersucht werden.

Die willkürlichen Messdaten (Trainingsdaten) lauten

$$\mathbf{u} = [5 \ 40 \ 50 \ 90 \ 110 \ 160 \ 180 \ 200 \ 230 \ 210]^\top,$$

$$\mathbf{y} = [10 \ 30 \ 18 \ 20 \ 26 \ 31 \ 37 \ 28 \ 32 \ 41]^\top,$$

wobei die Paare (u_i, y_i) entsprechend u_i aufsteigend sortiert sind. Diese 10 Messwerte ($N = 10$) könnten auch unsortiert sein, solange die Paare (u_i, y_i) erhalten bleiben. Aus \mathbf{u} wird gemäß (2.13) die Messmatrix \mathbf{A} aufgestellt:

$$\mathbf{A} = [\mathbf{1} \ \mathbf{u}].$$

Hiermit wird der Fehler angeschrieben als

$$\mathbf{e} = \mathbf{y} - \mathbf{A}\mathbf{x}.$$

Durch Minimierung der entsprechenden Kostenfunktion aus (2.16) wird der Parametervektor \mathbf{x} optimal bestimmt. Somit ist ein funktionaler (linearer) Zusammenhang zwischen der Schätzung der abhängigen Variable \hat{y} und der unabhängigen Variable u gefunden. Für OLS lautet dieses Regressionsmodell entsprechend (2.17)

$$\hat{y}(u) = 15,6366 + 0,0915 u.$$

2. Statistische Begriffe und Grundlagen

Die Messdaten und die geschätzten Geraden für OLS und LAV sind in Abbildung 2.1 dargestellt. In Abbildung 2.2 ist der dazugehörige Fehler über alle Indizes aufgetragen. Betrachtet man den Verlauf der Messpunkte so ist die Annahme eines linearen Zusammenhangs zwischen y_i und u_i gerechtfertigt. Die Kostenfunktion entscheidet welche Gerade durch diese Messpunkte gelegt wird.

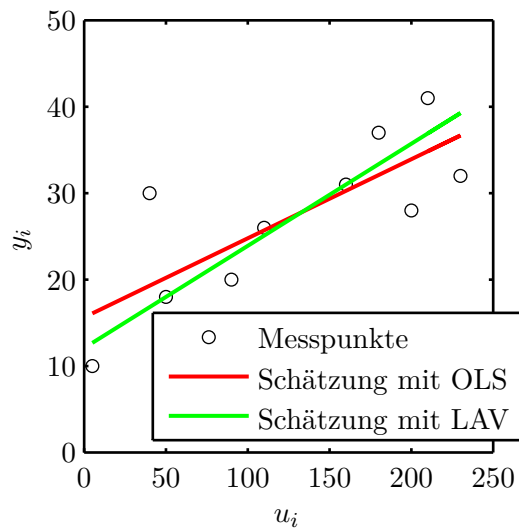


Abbildung 2.1.: Beispiel 2.1 – Fiktive Messpunkte inklusive Schätzungen mit OLS und LAV

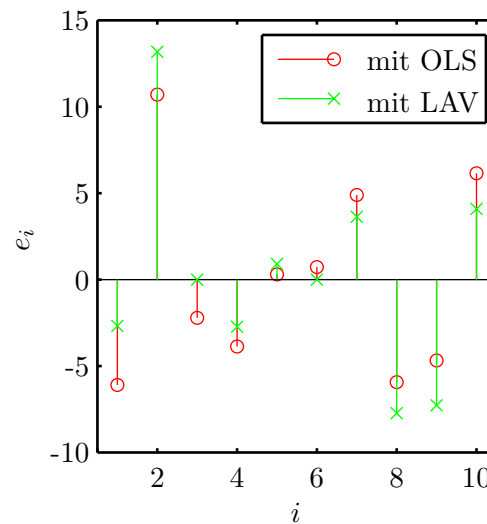


Abbildung 2.2.: Beispiel 2.1 – Fehler (Residuum) für Schätzung mit OLS und LAV

Dieses Beispiel ist so gewählt, dass für $i = 2$ ein Ausreißer enthalten ist, um die Eigenschaften der beiden Kostenfunktionen OLS und LAV zu veranschaulichen. Bei OLS wird durch das Quadrat in der Kostenfunktion ein großer Fehler e_2 deutlich stärker gewichtet als ein kleiner Fehler. Daher zielt die Optimierung darauf ab große Fehler zu reduzieren, wodurch die Gerade für OLS am Anfang angehoben wird. Bei LAV geht der Fehler unskaliert in die Kostenfunktion ein, was diese Kostenfunktion robuster gegenüber Ausreißern macht. In Abbildung 2.2 ist gut zu sehen, dass bei LAV größere Fehler e_i erlaubt sind, als bei OLS.

In diesem Beispiel wurde der Zusammenhang zwischen zwei Messdaten y und u durch eine lineare Regression der Form $\hat{y}(u) = \alpha + \beta u$ geschätzt. Für die Bestimmung der Parameter α und β wurden die OLS- und die LAV-Kostenfunktion verglichen, wobei sich die LAV-Kostenfunktionen als robuster herausgestellt hat, was bereits aus ihrer Definition (2.16) ersichtlich ist. Daher wird in dieser Arbeit die LAV-Kostenfunktion für die Bestimmung der Regressionsparameter eingesetzt. Für diese Bestimmung wird in dieser Arbeit die Matlab-Funktion `linprog` verwendet, da hierfür gezielt der Simplex-Algorithmus ausgewählt werden kann und sie nur eine geringe Rechenzeit benötigt (vgl. Anhang A.2).

Wären nun Werte u_i in der Zukunft bekannt, so können mit dem linearen Zusammenhang Prognosewerte \hat{y}_i bestimmt werden, was der Prädiktion in dieser Arbeit entspricht.

Die Qualität der Schätzung wurde nur visuell anhand der Übereinstimmung in Abbildung 2.1 bewertet und nicht anhand statistischer Auswertung des Residuum e evaluiert. Solch eine statische Auswertung enthält aufgrund der wenigen Trainingsdaten für die Parameterbestimmung keine Aussagekraft und wird daher auch in der restlichen Arbeit nicht durchgeführt.

Für eine umfassendere Einführung in die Regressionsanalyse ist [5, Kap. IV] empfehlenswert.

2.3. Zeitreihenanalyse

Bei der Zeitreihenanalyse ist die einzige unabhängige Variable die Zeit, womit sie ein Spezialfall der Regressionsanalyse ist. Eine Zeitreihe wird häufig entsprechend (2.2) durch Abtastung eines analogen Signals erzeugt und es stehen bei der reinen Zeitreihenanalyse keine weiteren Daten zur Verfügung. Dieser Fall tritt vor allem in den Wirtschaftswissenschaften häufig auf, wie zum Beispiel bei der Analyse von Börsenkursen.

In [6] wird ein guter Überblick der möglichen Methoden gegeben. Der Inhalt reicht von der Bestimmung des Trends bis zu Analysen im Frequenzbereich.

Ein Schwerpunkt der Zeitreihenanalyse ist die Beschreibung der Zeitreihe mit einem stochastischen Modell. Hierfür wird angenommen, dass die Zeitreihe y einem stochastischen Prozess entspringt, und es somit zielführend ist, diesen Prozess mathematisch zu beschreiben. Dieses mathematische Modell entspricht einer Übertragungsfunktion $H(z)$ deren Eingangsgröße $u[n]$ Weißes Rauschen ist. Die z-transformierte Schätzung für die Zeitreihe ergibt sich somit als

$$\hat{y}(z) = H(z) u(z). \quad (2.18)$$

Die Übertragungsfunktion $H(z)$ wird als gebrochen rationale Funktion in z angesetzt:

$$H(z) = \frac{\sum_{k=0}^q b_k z^{-k}}{1 + \sum_{k=1}^p a_k z^{-k}}, \quad (2.19)$$

und wird aufgrund ihrer Struktur als **Autoregressive-Moving-Average (ARMA) Prozess** bezeichnet. So gehen durch das Nennerpolynom die vergangenen p y -Werte gewichtet ein, was als Autoregression bezeichnet wird. Durch das Zählerpolynom wird ein gewichteter Mittelwert des aktuellen und der vergangenen $(q - 1)$ u -Werte gebildet, was als Moving-Average bezeichnet wird. Häufig wird in Klammern die Ordnung des Zähler- und Nennerpolynoms angegeben: ARMA(p, q).

Da das Eingangssignal $u[n]$ Weißes Rauschen ist, werden die unbekannt Parameter a_k und b_k nicht, wie in der Parameteridentifikation üblich, mittels OLS bestimmt (siehe z.B. [7, Kap. 7]). Stattdessen werden stochastische Eigenschaften der Zeitreihe y , wie die Autokorrelation, für die Bestimmung der Parameter verwendet. Die Parameter des AR-Teils können geschlossen bestimmt werden und für den MA-Teil ist eine mehrstufige Optimierung notwendig.

2. Statistische Begriffe und Grundlagen

Somit kann eine Zeitreihe mit einem ARMA-Prozess beschrieben werden der einer linearen Übertragungsfunktion im z -Bereich entspricht, deren Parameter statistisch bestimmt werden, da die Eingangsgröße Weißes Rauschen ist.

Für weitere Ausführungen siehe [8] oder [6, Kap. 6]. Bei letzterem muss beachtet werden, dass in der mathematischen Statistik die z -Transformierte üblicherweise anders als in den Ingenieurwissenschaften definiert ist.

2.4. Qualitätsmaße

In diesem Abschnitt werden drei Qualitätsmaße definiert, anhand derer die unterschiedlichen Prognosemethode-Varianten qualitativ verglichen, systematische Fehler in der Prognosemethode erkannt und die Ergebnisse der Prognose anschaulich präsentiert werden können. So wie bei einer Optimierung das Ergebnis von der Kostenfunktion abhängt, so hängt die finale Prognosemethode vom Qualitätsmaß für die Auswahl und Entwicklung der Prognosemethode-Varianten ab.

Als Qualitätsmaß werden üblicherweise Funktionen des Prognosefehlers betrachtet, der analog zum Residuum (2.10) definiert ist. In diesem Abschnitt bezeichnet \hat{y} die Prognose für y , weshalb der dazugehörige Fehler als Prognosefehler und nicht als Residuum bezeichnet wird.

Das erste Qualitätsmaß ist der **Mean-Absolute-Error (MAE)** welcher als

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, \quad (2.20)$$

definiert ist und einem absoluten Qualitätsmaß entspricht. Zusätzlich werden zwei relative Qualitätsmaße betrachtet: Der **Mean-Percentage-Error (MPE)**

$$\text{MPE} = \frac{1}{N} \sum_{i=1}^N \frac{y_i - \hat{y}_i}{y_i} \cdot 100\%, \quad (2.21)$$

und der **Mean-Absolute-Percentage-Error (MAPE)**

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100\%. \quad (2.22)$$

Die Eigenschaften dieser drei Qualitätsmaße sollen anhand eines Beispiels veranschaulicht werden.

Beispiel 2.2. Zwei Datenvektoren werden anhand der beschriebenen Qualitätsmaße miteinander verglichen und das Ergebnis interpretiert. Als Istverlauf wird eine Gerade durch den Ursprung mit Steigung 1 betrachtet. Die Schätzung (Prognose) für diese Gerade entspricht einer um 10 nach oben geschobenen Geraden. Somit ist der Fehler $e_i = y_i - \hat{y}_i = -10$ für alle Indizes i . Es ist zu beachten, dass mit dieser Definition des Fehlers eine zu große Schätzung zu einem negativen Fehler führt.

Da MPE und MAPE für $y \rightarrow 0$ gegen $\pm\infty$ streben, wird für die Betrachtung eine Schranke bei 10 eingeführt. Alle Daten unter dieser Schranke werden für die Berechnung von MPE und MAPE ignoriert. Die Daten und die Schranke sind in Abbildung 2.3 dargestellt.

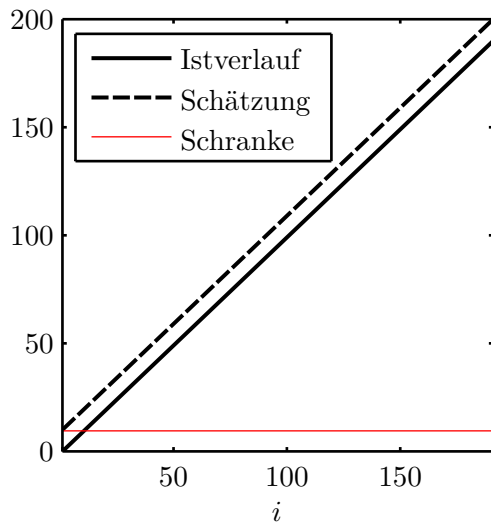


Abbildung 2.3.: Beispiel 2.2 – Fiktive Daten und Schranke für MPE und MAPE

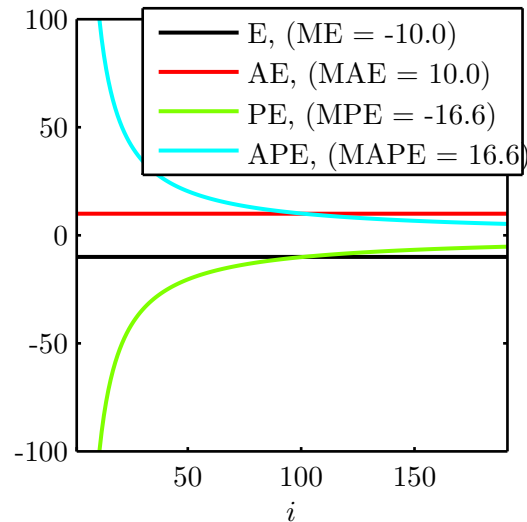


Abbildung 2.4.: Beispiel 2.2 – Qualitätsmaße

Um die Natur der Qualitätsmaße zu untersuchen wird auf die Mittelwertbildung verzichtet. Somit wird zum Beispiel aus dem MAE ein AE $|y_i - \hat{y}_i|$, der den absoluten Fehler zu jedem Index i angibt. Der Verlauf aller Qualitätsmaße ohne Mittelwertbildung ist in Abbildung 2.4 dargestellt. Man sieht gut, dass der AE unabhängig von der Größe des Istwerts ist und dem Betrag des Fehlers $y_i - \hat{y}_i$ (engl. Error bzw. E) entspricht. Ganz anders PE und APE, die, bei konstantem Fehler, für kleinere Istwerte stark ansteigen. Hier ist auch die Notwendigkeit der Schranke ersichtlich, da sonst der PE und APE gegen $\pm\infty$ streben. In der Legende sind zusätzlich Mean-Error (ME), MAE, MPE und MAPE angegeben, da für die Bewertung der Qualität schlussendlich einzelne Zahlenwerte und keine Kurvenverläufe benötigt werden. Gemäß der Definitionen sind ME bzw. MA in der Einheit der Daten, und MPE bzw. MAPE in % angegeben. An diesem Beispiel (und schon anhand der Definition) ist ersichtlich, dass für den MPE auch negative Prozentzahlen möglich sind.

Alle drei Qualitätsmaße (MAE, MPE und MAPE) haben also bestimmte Eigenschaften, die ihre Anwendbarkeit einschränken. Der MAE bewertet den Fehler absolut zum Istverlauf, wodurch ein Prognosefehler unskaliert eingeht; er liefert aber eine schlecht vergleichbare Zahl in der Größenordnung und Einheit der Daten. MPE und MAPE be-

2. Statistische Begriffe und Grundlagen

werten den Fehler relativ zum Istverlauf, wodurch der Prognosefehler mit dem Istwert skaliert eingeht; sie liefern aber eine gut vergleichbare Prozentzahl.

Für die Auswahl und Entwicklung der Prognosemethode-Varianten sollte ein Qualitätsmaß eingesetzt werden, welches erstens den Prognosefehler nicht skaliert und welches zweitens auch zur LAV-Kostenfunktion für die Regressionsparameter passt. Der Prognosefehler sollte nicht skaliert werden, da eine Abweichung von Prognose zu Istverlauf unabhängig vom Wertebereich zu einem Wärmeüberschuss bzw. -mangel im Pufferspeicher führt. Da die relativen Qualitätsmaße MPE und MAPE den Prognosefehler mit dem Istwert skalieren erfüllt nur der MAE die erste Anforderung. Die verwendete LAV-Kostenfunktion minimiert gemäß (2.16) die Summe der Absolutbeträge des Residuums $\sum_i |y_i - \hat{y}_i|$, was dem MAE sehr ähnlich ist, der gemäß (2.20) statt der Summe den Mittelwert bildet $1/N \sum_{i=1}^N |y_i - \hat{y}_i|$. Somit liegt der Unterschied zwischen LAV und MAE, neben ihrer unterschiedlichen Anwendung, im skalaren Faktor $1/N$ und der Bezeichnung von $y - \hat{y}$ als Residuum bzw. als Prognosefehler. Da der MAE somit erstens den Prognosefehler unskaliert verwendet und zweitens zur LAV-Kostenfunktion passt, wird er für die Auswahl und Entwicklung der Prognosemethode-Varianten eingesetzt.

Der MPE wird verwendet um systematische Fehler in der Prognosemethode zu erkennen. Wird zum Beispiel immer zu wenig prognostiziert, so würde das Vorzeichen dieses systematischen Fehlers durch den Absolutbetrag im MAPE eliminiert werden, scheint aber im MPE auf.

Für alle Auswertungen, die in dieser Arbeit als Ergebnis präsentiert werden, wird der MAPE herangezogen, da dieser intuitive Prozentzahlen liefert. Somit sind die Ergebnisse innerhalb dieser Arbeit für verschiedene Verbraucher vergleichbar, sind Außenstehenden gut vermittelbar und können zusätzlich mit den Ergebnissen in den meisten wissenschaftlichen Artikeln verglichen werden, da dort häufig der MAPE verwendet wird [9].

Eine ausführliche Beschreibung und Diskussion gängiger Qualitätsmaße für die Prognose ist in [9] zu finden. Darin wird unter Anderem ein Qualitätsmaß empfohlen, bei dem das Ergebnis einer simplen Methode als Vergleichsgröße betrachtet wird. In anderen Worten, es wird bewertet, um wie viel eine betrachtete Methode besser oder schlechter ist als die simple Methode. Da für die in dieser Arbeit betrachtete Anwendung noch keine simple Methode etabliert ist, hat so ein Qualitätsmaß hier keinen Sinn. In einigen wissenschaftlichen Artikeln wird statt des MAPE die Verwendung eines symmetric-MAPE (sMAPE) empfohlen. Dieser hat jedoch – entgegen seines Namens – ein asymmetrisches Verhalten und bewertet positive und negative Fehler deutlich anders, wie in [10] nachgelesen werden kann. Daher wird in dieser Arbeit der MAPE statt des sMAPE verwendet.

3. Prognosemethoden in der Literatur

In diesem Kapitel wird zuerst ein ausführlicher Überblick über die in der Literatur üblichen Methoden für einen Prädiktor gegeben. Danach wird kurz auf mögliche Methoden für einen Korrektor eingegangen.

Sowohl die Prädiktion, die im Allgemeinen als Prognose bezeichnet wird, als auch die Korrektur, häufig als Bias-Korrektur bezeichnet, sind eigenständige Gebiete, und werden deshalb auch hier getrennt behandelt. So kommt die Prädiktion als Prognose in der Regel ohne Korrektur zum Einsatz, wohingegen die Korrektur eine vorangegangene Prädiktion voraussetzt.

3.1. Prädiktor

Es soll ein Überblick der Methoden zur Lastprognose, genauer gesagt zur Lastprädiktion, gegeben werden. Ein Großteil der untersuchten wissenschaftlichen Artikel kommt aus dem Gebiet der elektrischen Lastprognose, da dies ein stark beforschtes Gebiet ist. Diese Lastprognose hat einen großen Einfluss auf die Netzstabilität und letztendlich die Kosten. So wurden beispielsweise in [1] bereits 1987 gängige Methoden zur Prognose der elektrischen Last verglichen und auf ein Paper von 1966 referenziert. Das Thema der Wärmelastprognose steht erst in den letzten Jahren im Fokus der Wissenschaft, wobei hier die Dissertation [11] aus dem Jahr 2000 einen frühen Beitrag darstellt.

Nach diesem kurzen Exkurs in die Geschichte ist klar, dass insbesondere im Bereich der elektrischen Lastprognose bereits enormes Wissen angesammelt wurde. Dieses Wissen kann für die Wärmelastprognose eingesetzt werden, wobei folgende Unterschiede zu beachten sind:

- Gebäude und Wärmenetze reagieren viel **träger** als elektrische Netze.
- Die Spitzenlast ist bei Wärme deutlich **unkritischer**, da diese durch thermische Speicher kompensiert werden kann. Bei einem elektrischen Netz gefährdet eine Nicht-Abdeckung der Spitzenlast die Netzsicherheit (siehe z.B. [1]).
- Die Wärmelast hat eine stärkere **saisonale** Komponente und andere Einflussgrößen als die elektrische Last.

Unter Berücksichtigung dieser Unterschiede können die Methoden der elektrischen Lastprognose als Basis für Methoden der Wärmelastprognose verwendet werden.

Im Folgenden sind Methoden aus der Wärmelastprognose und der elektrischen Lastprognose angeführt. Die grobe Einteilung erfolgt in Klassische Methoden, die auf Ansätzen mit klassischen Funktionen basieren, und Künstliche-Intelligenz Methoden, wo

3. Prognosemethoden in der Literatur

die Recheneinheit anhand von vordefiniertem Wissen bzw. durch Lernalgorithmen die Lastverläufe vorhersagt. Nach diesem Überblick werden die einzelnen Methoden diskutiert und auszugsweise näher beschrieben. Dieser grobe Überblick wird in Abschnitt 5.1 noch vertieft. Dort werden die Erkenntnisse der Messdatenauswertung mit deren Umsetzung in Prognosemethoden der Literatur verknüpft, woraus dann der erste Prädiktor resultiert.

Ein guter Überblick über in der elektrischen Lastprognose verwendete Methoden wird in [12] gegeben, auf dem auch die Unterteilung in diesem Kapitel basiert. In [13] werden 5 Methoden beschrieben und miteinander verglichen, und [1] liefert eine noch immer gute, praxisrelevante Zusammenfassung des Wissensstandes von 1987.

3.1.1. Klassische Methoden

In den klassischen Methoden wird versucht den Lastverlauf als Funktion der Zeit und anderer Faktoren zu beschreiben. Hierbei wird in einem ersten Schritt der funktionale Zusammenhang (Ansatz) festgelegt. In einem zweiten Schritt werden dann die Parameter der Funktion bestimmt. So könnte zum Beispiel der Ansatz $y(t, u) = k_1 \sin(t) + k_2 u$ lauten, wo in einem zweiten Schritt die unbekannt Parameter k_1 und k_2 bestimmt werden.

Die klassischen Methoden der Regression und stochastischen Modelle werden kombiniert in [14, 15] für die Wärmelastprognose und in [16] für die elektrische Lastprognose eingesetzt.

Regression

Die Regression (siehe Abschnitt 2.2) entspricht einer intuitiven Herangehensweise, die der physikalischen Modellbildung ähnelt. Es wird versucht den unbekannt Lastverlauf mit anderen bekannten Faktoren (Wetter, Kalendertag, etc.) in Zusammenhang zu bringen, und diesen mit einer Funktion zu beschreiben.

In [17, 18] sind ältere und in [19, 20] modernere Ansätze für die elektrische Lastprognose zu finden. Zur Wärmelastprognose wird in [21] eine simple Methode vorgestellt.

Zeitreihen und stochastische Modelle

Die Zeitreihenanalyse (siehe Abschnitt 2.3) betrachtet ausschließlich den zeitlichen Lastverlauf und versucht hierfür eine mathematische Beschreibung zu finden. Sie wird angewandt, wenn die Einflussfaktoren unbekannt sind bzw. ihr Einfluss nicht ausreichend mathematisch beschrieben werden kann.

Zeitreihen werden häufig mit stochastischen Modellen beschrieben, wobei hier in der Regel ein ARMA-Prozess verwendet wird, von dem es viele Erweiterungen gibt. Die tatsächliche Prognose erfolgt dann zum Beispiel mit einem Kalman-Filter, der die Zustände des ARMA in Zustandsraumdarstellung schätzt.

So wird in [22] ein aufwendiges Seasonal-Autoregressive-Integrated-Moving-Average (SARIMA) Zustandsraummodell inklusive Kalman-Filter verwendet. Weitere ARMA

Kombinationen finden sich beispielsweise in [23, 24, 25, 26], wobei zum Teil auf fertige Toolboxes zurück gegriffen wird.

3.1.2. Künstliche-Intelligenz-Methoden

Die Künstliche-Intelligenz-Methoden unterscheiden sich von den klassischen Methoden darin, dass der mathematische Zusammenhang nicht geschlossen vorgegeben wird. Die Recheneinheit „sucht“ sich den Zusammenhang selbst, wobei dieser oft mit einem klassischen Ansatz (z.B. $y(t, u) = k_1 \sin(t) + k_2 u$) nichts mehr gemein hat, sondern zum Beispiel auch auf der Verknüpfung großer Datenbanken basieren kann. Diese (scheinbare) Künstliche-Intelligenz entspricht somit der Fähigkeit der Recheneinheit enorme Datenmengen in kurzer Zeit zu verarbeiten und zu verknüpfen. Da bei diesen Methoden die Einsicht in das Verhalten des zu beschreibenden Systems verloren geht, werden diese Systembeschreibungen auch als Black-Box bezeichnet.

In der Anwendung werden häufig Künstliche-Intelligenz-Methoden miteinander kombiniert, wobei zum Beispiel Similar-Day-Approach gemeinsam mit Fuzzy-Logik [27, 28], und zusätzlich noch mit Neuronalen-Netzen zum Einsatz kommt [29]. Auch Kombinationen mit klassischen Methoden [30, 31] kommen vor, sind aber die Ausnahme.

Neuronale-Netze

Neuronale-Netze (NN) werden eingesetzt, wenn ein Zusammenhang zu anderen Faktoren angenommen wird, aber nicht weiter beschrieben werden kann oder will. Es handelt sich um eine nichtlineare Black-Box-Variante der Regressions-Methode. Durch ausreichende Datenmengen „lernt“ das NN den Zusammenhang zwischen Einflussfaktoren und Last und kann dann für die Prognose eingesetzt werden.

Sehr viele wissenschaftliche Artikel im Bereich der elektrischen Lastprognose, wie beispielsweise [32, 33, 34], beschäftigen sich mit diesem Thema. In der Wärmelastprognose setzen erst 2014 [35, 36] auf NN.

Experten-Systeme

Ein Experten-System basiert auf Regeln nach denen auch ein Experte entscheiden würde. So werden Faustregeln, Erfahrungswerte und anderes Expertenwissen in der Recheneinheit hinterlegt, womit diese eine Prognose erstellt.

Diese Methode wird häufig in Verbindung mit anderen Methoden eingesetzt, um die Ergebnisse für bestimmte Szenarien (z.B. Weihnachtsabend) zu verbessern.

Fuzzy-Logik

Mit der Fuzzy-Logik können unscharfe Begriffe wie „kalt - warm - heiß“ mathematisch modelliert werden. Dies wird, wie die Experten-Systeme, häufig in Verbindung mit anderen Methoden eingesetzt.

3. Prognosemethoden in der Literatur

Similar-Day-Approach

In einer umfassenden Datenbank wird nach einem ähnlichen historischen Tag wie dem Prognosetag gesucht. Wurde in der Datenbank ein Tag gefunden, der ausreichend gut mit dem Prognosetag übereinstimmt, so wird der damalige Lastverlauf als Prognose für den Prognosetag verwendet. Hierbei wird ein Modell benötigt, um die Überprüfung auf Ähnlichkeit durchzuführen (Mustererkennung).

Dieser Ansatz wird wieder häufig mit anderen Methoden kombiniert, kommt aber in [37, 38] auch alleine zur Anwendung. In [38] wird diese Methode für die bessere Vorauswahl von Trainingsdaten ausgewählt, auf die dann wieder der Similar-Day-Approach angewandt wird. Diese bessere Vorauswahl von Trainingsdaten ist aber auch für andere Methoden denkbar.

Support-Vector-Machines

Dieser Ansatz basiert auf einer mathematisch anspruchsvollen Theorie, deren Ziel es ist Objektmengen durch Hyperebenen (linear) zu trennen. Da diese lineare Trennung in der Anwendung fast nie möglich ist, wird ein *Kernel* verwendet. Durch diesen Kernel werden die Objekte in einen hoch-dimensionalen Raum transformiert, in dem sie linear trennbar sind.

Für die Lastprognose bedeutet dies, dass nichtlineare Regression in einen höherdimensionalen Raum transformiert wird, in dem eine lineare Regression durchgeführt werden kann. Diese Methode liefert eine nichtlineare Black-Box (ähnliche einem NN) mittels der die Last prognostiziert werden kann. Sie wird im Vergleich zu NN deutlich seltener eingesetzt, aber zum Beispiel in [39] verwendet um die Kühllast eines Gebäudes zu prognostizieren.

3.1.3. Diskussion

In den klassischen Methoden geht der Methodenentwicklung in der Regel eine intensive Auswertung der Messdaten voraus. Ein Musterbeispiel hierfür ist [19] wo mit einem mächtigen Softwarepaket von R¹ eine bemerkenswerte Auswertung durchgeführt wird. Solch eine Auswertung der Messdaten ist notwendig, um den Ansatz gezielt zu wählen, und liefert zusätzlich tiefe Einblicke in die untersuchten Daten.

Betrachtet man die zeitliche Entwicklung der wissenschaftlichen Artikel, so werden mit zunehmender Rechenleistung immer komplexere Methoden eingesetzt. Laut [1, S. 1565] wurden in der elektrischen Lastprognose bereits im Jahr 1987 alle statischen Regressions-Methoden durch dynamische Methoden ersetzt, da diese zeitabhängige, zufällige Effekte und Wettereinflüsse besser berücksichtigen können. Der damalige Stand der Technik entspricht einer Mischung klassischer Methoden [1, Formeln (4),(5),(6)]. Hierbei wird die Last beschrieben als additives Modell der Form

$$\hat{y} = f_1 + f_2. \quad (3.1)$$

¹R ist ein Werkzeug zur statistischen Datenanalyse und zur grafischen Darstellung der Daten bzw. der Ergebnisse. <http://www.r-project.org>

Die Funktion f_1 entspricht einer Summe von Sinusfunktionen mit einer Periode von 1 und 7 Tagen und beschreibt somit grob den täglichen und wöchentlichen Verlauf. Der Rest wird mit f_2 als ARMA Modell mit Wetterdaten als Eingang modelliert. In moderneren Ansätzen wird der periodische Anteil (hier f_1) bzw. die Saison (S) im ARMA Modell inkludiert, was als Seasonal-ARMA (SARMA) bezeichnet wird. Wird noch durch Ableitung (I) der Trend eliminiert und werden externe Eingänge (X), wie bei f_2 , berücksichtigt, so erhält man den aktuellen Stand der Technik – ein SARIMAX Modell wie in [24] von 2013.

In den Künstliche-Intelligenz-Methoden befassen sich die Autoren kaum mit der statistischen Auswertung der Daten. Stattdessen werden große Datenmengen herangezogen und durch Performance-Vergleich verschiedener Konfiguration wird die passende Methode gefunden.

In den betrachteten wissenschaftlichen Artikeln werden verschiedenste Methoden, von NN bis Support-Vector-Machines, angewendet. Die Autoren finden auch immer Konfigurationen, in denen die neue Methode bessere Ergebnisse liefert als etablierte Methoden. Diese Resultate sind aber kritisch zu hinterfragen, da das Ergebnis maßgeblich von der Verfügbarkeit großer Datenmengen abhängt, häufig nur ein spezifischer Lastverlauf untersucht wird und somit die neue Methode möglichst gut an das spezifische Problem angepasst wird.

Zusammenfassend ist in den wissenschaftlichen Artikeln ein Trend zu immer rechenintensiveren Methoden zu erkennen. Ob diese Entwicklung auch auf die Praxis zutrifft ist fraglich, da hier die Rechenzeit und Stabilität der reinen Prognosequalität gegenüberstehen. In [21] werden aufwändige Methoden für die Lastprognose hinterfragt. Eine Kernaussage ist hier, dass zunächst die Wettervorhersage zuverlässiger werden muss, bevor komplexe Methoden für die Lastprognose zur Anwendung kommen können.

3.2. Korrektor

In dieser Arbeit wird eine Prädiktion mit einer Korrektur – dem Korrektor – verbessert. Diese Trennung ermöglicht eine schrittweise Evaluierung, eine einfache Implementierung und klare Erweiterungsmöglichkeiten. Diese Prädiktor-Korrektor-Methode wird in den verschiedensten Gebieten erfolgreich angewandt. Zum Beispiel basieren der asymptotische Beobachter in der Regelungstechnik (siehe z.B. [40, S. 444]) und das Runge-Kutta-Verfahren zur numerischen Lösung von Differentialgleichungen (siehe z.B. [41, Kap. 3]) auf diesem Ansatz.

In keinem der wissenschaftlichen Artikel zum Thema „Lastprognose“ (hier Prädiktor) wird ein Korrektor beschrieben oder evaluiert. Dies ist darauf zurückzuführen, dass der Korrektor in den Prädiktor inkludiert werden kann, wodurch aber die vorteilhaften Eigenschaften der Trennung verloren gehen. Allerdings findet sich in [1, S.1568] von 1987 zum Thema „Echtzeitmodus“ eine allgemeine Beschreibung, die auch auf eine

3. Prognosemethoden in der Literatur

Prädiktor-Korrektor-Methode zutrifft und durch die damals stark begrenzte Rechenleistung motiviert sein dürfte:

The 24 h forecast must be generated at least once a day. In addition, in the real-time mode, there may be frequent re-forecasting whenever weather forecasts change markedly, abnormal events occur, telemetered data indicate a significant deviation of the values of the actual load from the forecast ones, or simply to update and refine the current day's forecast based on the most recent load and weather information.

In diesem Zitat beschreibt der „24 h forecast“ eine Prädiktion und der Rest mögliche Anwendungen eines Korrektors.

Ein Korrektor scheint aber in anderen Prognose-Disziplinen nur zur Anwendung zu kommen, wenn die häufige Prädiktion zu rechenaufwendig ist, oder lokale Ergebnisse einer globalen Prädiktion verbessert werden sollen. Hier haben alle betrachteten Methoden gemein, dass sie den Prädiktionsfehler der vergangenen Zeitschritte für die Korrektur der aktuellen Prognose verwenden, und somit der systematische Fehler (Bias) der Prognosemethode reduziert wird.

Im Bereich der Wirtschaftswissenschaften stellte 2015 [42] eine Methode vor, die sich eines Autoregressiven (AR) Prozesses für den Prädiktionsfehler bedient, um die Prognose zu verbessern.

Auch in der Wetterprognose wird zum Teil ein Korrektor eingesetzt, da diese Modelle sehr rechenintensiv sind. So verwendet [43] den Best-Easy-Systematic-Estimator, der aus den ersten drei Quartilen des Prädiktionsfehlers der vergangenen Tage berechnet wird. In [44] wird der Prädiktionsfehler-Mittelwert der vergangenen Tage für die Korrektur herangezogen.

4. Messdaten

Einen Schwerpunkt dieser Arbeit stellt die Auswertung von Messdaten verschiedenster Verbraucher dar. Diese Auswertung liefert Erkenntnisse, auf der die spätere Prognosemethode basiert. Sofern nicht anders erwähnt, betreffen die angeführten Schritte den Wärmestrom \dot{Q} , den es schlussendlich zu prognostizieren gilt.

Im Allgemeinen setzt sich dieser Wärmestrom aus der benötigten Wärme für das Warmwasser der Heizung, sowie zusätzliches Warmwasser zusammen. Je nach Verbraucher wird dieses zusätzliche Warmwasser für den Küchen- und Sanitärbereich (Duschen, etc.) oder auch als Prozesswärme benötigt.

Im Folgenden wird zuerst der Wärmestrom als Messgröße genauer betrachtet, und danach wird die Aufbereitung der Daten motiviert und beschrieben. Abschließend werden die Messdaten ausgewertet und daraus Schlussfolgerungen für die Wahl der Prognosemethode abgeleitet.

4.1. Wärmestrom als Messgröße

Der Wärmestrom wird nicht direkt gemessen, lässt sich aber aus messbaren Größen berechnen:

$$\dot{Q} = c_W \dot{m}_W (T_{VL} - T_{RL}), \quad (4.1)$$

wobei c_W die spezifische Wärmekapazität von Wasser, \dot{m}_W den Massenstrom des Wassers und T_{VL} bzw. T_{RL} die Vorlauf- bzw. Rücklauf-temperatur des Wassers beschreibt. Der Massenstrom des Wasser lässt sich wiederum mittels

$$\dot{m}_W = \rho_W \dot{V}_W \quad (4.2)$$

aus der Dichte von Wasser ρ_W und aus dem Volumenstrom des Wassers \dot{V}_W berechnen. Die spezifische Wärmekapazität bzw. die Dichte von Wasser sind temperatur- und druckabhängig (siehe z.B. [45, Db2]) und betragen beispielsweise bei einer Temperatur von 60 °C und einem Druck von 1 bar:

$$c_W = 4,184 \text{ kJ/kg K} \quad \text{bzw.} \quad \rho_W = 983,20 \text{ kg/m}^3. \quad (4.3)$$

In der betrachteten Anwendung wird die Abhängigkeit dieser Größen aber häufig vernachlässigt und sie werden innerhalb des betrachteten Bereichs als konstant angenommen. Somit kann der Wärmestrom aus messbaren Größen und bekannten Konstanten bestimmt werden. Die üblichen Messsysteme (Wärmemengenzähler) liefern den Wärmestrom bereits als berechnete Größe. Der Wärmestrom hat die Einheit Watt W und wird in der Praxis auch häufig als Leistung bezeichnet.

4.2. Klassifizierung der Verbraucher

In dieser Arbeit werden sieben Verbraucher betrachtet, wobei das Spektrum vom einfachen Wohnhaus bis zum Fernwärmenetz reicht. In Tabelle 4.1 sind die Verbraucher mit Klasse und Anschlussleistung angeführt.

Verbraucherbezeichnung	Verbraucherklasse	Anschlussleistung
VB1	Wohnhaus	35 kW
VB2	Bürogebäude	300 kW
VB3	Bürogebäude und Gewerbe	125 kW und 17 kW
VB4	Doppelhaus	22 kW
VB5	Mikrowärmenetz	78 kW
VB6	Nahwärmenetz	2,1 MW
VB7	Fernwärmenetz	30 MW

Tabelle 4.1.: Verbraucher mit Bezeichnung, Klasse und Nennleistung.

Die Daten der Verbraucher sind nur für bestimmte Zeiträume in den Jahren 2012, 2013 und 2014 vorhanden. Daher sind in den nachfolgenden Auswertungen einige Verbraucher in bestimmten Zeiträumen ausgelassen. Insgesamt wird aber versucht, für jede Auswertung, aus der eine Entscheidung für den Algorithmus abgeleitet wird, alle Verbraucher gleichermaßen zu berücksichtigen. Für eine Aufschlüsselung, in welchem Zeitraum Daten der jeweiligen Verbraucher vorhanden sind, sei auf die Tabellen A.1, A.1 und A.2 im Anhang A.1 verwiesen.

4.3. Messdatenaufbereitung

Die Messdaten müssen für die spätere Auswertung aufbereitet werden. Hierfür werden Datenlücken durch lineare Interpolation zwischen den Nachbarwerten behoben.

Nun ist es noch notwendig die Messdaten zu filtern, damit die essentiellen Dynamiken analysiert und schlussendlich prognostiziert werden können. Hierfür wird ein gleitender Mittelwert mit einem Blackman-Fenster verwendet, welches sich zum Rand hin verjüngt, womit die Daten am Rand schwächer in die Mittelwertbildung eingehen. Dies entspricht einem akausalen Filter, welches keine Phasenverschiebung hervorruft.

Die Notwendigkeit der Filterung ist in Abbildung 4.1 leicht nachzuvollziehen, wo der ungefilterte und der gefilterte Wärmestrom für VB2 einander gegenübergestellt werden. In der Abbildung ist klar zu erkennen, dass die Spitzen des Signals stark gedämpft werden. Dies ist bei der Wärmelast zulässig, da erstens eine Nicht-Abdeckung dieser Spitzen kaum Auswirkungen auf die Wärmeverbraucher hat und zweitens kurzzeitige Lastspitzen in der Wärmelast durch den Pufferspeicher ausgeglichen werden können. Daher ist es ausreichend einen groben Verlauf zu prognostizieren und somit auch nur die gefilterte Last genauer zu untersuchen. Diese Eigenschaften der Wärmelast sind ein großer Unterschied zur elektrischen Last, wie in Kapitel 3 beschrieben.

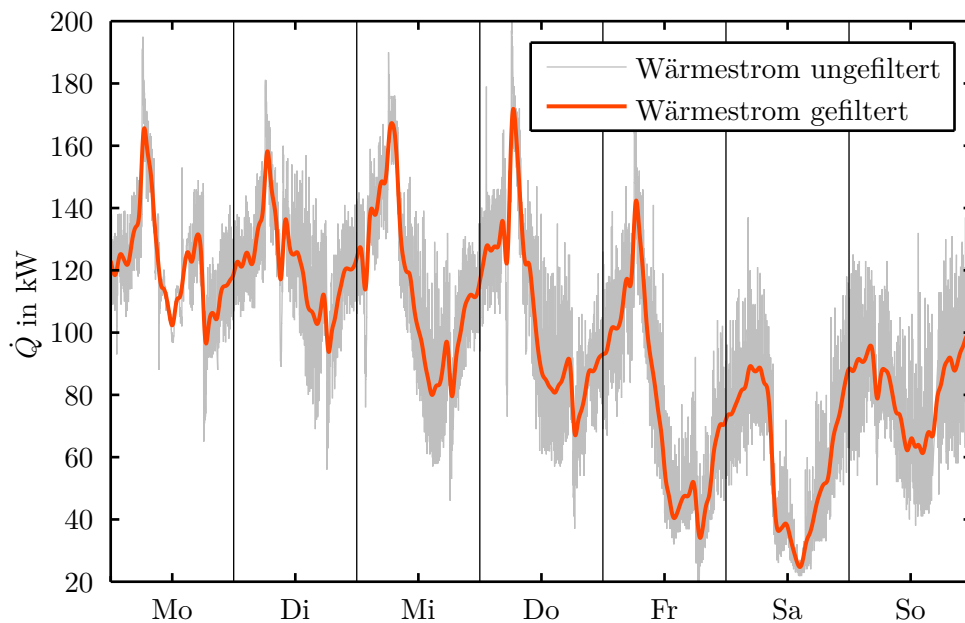


Abbildung 4.1.: VB2 – Wärmestrom einer Woche ungefiltert und gefiltert

4.4. Messdatenauswertung

Die Auswertung der Messdaten stellt einen entscheidenden Schritt dar, da sie das Fundament für die spätere Prognosemethode liefert. Dabei wird das zeitliche Verhalten des Wärmestroms, sowie seine Abhängigkeit von Wetterfaktoren ausführlich analysiert und ein Einblick in die Verbraucherdaten gegeben.

Aufgrund der großen Datenmengen werden nur exemplarische Auswertungen beschränkter Zeiträume einzelner Verbraucher angeführt. Diese Auswertungen wurden aber für alle Verbraucher durchgeführt und lieferten vergleichbare Ergebnisse.

Als Erstes wird der vermutlich starke Einfluss des Wetters untersucht, wobei zunächst die Auswahl der Wettergrößen explizit motiviert wird. Als Zweites werden zeitliche Einflüsse untersucht, die auf Benutzerverhalten und Gebäudeautomation zurückzuführen sind. Zuletzt wird der Einfluss des vergangenen Lastverlaufs untersucht, wobei hier ebenfalls zeitliche Zusammenhänge eine Rolle spielen.

4.4.1. Einfluss des Wetters

Das Wetter hat einen entscheidenden Einfluss auf die benötigte Wärme für die Heizung der Gebäude (kurz: Heizlast). Zum einen ist dies aus physikalischen Überlegungen und persönlichen Erfahrungen einsichtig. So dürfte an einem sonnigen, warmen Tag weniger Wärme benötigt werden, als an einem regnerischen, kalten Tag. Zum anderen wird in der Literatur das Wetter als die dominante Eingangsgröße für Prognosemethoden verwendet,

4. Messdaten

wobei dies sowohl auf die elektrische, als auch auf die Wärme-Last zutrifft. Das Wetter wird auch deswegen häufig verwendet, weil durch Wetterprognosen eine Schätzung für die Zukunft vorhanden ist.

Nun kann das Wetter durch viele verschiedene Messgrößen, wie Außentemperatur, Luftfeuchte oder Windgeschwindigkeit, beschrieben werden. Für die Wärmelastprognose sind aber nur Messgrößen interessant, welche einen entscheidenden Einfluss auf die Wärmelast haben und ausreichend genau prognostiziert werden können. Man beachte, dass für die Wärmelastprognose eine Wetterprognose benötigt wird, sofern Wetterdaten als Eingangsgrößen verwendet werden.

Den stärksten Einfluss auf den Wärmebedarf hat die Außentemperatur. Dies ist leicht nachvollziehbar, wenn man den Wärmebedarf im Sommer bei 35 °C mit dem Wärmebedarf im Winter bei -10 °C vergleicht. Die Sonneneinstrahlung wärmt die Gebäude, und durch die Fenster auch direkt die Innenräume, und hat somit ebenfalls einen Einfluss auf den Wärmebedarf. Schlussendlich hat die Windgeschwindigkeit noch einen Einfluss, der jedoch schwierig zu prognostizieren ist, da der Wind durch umliegende Vegetation bzw. Gebäude je nach Windrichtung abgeschwächt oder sogar abgeblockt werden kann.

Aus diesen Überlegungen folgt, dass die Außentemperatur und die Sonneneinstrahlung (Globalstrahlung) als mögliche Eingangsgrößen für die Prognosemethode näher untersucht werden sollen. Da die zu entwickelnde Prognosemethode für möglichst viele Verbraucher, ohne aufwendige Parametrierung, eingesetzt werden soll, wird die Windgeschwindigkeit nicht betrachtet.

Einfluss der Außentemperatur

Die **Außentemperatur** T_A hat einen entscheidenden Einfluss auf den Wärmebedarf eines Gebäudes und wird daher als Erstes untersucht. Da der Einfluss der Außentemperatur auf den Wärmestrom untersucht werden sollen, werden diese beiden Größen gemeinsam betrachtet. Es ist davon auszugehen, dass eine steigende Außentemperatur zu einer sinkenden Wärmelast führt und umgekehrt.

Zunächst wird die Außentemperatur zum gefilterten Lastverlauf aus Abbildung 4.1 betrachtet. Dies ist in Abbildung 4.2 zu sehen, wobei die Außentemperatur eine separate Skala auf der rechten Seite hat. Die Verläufe bestätigen die Vermutung, des Zusammenhangs zwischen einer steigenden Außentemperatur und sinkenden Wärmelast.

Diese erste Erkenntnis soll nun vertieft werden, indem versucht wird den Zusammenhang mit einer Funktion $\dot{Q} = f(T_A)$ zu beschreiben. Dieses Vorgehen entspricht einer Regressionsanalyse, wie in Abschnitt 2.2 beschrieben. Die abhängige Variable y entspricht dem Wärmestrom \dot{Q} , und die unabhängige Variable u entspricht der Außentemperatur T_A . Für eine visuelle Untersuchung werden Stundendaten von Abbildung 4.2 gegeneinander aufgetragen, was in Abbildung 4.3 zu sehen ist. Hierbei wurden die Verläufe in Abbildung 4.2 entsprechend Abschnitt 4.3 mittels gewichtetem gleitenden Mittelwert bestimmt um die essentiellen Dynamiken auswerten zu können. Somit entspricht die Verwendung von Stundendaten einer stündlichen Abtastung der essentiellen Dynamik der Messdaten.

Da die Verteilung der Messpunkte eine Beschreibung mit einer linearen Regression

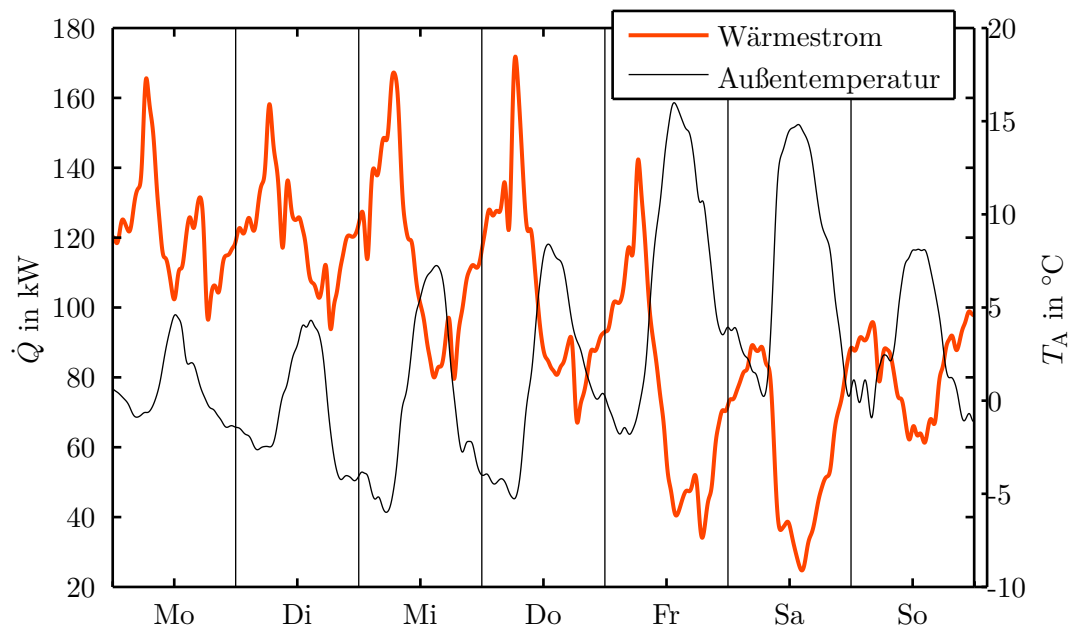


Abbildung 4.2.: VB2 – Wärmestrom und dazugehörige gemessene Außentemperatur einer Woche Ende Februar

nahelegt, wird ein Regressionsmodell entsprechend (2.17) angesetzt

$$\hat{Q}(T_A) = \alpha + \beta T_A, \quad (4.4)$$

das als Schätzung für den Zusammenhang $\dot{Q} = f(T_A)$ verwendet werden kann. Die Parameter dieses Regressionsmodells werden analog zur Vorgehensweise in Beispiel 2.1 für OLS und LAV getrennt bestimmt. Die beiden resultierenden Regressionsmodelle sind in Abbildung 4.3 eingezeichnet und geben den Zusammenhang sichtlich gut wieder.

Zusammengefasst kann der Zusammenhang $\dot{Q} = f(T_A)$ mit einer linearen Funktion angenähert werden. Hier sei angemerkt, dass nicht für alle Verbraucher und Wochen ein derart klarer linearer Zusammenhang wie in Abbildung 4.3 gegeben ist.

Einfluss der Globalstrahlung

Die **Globalstrahlung** I_G ist gemäß [46, S. 7] als Summe direkter, diffuser und reflektierter Strahlung definiert. Somit beinhaltet sie alle Sonneneinstrahlung, die als Energieeintrag auf die Erdoberfläche trifft, und ist immer positiv. Dieser Energieeintrag erwärmt Gebäude und über Glasflächen auch direkt die Innenräume, und reduziert somit die Heizlast. Für diese Untersuchung wird gezielt das Bürogebäude (VB2) mit seinen großen Glasflächen betrachtet, da hier ein besonders starker Zusammenhang zu erwarten ist.

Analog zur Außentemperatur wird zunächst der zeitlichen Verlaufes des Wärmestroms und der Globalstrahlung betrachtet. Dies ist in Abbildung 4.4 zu sehen, wobei die Skala

4. Messdaten

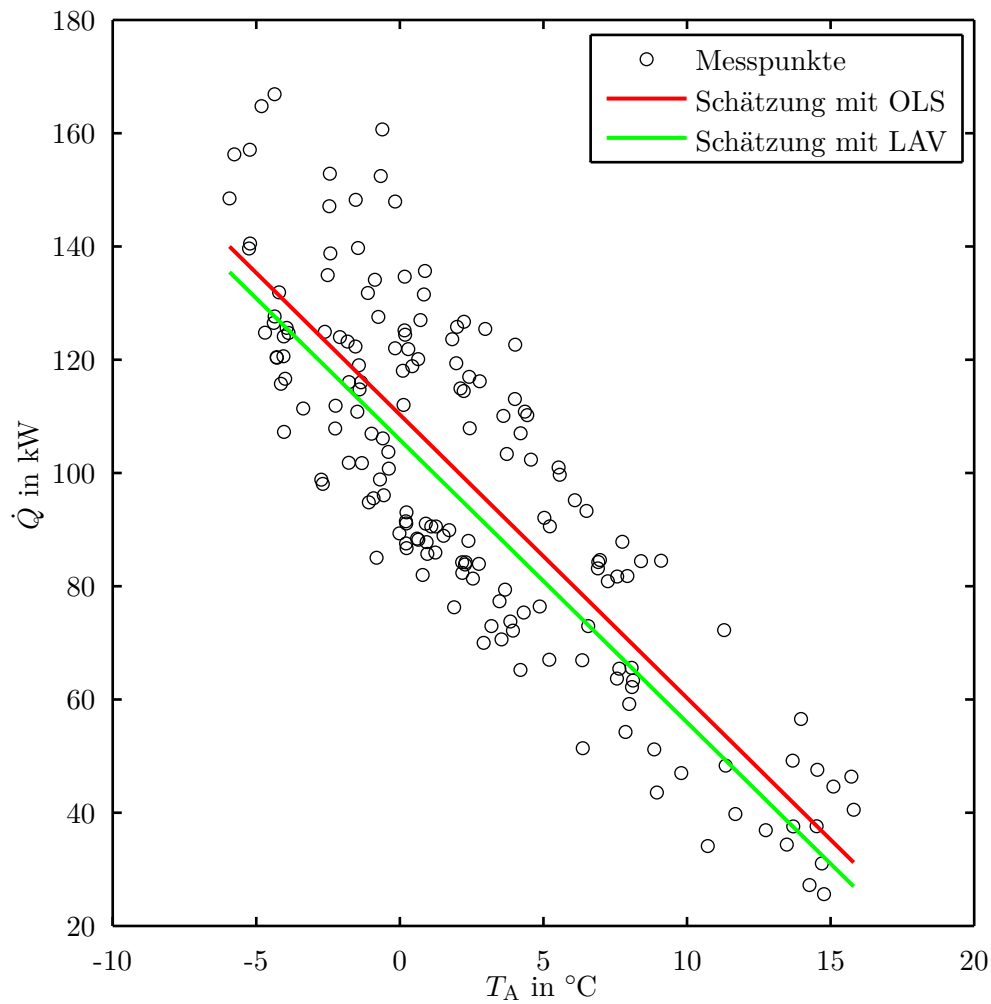


Abbildung 4.3.: VB2 – Wärmestrom in Abhängigkeit der Außentemperatur aufgetragen, inklusive Schätzungen mit linearer Regression. Die Messpunkte entsprechen punktuellen Stundendaten einer Woche.

für die Globalstrahlung auf der rechten Seite zu finden ist. Der zeitliche Verlauf der Globalstrahlung gibt gut wieder, wie die Sonne am Morgen aufgeht, zu Mittag ihren Höchststand erreicht und am Abend untergeht. In der Nacht ist die Globalstrahlung gleich Null. In dieser Abbildung ist kein statischer Zusammenhang zwischen Wärmestrom und Globalstrahlung zu erkennen.

Obwohl aus Abbildung 4.4 kein statischer Zusammenhang erkennbar ist, wird im nächsten Schritt der Wärmestrom auch über die Globalstrahlung aufgetragen, da dies Erkenntnisse zu Globalstrahlung und linearer Regression liefert. Für diese Regressions-

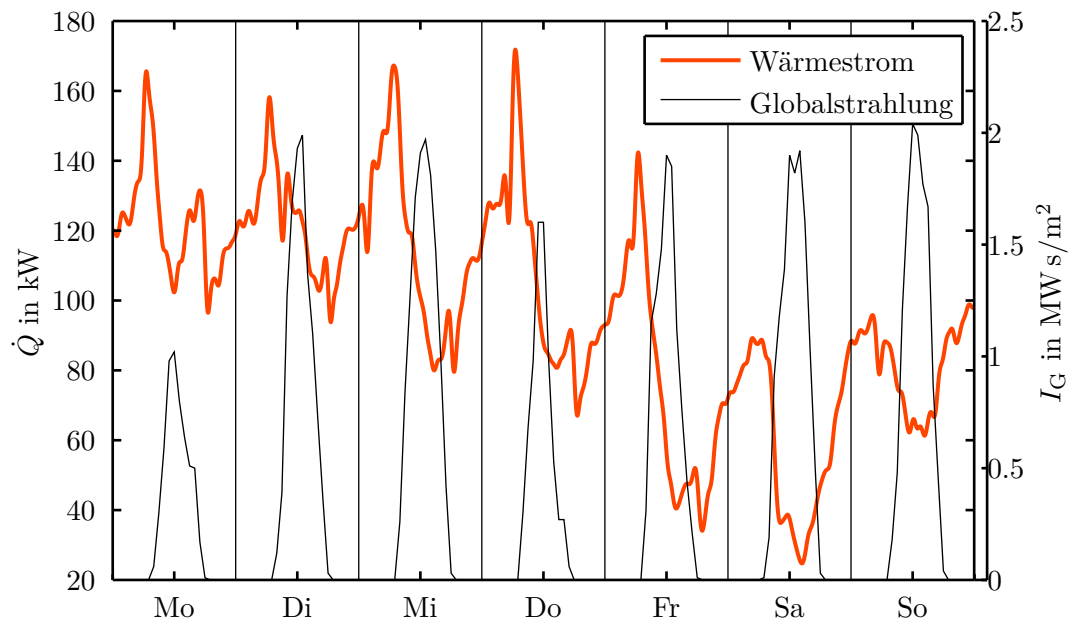


Abbildung 4.4.: VB2 – Wärmestrom und dazugehörige gemessene Globalstrahlung einer Woche Ende Februar

analyse entspricht die abhängige Variable y wieder dem Wärmestrom \dot{Q} , aber die unabhängige Variable u der Globalstrahlung I_G . Betrachtet man die Messpunkte in Abbildung 4.5 so ist erstens kein Zusammenhang zu erkennen und es liegen zweitens die Messpunkte der Nacht bei $I_G = 0$, was für eine Beschreibung des Zusammenhangs natürlich berücksichtigt werden müsste. Obwohl kein statischer, linearer Zusammenhang sinnvoll ist, kann mittels OLS oder LAV eine Gerade berechnet werden. In diesem Fall ist das Residuum, welches die Abweichung zwischen Messung y und Schätzung \hat{y} beschreibt, sehr groß, und das Regressionsmodell ist unbrauchbar.

Dies zeigt den begrenzten Einsatzbereich der linearen Regression, wobei eine zeitverschobene Globalstrahlung $I_G[n - m]$, $m = 1, 2, \dots$ als unabhängige Variable u eventuell besser für ein Regressionsmodell passen würde. Da die Globalstrahlung eine gewisse Zeit benötigt um Umgebung und Gebäude zu erwärmen, dürfte ein dynamisches Modell deutlich bessere Übereinstimmung liefern. Solch ein dynamisches Modell wird aber nicht weiter entwickelt, da die Außentemperatur eine vielversprechendere Eingangsgröße darstellt und der Einfluss der Globalstrahlung auf die Umgebung bereits teilweise in der Außentemperatur abgebildet ist.

Zusammengefasst wurde kein lineares Regressionsmodell gefunden, welches den Zusammenhang zwischen Wärmestrom und Globalstrahlung beschreibt.

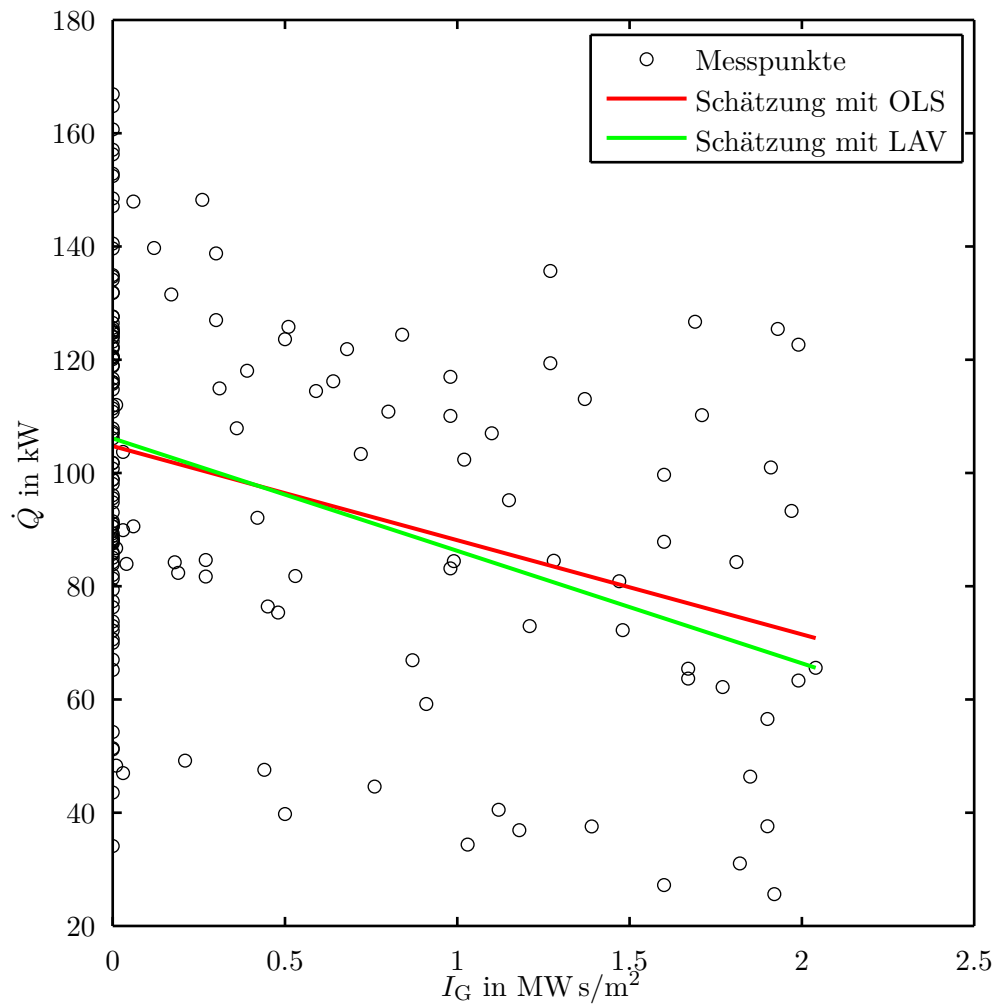


Abbildung 4.5.: VB2 – Wärmestrom in Abhängigkeit der Globalstrahlung aufgetragen, inklusive Schätzungen mit linearer Regression. Die Messpunkte entsprechen wieder punktuellen Stundendaten einer Woche.

4.4.2. Einfluss zeitlicher Faktoren

Zusätzlich zu Wetterfaktoren haben auch Wochentag und Uhrzeit einen noch zu untersuchenden Einfluss auf den Wärmebedarf. Der Einfluss von Wochentag und Uhrzeit ist auf des Verhalten der Personen in den Gebäuden, das Verhalten der Gebäudeautomation und auf das Tageslicht zurückzuführen. Zum Beispiel gehen manche Personen regelmäßig zu einer bestimmten Uhrzeit duschen, die Gebäudeautomation reduziert die Leistung der Heizung über Nacht (Nachtabenkung), und die Sonne geht zu einer bestimmten Uhrzeit

auf und beginnt Gebäude und Umgebung zu erwärmen.

Man könnte an dieser Stelle auch Monate und Jahreszeiten betrachten, worauf verzichtet wurde, da diese Einflüsse stark mit den jeweiligen Wetterbedingungen verknüpft ist. Zum Beispiel wird im Jänner mehr geheizt als im August, was aber auf die unterschiedlichen Außentemperaturen zurückzuführen ist. Würden also auch die Jahreszeiten berücksichtigt, so würde der Einfluss von Wetterfaktoren doppelt in die Prognosemethode eingehen, was nicht erwünscht ist.

Einfluss des Wochentages

Diese Arbeit beschäftigt sich mit unterschiedlichsten Verbrauchern, hinter denen aber immer reale Gebäude oder Wärmenetze, die ein Zusammenschluss mehrerer Gebäude sind, stehen. Da diese Gebäude von Personen genutzt werden, die einen wöchentlichen Rhythmus haben, unterliegen auch die Vorgänge in den Gebäuden wöchentlichen Rhythmen. Aus diesem Grund wird nun untersucht, welchen Einfluss die Wochentage (Montag bis Sonntag) auf den Wärmestrom haben.

Zuerst soll der bereits bekannte Wärmestrom in Abbildung 4.2 näher untersucht werden. Betrachtet man die Tage der Arbeitswoche (Montag bis Freitag) so fällt auf, dass die Verläufe sich ähnlich verhalten. So ist eine scharfe Spitze am Morgen und ein tiefes Tal am Nachmittag zu erkennen, wobei im Tal noch eine kleinere Spitze auftritt. Der Freitag hat zwar einen ähnlichen Verlauf wie Mittwoch und Donnerstag, ist aber nach unten verschoben, was auf die angestiegene Außentemperatur zurückzuführen ist. Das Wochenende (Samstag und Sonntag) hat einen grundlegend anderen Verlauf, was bei einem Bürogebäude (siehe Verbraucherklasse in Tabelle 4.1) nicht weiter verwundert. Die Arbeitswoche und das Wochenende haben also für diese Woche deutlich andere Verläufe, was eine separate Behandlung dieser beiden Zeiträume nahelegt.

Diese Beobachtung aus einer Woche wird nun auf das gesamte Jahr und alle Verbraucher erweitert. Zu diesem Zweck werden je Verbraucher alle Wochen parallel dargestellt und zu jedem Zeitpunkt (Uhrzeit) der Mittelwert bestimmt, was in den Abbildungen 4.6 bis 4.12 zu sehen ist. Jede Woche ist einer eigenen Farbe dargestellt, wobei „kalte“ Farben (Blau, Grün, etc.) für die kalten Jahreszeiten und „warme“ Farben (Gelb, Rot, etc.) für die warmen Jahreszeiten verwendet wurden. Anhand der Anzahl und Farbe der Linien kann grob abgeschätzt werden, für welche Wochen Daten vorhanden sind. Manche Kurvenverläufe enthalten perfekte Geraden, welche aus der linearen Interpolation defekter oder fehlender Daten resultieren. Man beachte, dass der Wärmestrom in Abbildung 4.11 und 4.12 in MW und in den restlichen Abbildungen in kW aufgetragen ist.

In Abbildung 4.7, 4.9, 4.11 und 4.12 ist ein Unterschied zwischen Arbeitswoche und Wochenende zu erkennen. Bei diesen vier Verbrauchern ist die Morgenspitze am Wochenende deutlich geringer als in der Arbeitswoche. Für die restlichen drei Verbraucher hat der Wochentag keine sichtbare Auswirkung, was auf die Regelmäßigkeit der den Mittelwert dominierenden Heizlast durch die Gebäudeautomation zurückzuführen ist.

Zusammengefasst ist eine getrennte Behandlung von Arbeitswoche und Wochenende sinnvoll, da diese zum Teil ein stark unterschiedliches Verhalten aufweisen. Ansonsten sticht kein Wochentag als markant anders für alle Verbraucher heraus.

4. Messdaten

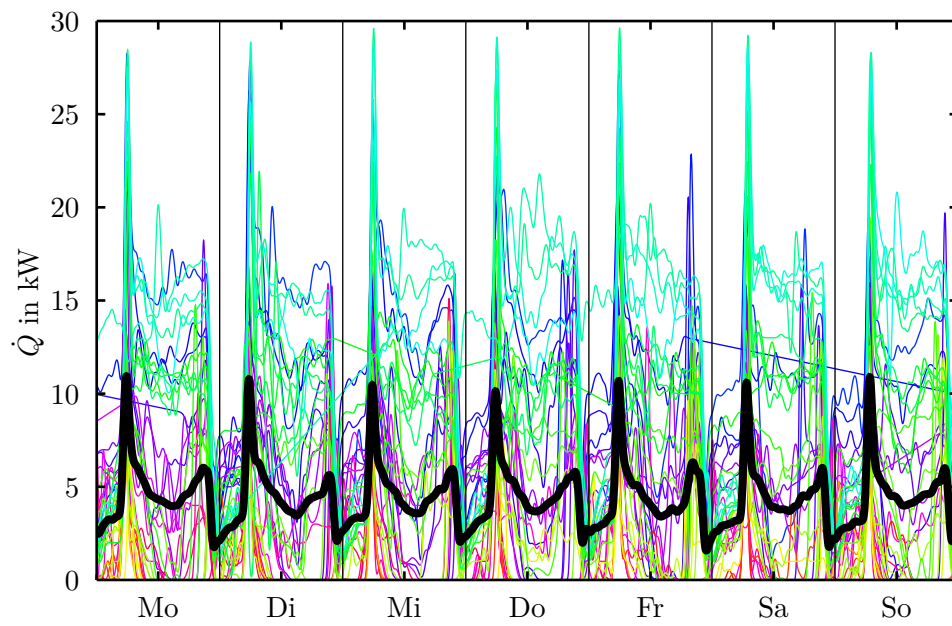


Abbildung 4.6.: VB1 – Alle Wochen eines Jahres parallel inkl. Mittelwert

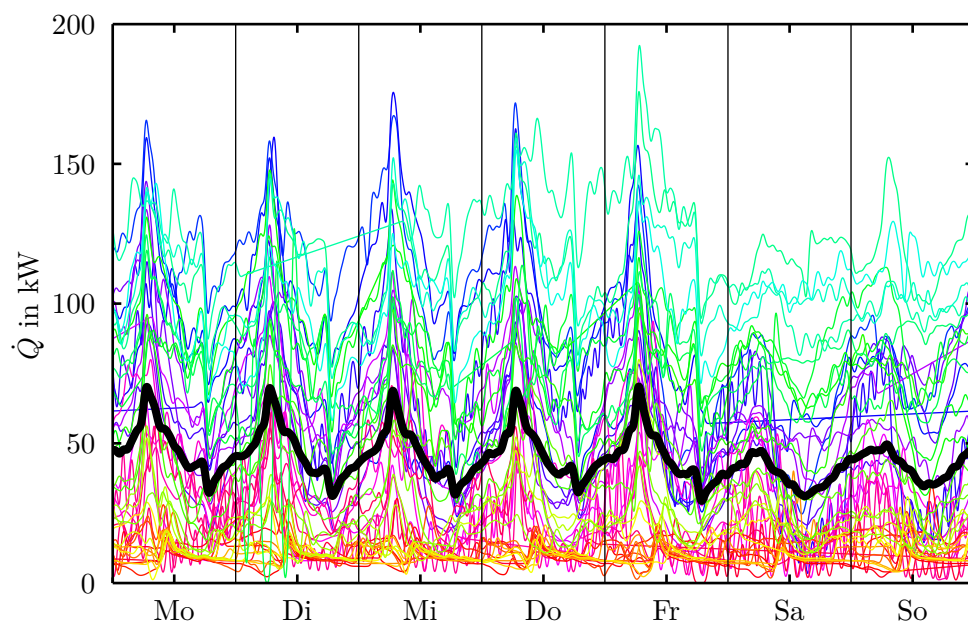


Abbildung 4.7.: VB2 – Alle Wochen eines Jahres parallel inkl. Mittelwert

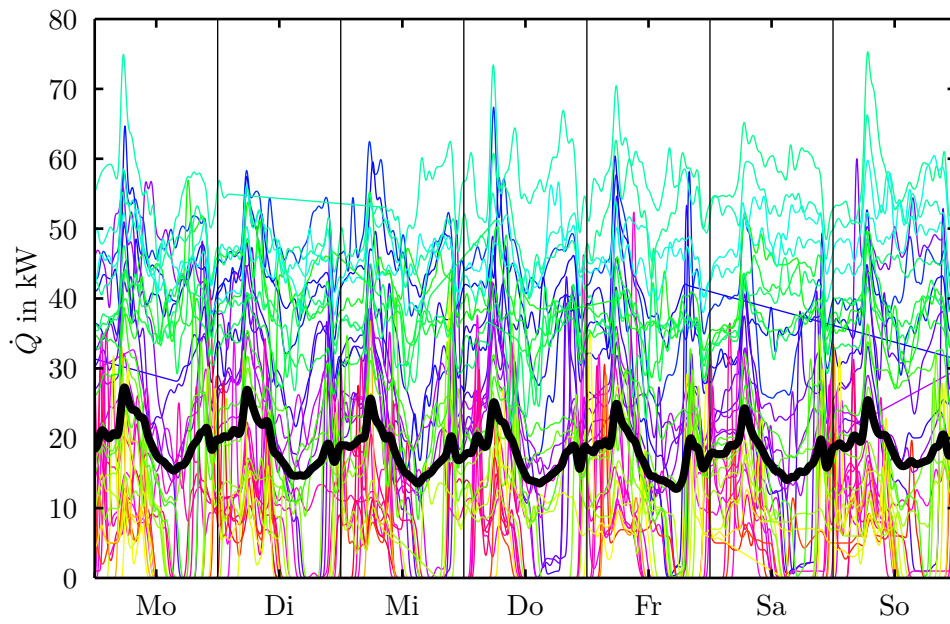


Abbildung 4.8.: VB3 – Alle Wochen eines Jahres parallel inkl. Mittelwert

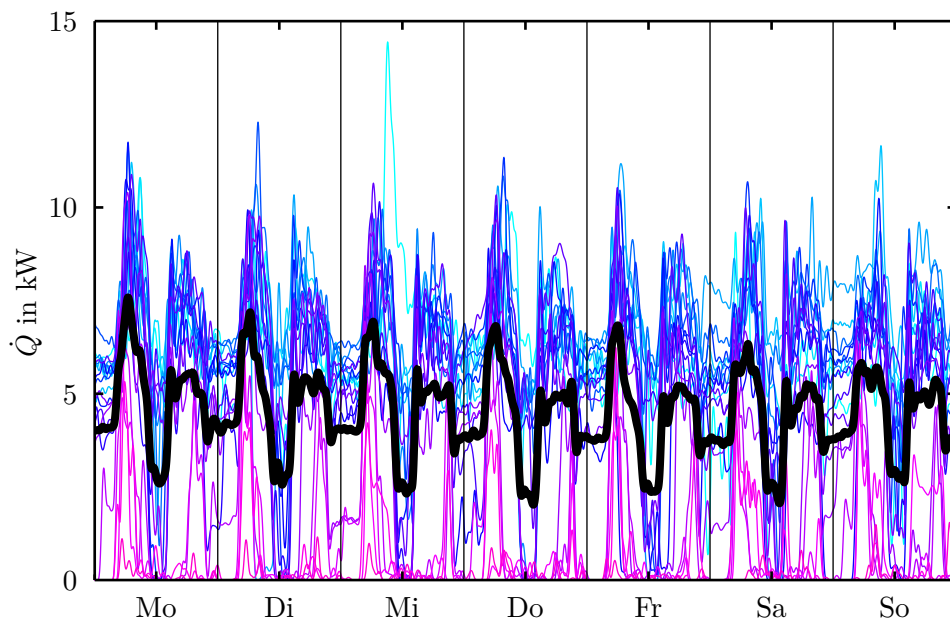


Abbildung 4.9.: VB4 – Alle Wochen eines Jahres parallel inkl. Mittelwert

4. Messdaten

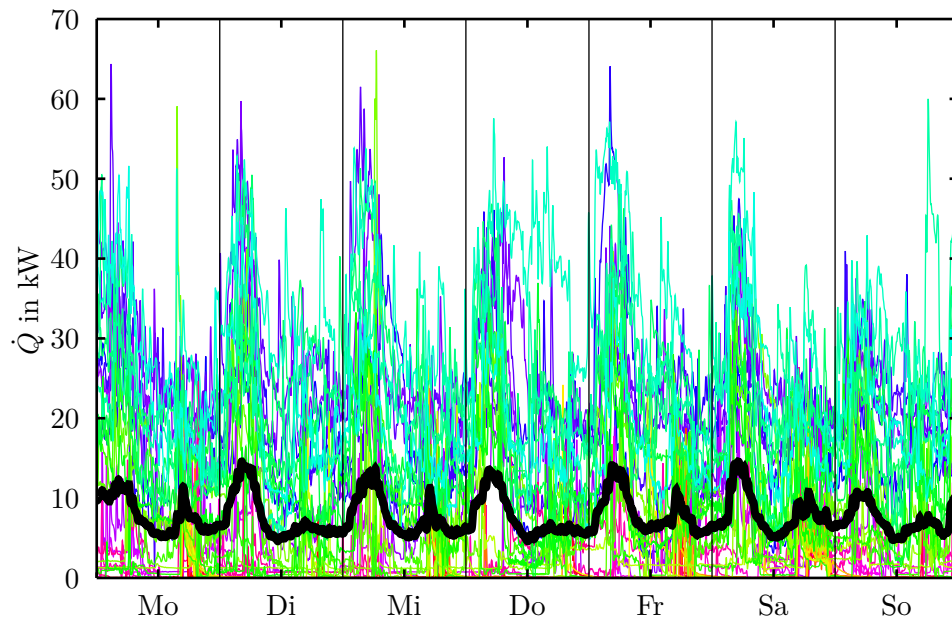


Abbildung 4.10.: VB5 – Alle Wochen eines Jahres parallel inkl. Mittelwert

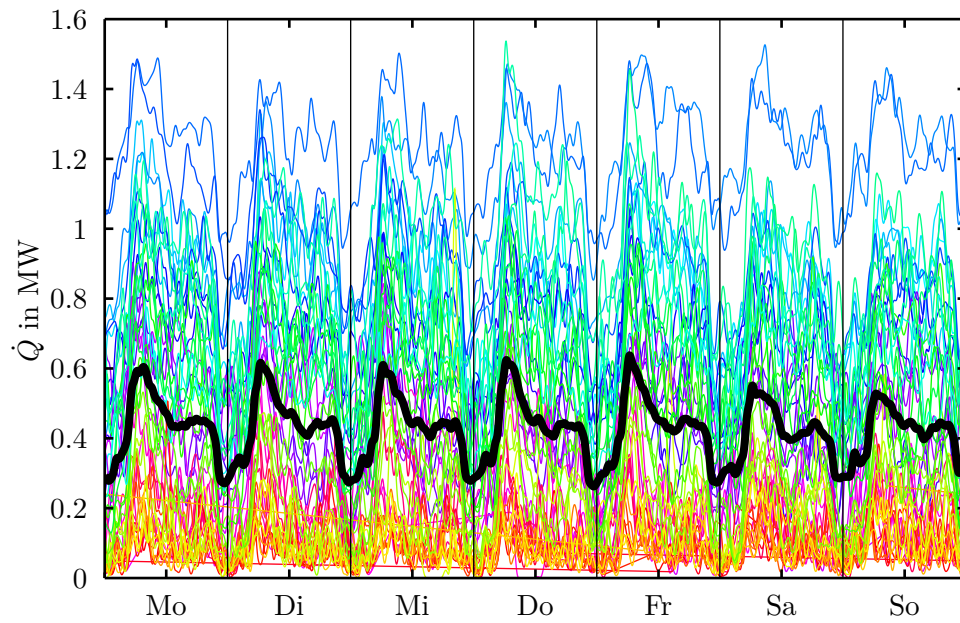


Abbildung 4.11.: VB6 – Alle Wochen eines Jahres parallel inkl. Mittelwert

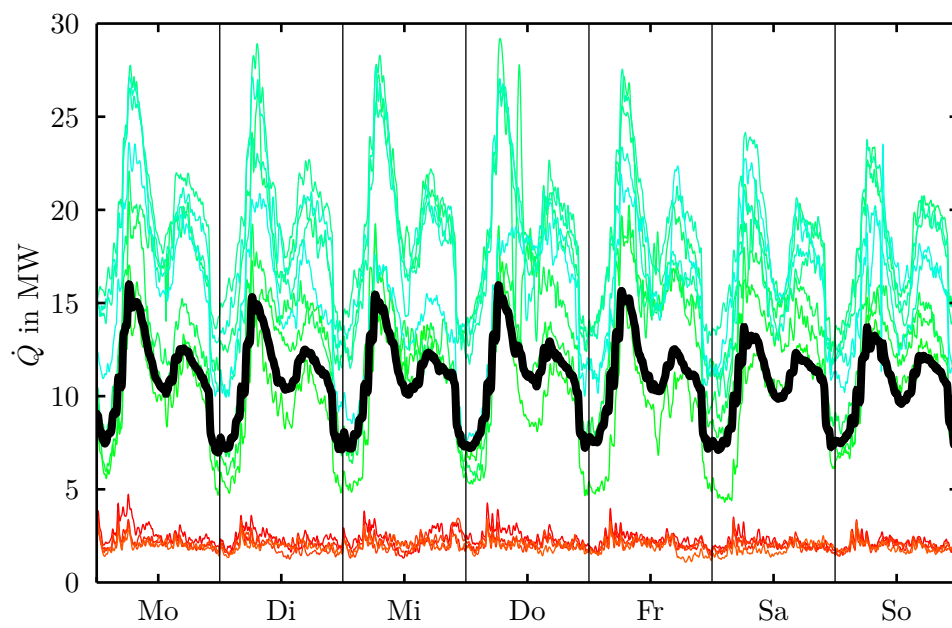


Abbildung 4.12.: VB7 – Alle Wochen eines Jahres parallel inkl. Mittelwert

Einfluss der Uhrzeit

Anhand zweier markanter Größen soll der Einfluss der Uhrzeit auf den Wärmebedarf untersucht werden. Betrachtet man die Lastverläufe in den bisherigen Abbildungen 4.2, 4.6, 4.7, 4.8, 4.9, 4.11 und 4.12 so fällt eine deutliche Lastspitze am Morgen auf. Vergleicht man in Abbildung 4.2 den Zeitpunkt der Lastspitze mit dem Temperaturtief des Tages, so scheinen diese nahezu zeitgleich zu sein. Anhand dieser beider Größen, der Lastspitze und dem Temperaturtief, werden nun Untersuchungen angestellt, und zusätzlich wird der Einfluss der Nacht untersucht. Für diese Untersuchung wird als Erstes ein Wohnhaus (VB1) und als Zweites ein Nahwärmenetz (VB6) betrachtet, um die Gemeinsamkeiten und Unterschiede herauszuarbeiten.

Von jedem Tag des Jahres wird das Maximum des Wärmestroms (Lastspitze) und das Minimum der Außentemperatur (Temperaturtief) bestimmt und die dazugehörige Uhrzeit ermittelt. In Abbildung 4.13 ist das Ergebnis dieser Berechnung für das ganze Jahr von VB1 zu sehen, wobei am Beginn des Jahres Daten fehlen. In dieser Abbildung ist die Nacht in Grau eingezeichnet, deren Beginn bzw. Ende mit der Zeitumstellung um eine Stunde versetzt ist.

Die Lastspitzen häufen sich entlang zweier Linien die ca. um 05:30 bzw. um 06:30 verlaufen. Die Linie bei 05:30 entspricht den Tagen der Arbeitswoche und die Linie bei 06:30 den Tagen des Wochenendes. Diese Lastspitzen am Morgen sind auch in den dazugehörigen Jahresdaten von Abbildung 4.6 deutlich zu erkennen. Hierbei handelt es sich um Effekte der Gebäudeautomation (Heizungsregelung) und nicht um regelmäßiges Benutzerverhalten. Daraus folgt die wichtige Erkenntnis, dass die betrachteten Messdaten von geregelten Systemen stammen und folglich eine Änderung der Gebäudeautomation andere Lastverläufe hervorrufen würde. Manche Lastspitzen sind verteilt über den Abend zu finden und dürften Benutzerverhalten (Duschen o.Ä.) repräsentieren. Im Sommer sind einige Lastspitzen bei 00:00 eingetragen. In diesen Fällen ist der Wärmebedarf für den ganzen Tag gleich Null, weshalb kein tägliches Minimum bestimmt werden kann.

Die Temperaturtiefs häufen sich ebenfalls am Morgen, scheinen jedoch keinen direkten Einfluss auf die Lastspitzen zu haben. Einige Temperaturtiefs finden sich um 23:59, was ein Resultat der täglichen Minimierung ist. Fällt die Außentemperatur über mehrere Tage ab, so tritt das tägliche Minimum am Ende der einzelnen Tage auf.

Der Beginn bzw. das Ende der Nacht scheinen keinen direkten Einfluss auf die Lastspitzen zu haben. Die Häufung der Lastspitzen am Morgen folgt eindeutig der Uhrzeit und nicht dem Zeitpunkt des Sonnenaufgangs.

Diese Untersuchung wird nun auf ein Nahwärmenetz angewandt, dessen Ergebnis in Abbildung 4.14 zu sehen ist. Auch hier ist eine Häufung der Lastspitzen am Morgen zu beobachten, eine eindeutige Uhrzeit kann aber nicht identifiziert werden. Dies ist durch die vielen verschiedenen Abnehmer zu erklären, womit die Morgenspitze abgeflacht wird, wie in den Jahresdaten in Abbildung 4.11 zu sehen ist. Daraus ist zu schließen, dass die Abnehmer dieses Nahwärmenetzes verschiedene Gebäudeautomatisierungen haben, womit sich die Lastspitze am Morgen zufällig in einem Band bewegt. Wieder ist kein

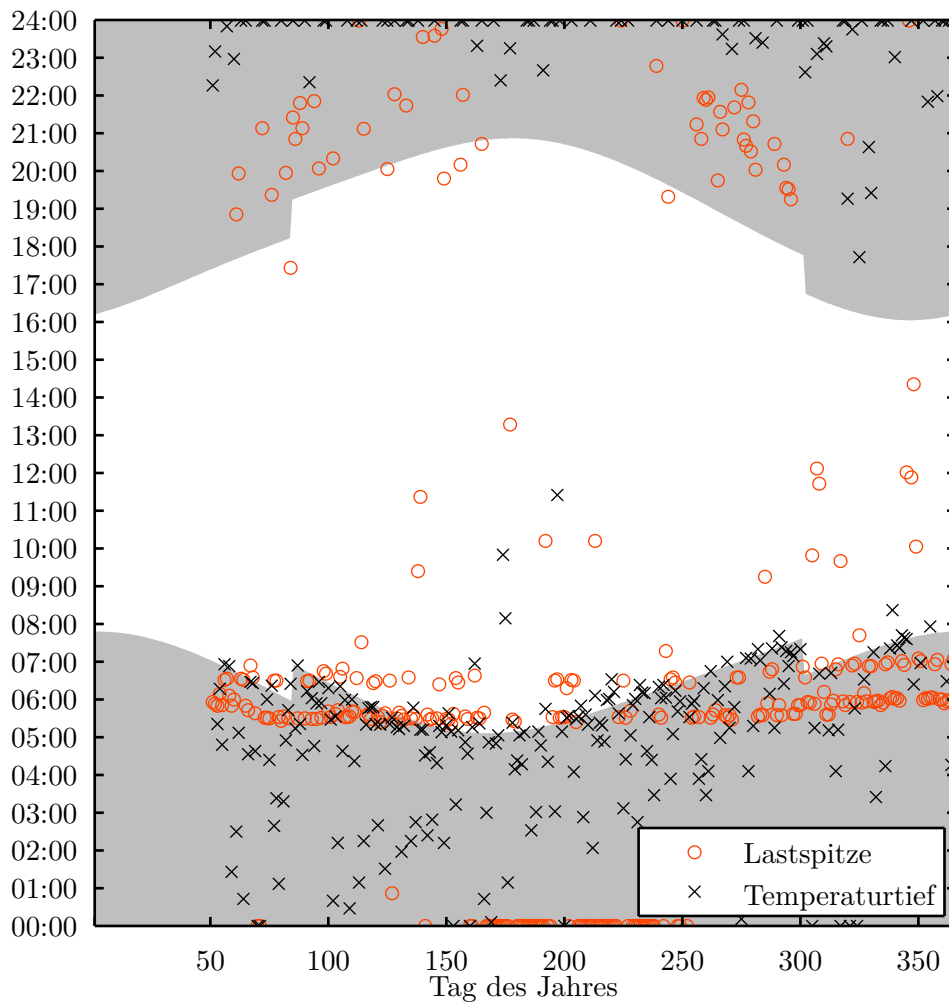


Abbildung 4.13.: VB1 – Uhrzeit der täglichen Lastspitze bzw. des täglichen Temperaturtiefs über das ganze Jahr. Die Nacht ist in Grau eingezeichnet.

direkter Zusammenhang zwischen Temperaturtief und Lastspitze erkennbar.

Zusammengefasst hat die Uhrzeit einen überraschend starken Einfluss auf den Wärmebedarf, was auf die Gebäudeautomation und regelmäßiges Benutzerverhalten zurückzuführen ist.

4.4.3. Einfluss des vergangenen Lastverlaufs

Der vergangene Lastverlauf beeinflusst zweifelsfrei den aktuellen Lastverlauf, was nun mit statistischen Methoden analysiert wird.

4. Messdaten

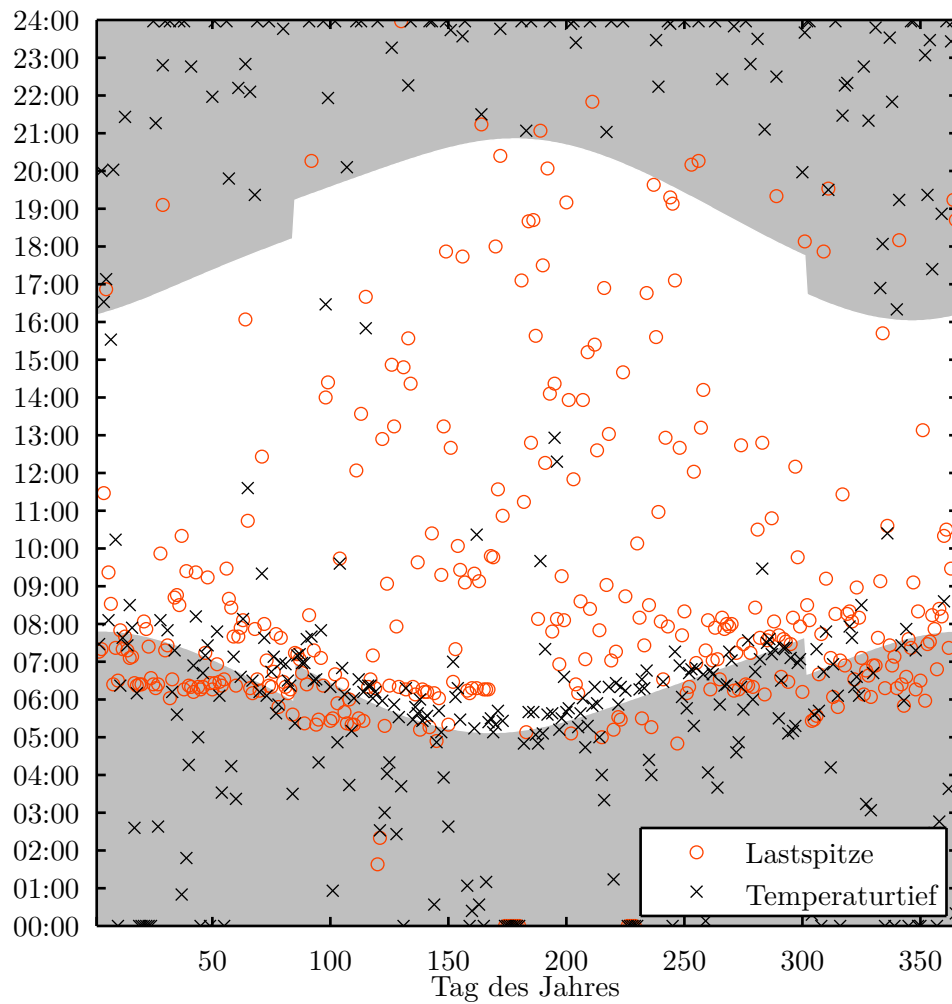


Abbildung 4.14.: VB6 – Uhrzeit der täglichen Lastspitze bzw. des täglichen Temperaturtiefs über das ganze Jahr. Die Nacht ist in Grau eingezeichnet.

Eine aussagekräftige Methode ist die Autokorrelationsfunktion (AKF) $r_{\dot{Q}\dot{Q}}[m]$ mit der zeitliche Regelmäßigkeiten innerhalb des Wärmestroms analysiert werden können. Für diese Untersuchung werden die Daten vom Bürogebäude (VB2) im November betrachtet, da hier starke Regelmäßigkeiten zu erwarten sind. Damit der Unterschied zwischen Arbeitswoche und Wochenende die Untersuchung nicht erschwert, werden nur die Arbeitswochen dieses Monats betrachtet. Für die nachfolgenden Untersuchungen wird der Monatsmittelwert abgezogen, um die Dynamiken stärker in der Vordergrund zu rücken.

Die AKF des vom Mittelwert befreiten Wärmestroms des beschriebenen Zeitraums ist in Abbildung 4.15 zu sehen. Hier ist die Zeitverschiebung m in Tagen aufgetragen,

wobei m alle 25 vorhandenen Tage (5 Arbeitswochen mit je 5 Tagen) durchläuft. Es sind nur Werte für $m > 0$ aufgetragen da, die AKF symmetrisch ist: $r_{xx}[m] = r_{xx}[-m]$. Der dominante, tägliche Zyklus des Wärmestroms, der dem Verlauf der AKF überlagert ist, ist deutlich erkennbar. Dies bedeutet, dass im Mittel der Wärmestrom $\dot{Q}[n]$ eine ähnliche Form wie der um ganzzahlige Vielfache von 24 h verschobene Wärmestrom $\dot{Q}[n + k \cdot 24 \text{ h}]$, $k = \pm 1, \pm 2, \dots$ aufweist.

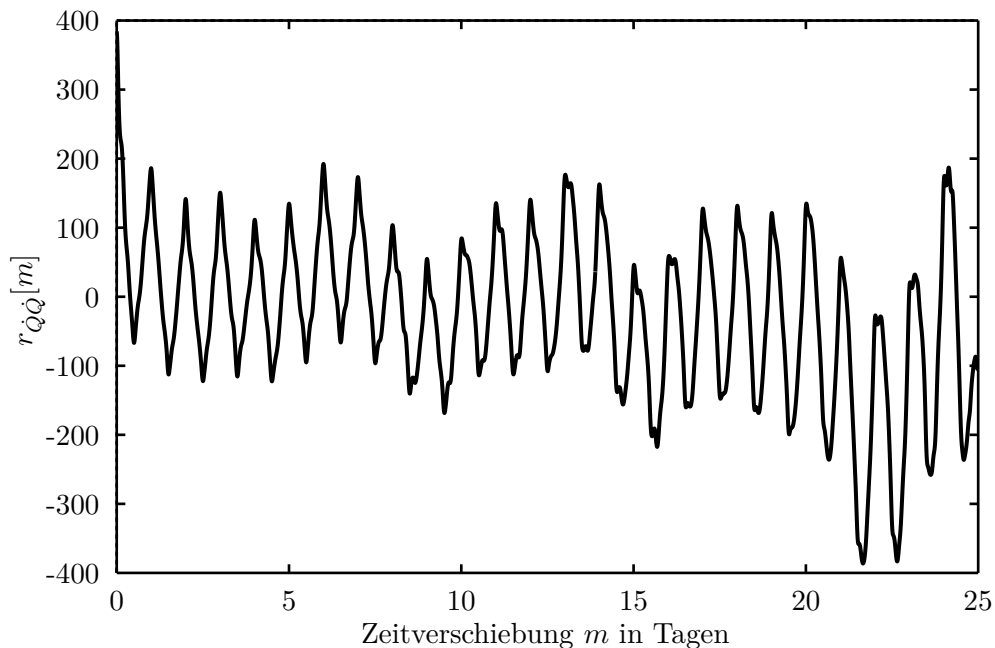


Abbildung 4.15.: VB2 – Autokorrelationsfunktion der mittelwertbefreiten Messdaten der Arbeitswoche von November

Diese Untersuchungen können mit Hilfe des Periodogramms $I(f)$ noch vertieft werden. Hierbei wird die AKF im Frequenzbereich betrachtet, wodurch die Frequenzen der Regelmäßigkeiten deutlicher zu erkennen sind.

Das Periodogramm des Wärmebedarfs ist in Abbildung 4.16 zu sehen. Da für diese Betrachtung der relative Verlauf interessant ist, wurde die y-Achse auf den Bereich $[0, 1]$ normiert. Das Periodogramm zeigt bei welchen Frequenzen dominante Regelmäßigkeiten in der AKF auftreten. In diesem Fall stechen drei Frequenzen ins Auge, die in Tabelle 4.2 aufgezählt sind und deren Periodendauer $1/f$ in Stunden und Tagen angegeben ist. Anhand der Periodendauer ist der schwache 5-tägige Rhythmus durch die 5-tägige Arbeitswoche gut erkennbar. Die bereits in der AKF erkannte tägliche Regelmäßigkeit ist hier genauso zu sehen, wie ein schwächerer 8-stündiger Rhythmus für den keine eindeutige Erklärung gefunden wurde.

Zusammengefasst bestätigen die Auswertungen mit AKF und Periodogramm den bereits bei den Lastspitzen beobachteten starken 24 h-Rhythmus des Wärmestroms. Die

4. Messdaten

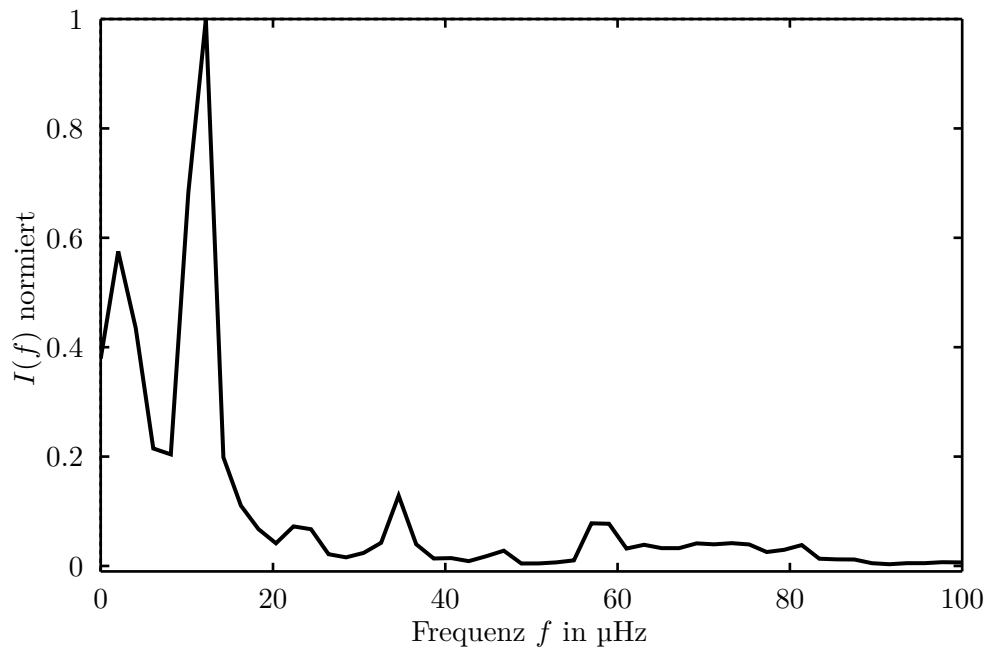


Abbildung 4.16.: VB6 – Periodogramm der mittelwertbefreiten Messdaten der Arbeitswochen von November

Frequenz in μHz	Periodendauer in Stunden	Periodendauer in Tagen
2,04	136,50	5,69
12,12	22,75	0,95
34,59	8,03	0,33

Tabelle 4.2.: Markante Frequenzen in Abbildung 4.16 und deren Periodendauer

anderen und neue Regelmäßigkeiten treten bei der Betrachtung anderer Monate bzw. anderer Verbraucher in verstärkter oder abgeschwächter Form auf. Diese sind aber nicht so stark und dominant wie der 24 h-Rhythmus, weshalb nur dieser 24 h-Rhythmus weiter verfolgt wird.

4.4.4. Zusammenfassung der Erkenntnisse aus der Messdatenauswertung

Anhand der durchgeführten Messdatenauswertung wurden mögliche Faktoren, welche den Verlauf des Wärmebedarfs beeinflussen, analysiert. Zum einen haben sich Vermutungen, wie der starke Einfluss der Außentemperatur oder die Unterschiede in Arbeitswoche und Wochenende, bestätigt. Zum anderen wurde der überraschend starke Einfluss der Uhrzeit durch den Zeitpunkt der Lastspitzen angedeutet und durch das Periodogramm nachgewiesen.

Einige Eigenschaften des Wärmebedarfs können mit verschiedenen Einflussfaktoren abgebildet werden. So kann die Heizlast durch die Außentemperatur und der Einfluss der Sonneneinstrahlung auf den Wärmebedarf durch die Globalstrahlung berücksichtigt werden. Vereinfacht können diese Eigenschaften des Wärmebedarfs aber auch mittels Jahreszeiten und Uhrzeit beschrieben und somit die Wetterfaktoren durch zeitliche Faktoren ersetzt werden.

Zusammengefasst liefert die Messdatenauswertung folgende wichtige Ergebnisse:

1. Der Zusammenhang zwischen Wärmestrom und Außentemperatur kann durch eine Gerade beschrieben werden.
2. Der Wärmestrom hat in der Arbeitswoche ein anderes Verhalten als am Wochenende.
3. Der Wärmestrom unterliegt einem 24 h-Rhythmus und ist somit stark von der Uhrzeit abhängig.

Diese Erkenntnisse bilden das Fundament für die Prognosemethode im folgenden Kapitel.

5. Prognosemethode

In diesem Kapitel wird das Ziel dieser Arbeit – eine Prognosemethode – schrittweise erarbeitet. Ausgehend von der Literaturrecherche in Kapitel 3 und der Messdatenauswertung in Abschnitt 4.4 werden zunächst grundlegende Überlegungen zur Prognosemethode angestellt. Daraus folgt der erste Prädiktor, der anhand der sieben Verbraucher evaluiert wird. Danach wird dieser Prädiktor erweitert und diese Erweiterungen wieder für alle Verbraucher bewertet. Da diese Erweiterungen keine Verbesserung bringen, ist der erste Prädiktor die finale Prädiktor-Komponente. Um das Prognoseergebnis zu verbessern, wird der Prädiktor mit einem Korrektor kombiniert. Hierbei wird wieder ein erster Korrektor entwickelt und dieser anschließend erweitert. Die daraus resultierende Prädiktor-Korrektor-Prognosemethode ist für alle Verbraucher getestet und enthält keine Parameter die von Hand gewählt werden müssen.

5.1. Grundlegende Überlegungen

Aus den drei wesentlichen Erkenntnissen der Messdatenauswertung in Abschnitt 4.4 werden nun Überlegungen für die Wahl der Prognosemethode abgeleitet. Hierbei wird auch die Literaturrecherche aus Kapitel 3 erweitert, um die übliche Berücksichtigung der jeweiligen Erkenntnis in der Literatur zu beschreiben. Zusätzlich zu diesen drei Erkenntnissen wird beachtet, dass die finale Prognosemethode auf einer Recheneinheit mit begrenzten Ressourcen implementierbar sein muss, weshalb ressourcensparende und leicht implementierbare Methoden bevorzugt werden.

Der starke Einfluss der Außentemperatur muss sich in der Prognosemethode wiederfinden. Da dieser mit einer linearen Funktion ausreichend genau beschrieben werden kann, ist eine aufwändigere Funktion unnötig. Ein linearer Ansatz erleichtert erstens die Berechnung und ist zweitens robust gegen Wetterprognose-Unsicherheiten, da ein Prognosefehler unskaliert eingeht.

Diese Überlegungen sprechen für die Verwendung einer linearen Regression und gegen Neuronale Netze und andere Black-Box Methoden. Darüber hinaus wird bei allen betrachteten wissenschaftlichen Artikeln, welche sich der linearen Regression bedienen, die Außentemperatur als Eingangsgröße verwendet. So wird der Zusammenhang zwischen Wärmestrom und Außentemperatur in [21] und [47] mit einer stückweise-linearen Funktion beschrieben, um mit einem einzigen Modell einen großen Temperaturbereich von -25 °C bis 25 °C abdecken zu können. In [17] geht die Außentemperatur linear, quadratisch, als Mittelwerte und als Maximalwerte ein. In [18] geht die aktuelle Außentemperatur sogar bis zur dritten Potenz und zusätzlich als vergangene Werte linear ein.

5. Prognosemethode

In [48] wird unter Anderem die Außentemperatur bis zu zwei Tage zurückliegend als Eingangsgröße verwendet.

In dieser Arbeit wird der erste Prädiktor als lineare Funktion der aktuellen Außentemperatur angesetzt. Eine stückweise-lineare Funktion ist unnötig, da durch die regelmäßige Aktualisierung der Parameter nur sehr beschränkte Temperaturbereiche in den Trainingsdaten auftreten. Die Berücksichtigung höherer Potenzen ist gemäß Messdatenauswertung unnötig und ist gemäß der Fehlerfortpflanzung durch die unsichere Wetterprognose problematisch.

Da der Wärmestrom während einer Arbeitswoche in vielen Fällen einen deutlich anderen Verlauf hat als an einem Wochenende, ist eine getrennte Behandlung dieser Zeiträume sinnvoll. Grundsätzlich könnte für diese Zeiträume jeweils eine andere, oder die gleiche Methode mit anderen Konfigurationen verwendet werden. Man könnte auch eine gesamte Methode verwenden, bei der mit binären Variablen, je nach Wochentag, bestimmte Komponenten zu- oder weggeschaltet werden. Dies macht jedoch, im endgültigen Verhalten keinen Unterschied zu einzelnen, übersichtlicheren Modellen.

Die Aufteilung in Arbeitswoche und Wochenende wird auch in [15] für die Wärmelast und in [18, Model A] sowie [49, 17, 48] für die elektrische Last durchgeführt. Eine Aufteilung in einzelne Tage ist unüblich und wird nur bei der elektrischen Last vereinzelt eingesetzt. So unterteilt [19] die Woche fünf Mal und behandelt Sonntage je nach Monat anders. In [26] werden Samstag und Sonntag getrennt behandelt und in [23] wird die Woche überhaupt anders unterteilt, da es sich hierbei um Daten eines islamischen Landes handelt. In allen in diesem Absatz beschriebenen Artikeln wird jeweils die gleiche Methode mit anderen Konfigurationen auf die spezifizierten Zeiträume angewandt.

In dieser Arbeit wird nur zwischen Arbeitswoche und Wochenende unterschieden, da in der Messdatenauswertung keine weitere Unterscheidung gefunden wurde, die auf alle Verbraucher zutrifft. Somit werden getrennte Parametersätze für die Arbeitswoche und das Wochenende bestimmt, wofür auch nur Trainingsdaten aus dem jeweiligen Zeitabschnitt verwendet werden dürfen. Sollen zum Beispiel, um die Flexibilität der Methode zu erhalten, maximal Trainingsdaten der vergangenen drei Wochen berücksichtigt werden, so stehen für die Arbeitswoche $3 \cdot 5 \text{ d} = 15 \text{ d}$ und für das Wochenende $3 \cdot 2 \text{ d} = 6 \text{ d}$ zur Verfügung. Eine feinere Unterteilung der Woche hat somit auch den Nachteil, dass für jeden dieser kürzeren Zeitabschnitte weniger Trainingsdaten zur Verfügung stehen, sofern die Flexibilität der Methode erhalten bleiben soll.

Der starke Einfluss der Uhrzeit sollte sich ebenfalls in der Methode wiederfinden. Die Uhrzeit ist lediglich eine feinere Unterteilung der Woche als die Trennung in Arbeitswoche und Wochenende. Daher kann analog je Uhrzeit eine andere Methode oder die gleiche Methode mit anderer Konfiguration verwendet werden.

Um den Einfluss der Uhrzeit besser abbilden zu können, setzen [18, Model A] und [21] auf die gleiche Methode mit anderer Konfiguration für jede Stunde des Tages. In [17] werden sehr viele mögliche Eingangsparameter für jede Stunde angesetzt und nach durchgeführter Parameterbestimmung werden Parameter mit geringem Einfluss verworfen. Dadurch ergibt sich für jede Stunde ein eigenes Modell, das jedoch ursprünglich

vom gleichen Ansatz aus entwickelt wurde. Diese populäre Methode wurde von [31] aufgegriffen und weiterentwickelt.

In dieser Arbeit wird für jede Stunde die gleiche Methode (lineare Regression) mit eigener Konfiguration (gleiches Modell mit eigenem Parametersatz) verwendet. In der Erweiterung des ersten Prädiktors in Abschnitt 5.3 wird zusätzlich die Globalstrahlung berücksichtigt. Da diese in der Nacht gleich Null ist, wird dort dann ein eigenes Modell für den Tag bzw. für die Nacht eingesetzt.

Man könnte alle drei Erkenntnisse der Messdatenauswertung auch mittels ARMA-Prozess modellieren, die auch häufig in der Literatur zur Anwendung kommen. Hierfür müsste der 24 h-Rhythmus (S) berücksichtigt werden, eine Elimination des Trends erfolgen (I) und die Außentemperatur als externer Eingang (X) berücksichtigt werden, womit ein SARIMAX-Modell benötigt würde. Zusätzlich wären eigene Parametersätze für Arbeitswoche und Wochenende notwendig. Dieses SARIMAX-Modell wäre von der Parameterbestimmung und Implementierung aber sehr aufwendig, weshalb die lineare Regression verwendet wird.

Zusammengefasst soll die lineare Regression für den ersten Prädiktor verwendet werden, wobei die Außentemperatur als einzige Eingangsgröße linear eingehen soll. Der Einfluss von Uhrzeit und Arbeitswoche bzw. Wochenende soll durch verschiedene Parametersätze für das Regressionsmodell berücksichtigt werden.

5.2. Erster Prädiktor

Der erste Prädiktor wird mittels linearer Regression realisiert, wobei der Einfluss von Arbeitswoche bzw. Wochenende, und der Einfluss der Uhrzeit durch jeweils eigene Parametersätze berücksichtigt wird. Da die Erweiterungen in Abschnitt 5.3 keine Verbesserung bringen, ist dieser erste Prädiktor die finale Prädiktor-Komponente.

Zuerst wird der Ansatz entsprechend eines linearen Regressionsmodells formuliert. Dann wird auf die Bestimmung der unbekannt Parameter des Regressionsmodells eingegangen. Anschließend wird die Anzahl der Trainingstage diskutiert und für alle Verbraucher passend bestimmt. Somit ist der erste Prädiktor vollständig beschrieben und kann abschließend für alle Verbraucher evaluiert werden.

5.2.1. Ansatz

Im Zuge der Messdatenauswertung ist der Zusammenhang $\dot{Q} = f(T_A)$ bereits mit einem linearen Regressionsmodell (4.4) beschrieben worden. Nun wird für Arbeitswoche und Wochenende ein eigenes Modell gleicher Form, aber mit anderen Zahlenwerte für die Parameter, sprich verschiedenen Parametersätze, verwendet. Somit sind die Unterschiede von Arbeitswoche und Wochenende berücksichtigt und es fehlt noch der Einfluss der Uhrzeit. Hierfür bleibt wieder der Ansatz unverändert und es werden eigene Parametersätze je Stunde n verwendet. Das Regressionsmodell hat also für jede Stunde,

5. Prognosemethode

sowohl für Arbeitswoche als auch für das Wochenende, die gleiche Form, aber $2 \cdot 24 = 48$ verschiedene Parametersätze.

Der Ansatz für den prädizierten Wärmestrom $\hat{Q}_P[n]$ für einen beliebigen Tag lautet

$$\hat{Q}_P[n] = \alpha[n] + \beta[n]\hat{T}_A[n], \quad n = 0, 1, \dots, 23, \quad (5.1)$$

mit

n ... diskrete Zeitvariable in Stunden (0 entspricht 00:00 und 23 entspricht 23:00),

\hat{Q}_P ... prädizierter Wärmestrom in W,

\hat{T}_A ... prognostizierte Außentemperatur in K,

α, β ... Parameter.

Somit wird die Last mit einem statischen, linearen Zusammenhang für jede Stunde des Tages durch einen Parametersatz α und β beschrieben.

Hierbei ist zu beachten, dass der Ansatz (5.1) die *prognostizierte* Außentemperatur benötigt. Somit ist die Qualität der Lastprognose von der Qualität der Wetterprognose abhängig.

5.2.2. Parameterbestimmung

In diesem Abschnitt wird zunächst die Bestimmung der unbekannt Parameter des Regressionsmodells beschrieben. Da diese Bestimmung aus den vorangegangenen Messdaten erfolgt, wird danach untersucht wie viele Messdaten verwendet werden sollen.

Regressionsparameter

Nachdem der Ansatz feststeht, sind nun die unbekannt 48 Parametersätze zu bestimmen. Diese Parametersätze können als Spaltenvektoren \mathbf{x}_i einer Parametermatrix \mathbf{X} angeschrieben werden

$$\begin{aligned} \mathbf{X} &= \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_{24} & \mathbf{x}_{25} & \mathbf{x}_{26} & \dots & \mathbf{x}_{48} \end{bmatrix} \\ &= \begin{bmatrix} \alpha_{\text{MoFr}}[0] & \alpha_{\text{MoFr}}[1] & \dots & \alpha_{\text{MoFr}}[23] & \alpha_{\text{SaSo}}[0] & \alpha_{\text{SaSo}}[1] & \dots & \alpha_{\text{SaSo}}[23] \\ \beta_{\text{MoFr}}[0] & \beta_{\text{MoFr}}[1] & \dots & \beta_{\text{MoFr}}[23] & \beta_{\text{SaSo}}[0] & \beta_{\text{SaSo}}[1] & \dots & \beta_{\text{SaSo}}[23] \end{bmatrix}. \end{aligned} \quad (5.2)$$

Die Parametersätze \mathbf{x}_1 bis \mathbf{x}_{24} werden für die Arbeitswoche benötigt, was mit dem Index MoFr gekennzeichnet ist. Analog werden die Parametersätze \mathbf{x}_{25} bis \mathbf{x}_{48} für das Wochenende (mit Index SaSo) benötigt. Somit sind in (5.2) die 48 Parametersätze mit insgesamt $2 \cdot 48 = 96$ Unbekannten übersichtlich dargestellt. Mit $\alpha_{\text{MoFr}}[1]$ wird zum Beispiel der Offset-Parameter für alle Tage der Arbeitswoche (Montag bis Freitag) zur Uhrzeit 01:00 bezeichnet.

Die Parameterbestimmung wird nun anhand der Arbeitswoche erklärt; beim Wochenende ist analog zu verfahren. Nachdem jetzt nur die Arbeitswoche betrachtet wird, werden auch nur Messdaten dieses Zeitraums betrachtet und mit dem Index MoFr gekennzeichnet. Falls mehr als fünf Arbeitstage betrachtet werden, so werden die Daten der Wochenenden entfernt und es wird Arbeitswoche an Arbeitswoche gereiht.

Da der Ansatz (5.1) einem Regressionsmodell entspricht, werden die Parameter entsprechend der linearen Regression in Abschnitt 2.2 bestimmt. Analog zur Messdatenauswertung in Abschnitt 4.4 entspricht die abhängige Variable y dem Wärmestrom \dot{Q}_{MoFr} , und die unabhängige Variable u entspricht der Außentemperatur $T_{\text{A,MoFr}}$. Somit kann entsprechend (2.14) ein Gleichungssystem mit $N_{\text{TT,MoFr}}$ berücksichtigten Trainingstagen für jede Stunde n des Tages aufgestellt werden:

$$\mathbf{y}[n] = \mathbf{A}[n]\mathbf{x}[n] + \mathbf{e}[n], \quad n = 0, 1, \dots, 23, \quad (5.3)$$

mit

$$\mathbf{A}[n] = \begin{bmatrix} \mathbf{1} & \mathbf{u}[n] \end{bmatrix} = \begin{bmatrix} 1 & T_{\text{A,MoFr}}[n - N_{\text{TT,MoFr}} \cdot 24] \\ 1 & T_{\text{A,MoFr}}[n - (N_{\text{TT,MoFr}} - 1) \cdot 24] \\ \vdots & \vdots \\ 1 & T_{\text{A,MoFr}}[n - 48] \\ 1 & T_{\text{A,MoFr}}[n - 24] \end{bmatrix} \quad (5.4)$$

und

$$\mathbf{y}[n] = \begin{bmatrix} \dot{Q}_{\text{MoFr}}[n - N_{\text{TT,MoFr}} \cdot 24] \\ \dot{Q}_{\text{MoFr}}[n - (N_{\text{TT,MoFr}} - 1) \cdot 24] \\ \vdots \\ \dot{Q}_{\text{MoFr}}[n - 48] \\ \dot{Q}_{\text{MoFr}}[n - 24] \end{bmatrix}. \quad (5.5)$$

Zur Bestimmung des optimalen Parametervektors \mathbf{x}^* für die Stunde n werden also Messdaten der vergangenen $N_{\text{TT,MoFr}}$ Trainingstage zu dieser Stunde n verwendet. Ist zum Beispiel $n = 6$ und $N_{\text{TT,MoFr}} = 11$, so werden für die Bestimmung von $\alpha_{\text{MoFr}}[6]$ und $\beta_{\text{MoFr}}[6]$ Messdaten der vergangenen 11 Arbeitstage um 06:00 verwendet. Die Anzahl der Trainingstage N_{TT} wird für die Arbeitswoche bzw. das Wochenende getrennt im nächsten Unterabschnitt bestimmt.

In der Praxis werden die Parameter laufend aus den vorangegangenen Messdaten aktuell bestimmt, womit die Methode adaptiv ist und automatisch die Änderungen im Laufe einer Saison berücksichtigt. Diese Herangehensweise stellt sicher, dass sich die Prognosemethode an ein geändertes Benutzerverhalten, an Änderungen der Gebäudesubstanz und an neue Wetterbedingungen stetig anpasst.

Der optimale Parametervektor wird durch Minimierung des Fehlers (Residuums) \mathbf{e} entsprechend der robusten LAV-Kostenfunktion von (2.16) bestimmt. Wie in Abschnitt 2.4 beschrieben ist es sinnvoll die anschließende Bewertung der Prognosequalität mittels des mittleren, absoluten Fehlers MAE durchzuführen. Es ist zu beachten, dass es sich hier um zwei verschiedene „Kostenfunktionen“ handelt. Mit der LAV-Kostenfunktion

5. Prognosemethode

wird die Summe der absoluten Residuen der Trainingsdaten $\sum_i |y_i - \hat{y}_i|$ minimiert. Dies drückt aus, wie gut das lineare Regressionsmodell an die Trainingsdaten angepasst werden konnte. Mit dem MAE-Qualitätsmaß $1/N \sum_{i=1}^N |\dot{Q}_i - \hat{Q}_{P,i}|$ wird bewertet, wie gut die daraus resultierende Lastprognose \hat{Q}_P mit dem tatsächlichen Lastverlauf \dot{Q} übereinstimmt.

Die vorgestellte Regressionsmethode kann graphisch interpretiert werden. Analog zu Abbildung 4.3 kann der lineare Zusammenhang zwischen Wärmestrom und Außentemperatur für jede Stunde einzeln dargestellt werden. Dies ist beispielhaft für VB1 in Abbildung 5.1 realisiert, wo die Messdaten farblich in 24 Stunden aufgeteilt sind.

Hier werden 11 Arbeitstage berücksichtigt ($N_{TT} = 11$), womit für jede Stunde 11 Messdatenpaare (Messpunkte) zur Verfügung stehen. Die Geraden entsprechen den 24 Regressionsmodelle gemäß Ansatz (5.1), deren Parameter mittels LAV-Kostenfunktion bestimmt wurden.

Diese Messdaten wurden gewählt, da hier für 21:00 ein steigend linearer Zusammenhang zwischen Wärmestrom und Außentemperatur geschätzt wird. Dieser Zusammenhang widerspricht physikalischen Überlegungen und sollte daher korrigiert werden. Dies wird mit einer Regel berücksichtigt, welche in Anlehnung an die Experten-Systeme in Kapitel 3 als Expertenregel bezeichnet wird.

Expertenregel 1. Falls für die Stunde n_0 ein steigend linearer Zusammenhang zwischen Wärmestrom und Außentemperatur geschätzt wird ($\beta[n_0] > 0$), so wird der Einfluss der Außentemperatur eliminiert ($\beta[n_0] = 0$), der Prädiktor für diese Stunde somit neu angesetzt als $\hat{Q}_P[n_0] = \alpha[n_0]$ und der Parameter $\alpha[n_0]$ entsprechend gewählter Kostenfunktion aus den Trainingsdaten bestimmt.

In der Implementierung mittels LAV wird hierfür die zweite Spalte der Matrix \mathbf{A} auf Null gesetzt, sprich $\mathbf{A}[n_0] = [\mathbf{1} \ \mathbf{0}]$, und daraufhin der Parametervektor $\mathbf{x}[n_0]$ neu bestimmt. Mit dieser Expertenregel wird die steigende Gerade in Abbildung 5.1 durch eine horizontale Gerade ersetzt, womit dieses selten auftretende Problem behandelt ist.

Anzahl der Trainingstage

Um die Zahl der Variablen in diesem Optimierungsproblem zu reduzieren, wird die Anzahl an Trainingstagen N_{TT}^* für alle sieben Verbraucher bestimmt und darauf fixiert. So wird der Einfluss von N_{TT} auf alle Verbraucher gemeinsam untersucht und basierend auf diesen Ergebnissen entschieden welches N_{TT}^* immer verwendet werden sollte. Das zu verwendende N_{TT}^* wird für die Arbeitswoche und das Wochenende getrennt untersucht, da hier andere Ergebnisse zu erwarten sind. Zum einen ist die Methode flexibler (adaptiver) je weniger Tage berücksichtigt werden, da sie sich schneller an Änderungen des Benutzerverhaltens, an Änderungen der Gebäudesubstanz und an neue Wetterbedingungen anpasst. Zum anderen ist die Methode robuster je mehr Tage berücksichtigt werden, da Messfehler und Ausreißer weniger stark in die Parameterbestimmung eingehen. Zwischen der Extreme einer sehr flexiblen bzw. sehr robusten Methode liegt eine

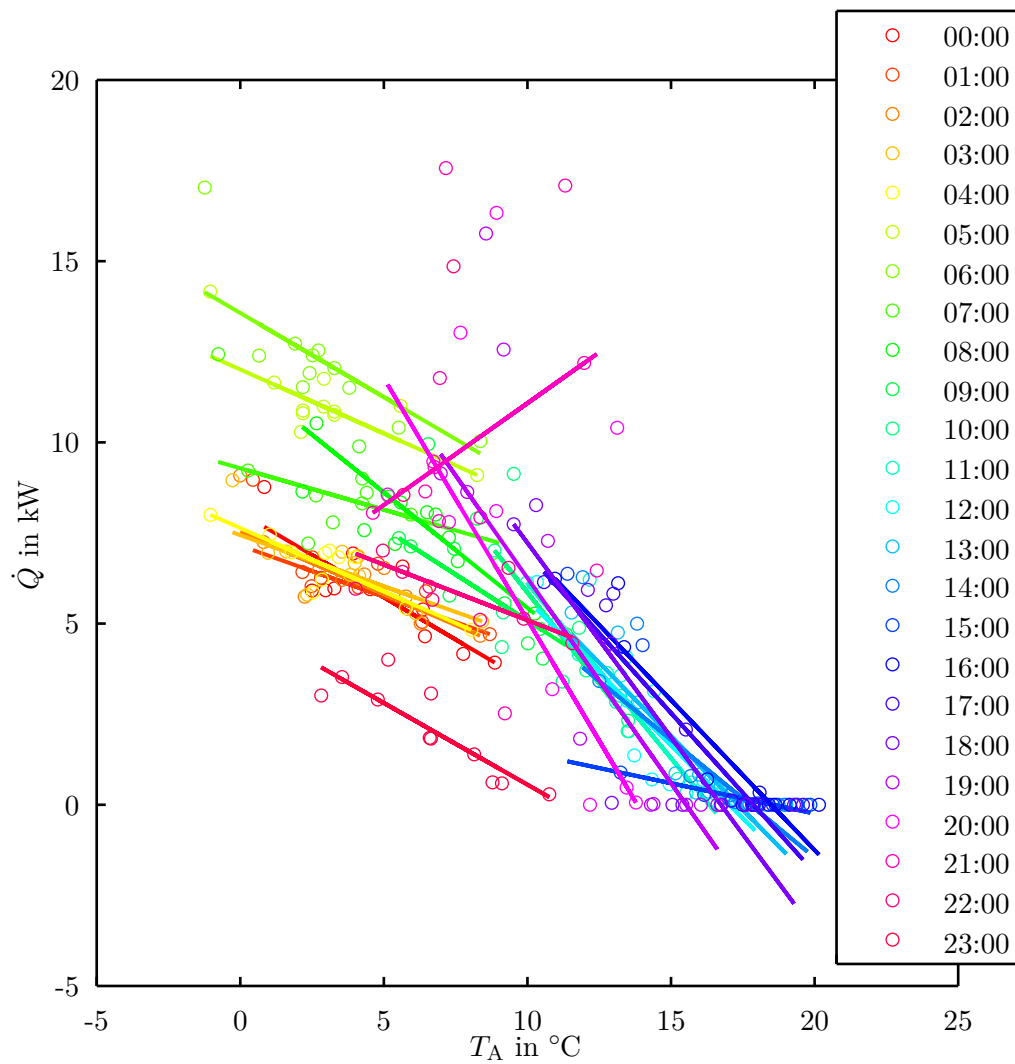


Abbildung 5.1.: VB1 – Stündliche Messdaten des Wärmestroms über die Außentemperatur für die Arbeitswoche aufgetragen. Die Daten stammen von 11 Arbeitstagen im Frühling und die Schätzungen wurden mit LAV bestimmt.

beste Anzahl an Trainingstagen, die es zu finden gilt.

Die Bestimmung der besten Anzahl der Trainingstage N_{TT}^* erfolgt in drei Schritten, wobei diese für Arbeitswoche und Wochenende getrennt durchzuführen sind:

Schritt 1: Bestimme für jeden Verbraucher getrennt die beste Anzahl der Trainingstage je Kalenderwoche $N_{\text{TT, KW}}^*$, was als Optimum der Kalenderwoche bezeichnet wird. Dadurch kann für jeden Verbraucher einzeln untersucht werden, ob für

5. Prognosemethode

gewisse Kalenderwochen (Monate, Jahreszeiten) ein Trend für $N_{TT, KW}^*$ zu erkennen ist.

Schritt 2: Bestimme für jeden Verbraucher einzeln die beste Anzahl der Trainingstage für alle Kalenderwochen $N_{TT, VB}^*$, was als Optimum des Verbrauchers bezeichnet wird. Somit wurde für jeden der sieben Verbraucher ein $N_{TT, VB}^*$ bestimmt und es kann auch hier ein Trend innerhalb der Verbraucher untersucht werden.

Schritt 3: Bestimme die beste Anzahl der Trainingstage für alle Verbraucher N_{TT}^* , was als globales Optimum aller Verbraucher bezeichnet wird. Dieser Zahlenwert wird für alle Verbraucher fixiert, womit ein sinnvoller Kompromiss für die praktische Implementierung gefunden wurde.

Im ersten Schritt wird das Optimum je Kalenderwoche $N_{TT, KW}^*$ eines Verbrauchers bestimmt. Hierfür wird, für jede Kalenderwoche extra, die Anzahl der Trainingstage von 1 bis N_{TT}^{\max} schrittweise erhöht und daraus eine Lastprognose für die nächste Kalenderwoche ermittelt. Es ist zu beachten, dass mit „Kalenderwoche“ für die Arbeitswoche nur fünf Tage und für das Wochenende nur zwei Tage gemeint sind. Da sich empirisch vier Kalenderwochen als sinnvolle Obergrenze herausgestellt haben wird die maximale Anzahl der Trainingstage für die Arbeitswoche bzw. das Wochenende auf

$$N_{TT, MoFr}^{\max} = 4 \text{ w} \cdot 5 \text{ d/w} = 20 \text{ d} \quad \text{bzw.} \quad N_{TT, SaSo}^{\max} = 4 \text{ w} \cdot 2 \text{ d/w} = 8 \text{ d} \quad (5.6)$$

festgelegt. Die Qualität dieser Lastprognose wird anhand des MAE für die gesamte Kalenderwoche bestimmt. Das Ergebnis dieser Vorgangsweise ist für VB1 in Abbildung 5.2 für alle Kalenderwochen zu sehen. In dieser Abbildung ist also die mittels MAE bewertete Qualität der Prognose in Abhängigkeit der Anzahl der Trainingstage aufgetragen.

Da gemäß (5.3) der Parametervektor je Stunde bestimmt wird, entspricht ein Trainingstag nur einem vergangenen Zahlenwert. Da der Parametervektor aber zwei Unbekannte (α und β) enthält, entspricht dies einer Gleichung mit zwei Unbekannten. Bei dem Lösungsalgorithmus wird bei solch einem unterbestimmten Gleichungssystem der zweite Parameter (hier β) auf Null gesetzt und dann erst die nun bestimmte Gleichung gelöst. In diesem Fall entspricht α einfach dem Zahlenwert vor 24h: $\alpha[n] = \dot{Q}[n - 24]$. Werden nun zwei Trainingstage verwendet so entspricht dies den vergangenen zwei Zahlenwerten, womit ein bestimmtes Gleichungssystem aufgestellt und ohne Optimierung gelöst werden kann. Erst ab drei Trainingstagen ist das Gleichungssystem überbestimmt und somit ist eine Minimierung des Residuums notwendig.

In Abbildung 5.2 ist zu erkennen, dass ein einzelner Trainingstag für einige Kalenderwochen bereits gute Ergebnisse liefern würde. In diesem Fall ist $\alpha[n] = \dot{Q}[n - 24]$ und $\beta = 0$, wodurch keine Abhängigkeit der Temperatur berücksichtigt wird. Für zwei Trainingstage verschlechtert sich das Ergebnis drastisch, da hier die Temperaturkomponente der vergangenen zwei Tage die Zukunft schlecht beschreibt. Werden mehr Trainingstage berücksichtigt, so wird das Ergebnis zuerst besser, da die Methode an Robustheit zunimmt. Ab einer gewissen Anzahl an Trainingstagen verschlechtert sich das Ergebnis, da die Methode an Flexibilität verliert.

Nun kann das Minimum der einzelnen Kurvenverläufe in Abbildung 5.2 bestimmt werden, womit der minimale MAE der jeweiligen Kalenderwoche ermittelt wurde. Die-

se minimalen MAE sind in Abbildung 5.3 für alle Kalenderwochen dargestellt, damit die Kurven in Abbildung 5.2 den Kalenderwochen farblich zugeordnet werden können. Ansonsten wird der Zahlenwert des minimalen MAE nicht weiter benötigt.

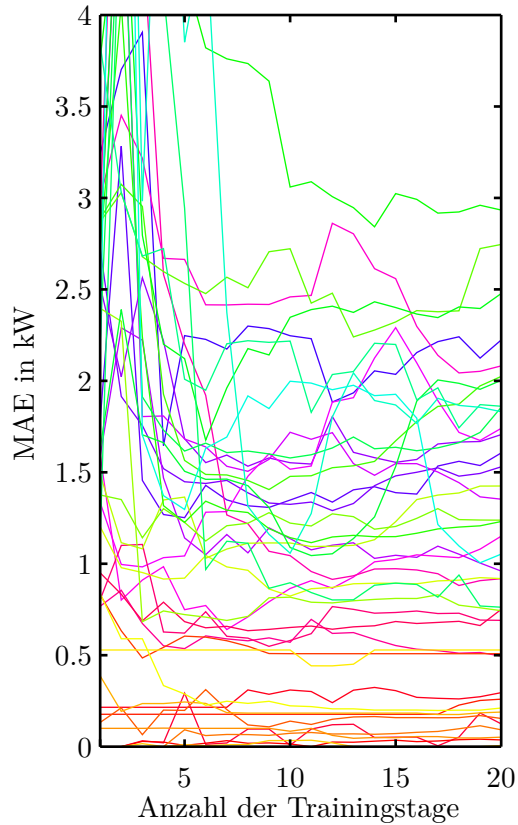


Abbildung 5.2.: VB1 – Qualität der Prognose (mittels MAE bewertet) in Abhängigkeit von der Anzahl der Trainingstage für die Arbeitswoche.

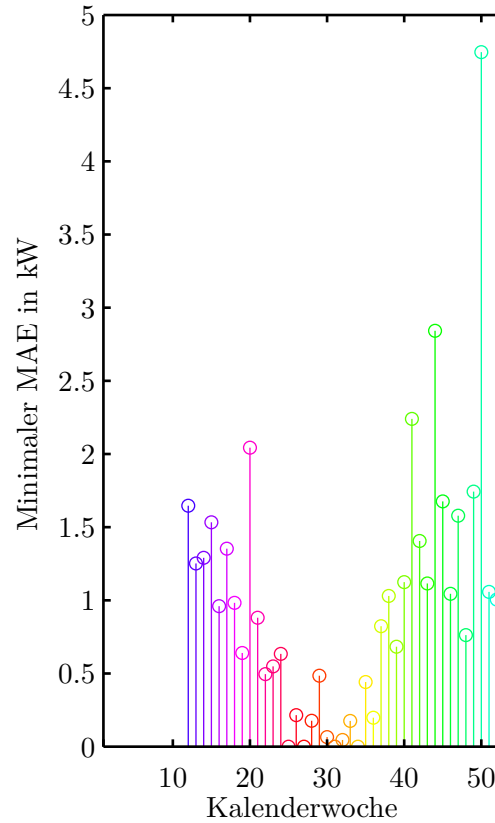


Abbildung 5.3.: VB1 – Minimum der einzelnen Kurvenverläufe aus Abbildung 5.2. Am Beginn des Jahres fehlen für diesen Verbraucher Daten.

Jetzt wird je Kalenderwoche die Anzahl der Trainingstage welcher zum minimalen MAE gehört bestimmt, womit ein Optimum $N_{TT, KW}^*$ für jede Kalenderwoche gefunden wird. Diese einzelnen Optima sind für VB1 (Wohnhaus) in Abbildung 5.4a und für VB6 (Nahwärmenetz) in Abbildung 5.4b dargestellt. Diese beiden Verbraucher wurden aufgrund ihres Kontrastes zueinander und der nahezu vollständig vorhandenen Daten für die beispielhafte, graphische Darstellung ausgewählt. Bei VB1 fehlen zu Beginn des Jahres Daten und bei VB6 sind alle Daten vollständig. Jedoch ist die Auswertung für die ersten 4 Kalenderwochen nicht möglich, da hierfür gemäß (5.6) wieder 4 vorangehende Kalenderwochen nötig wären.

5. Prognosemethode

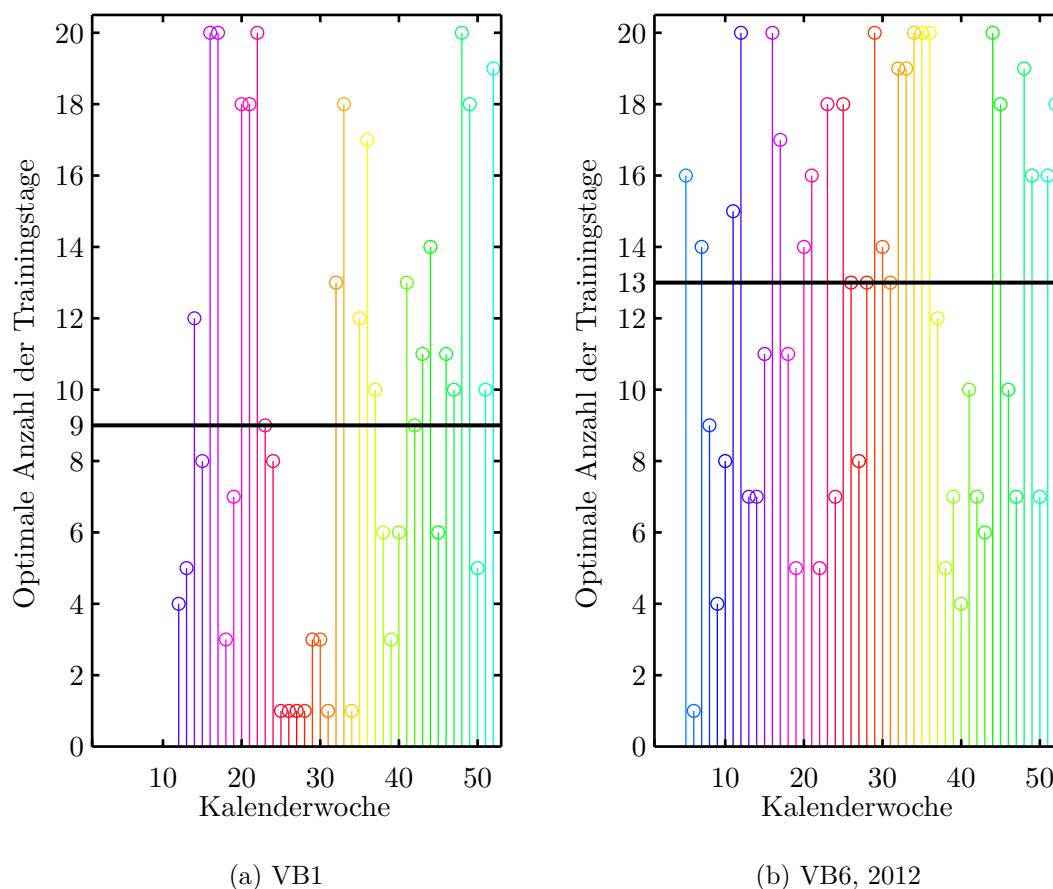


Abbildung 5.4.: Optimale Anzahl der Trainingstage je Kalenderwoche $N_{TT, KW}^*$ für die Arbeitswoche. Aus diesen einzelnen Optima für jede Kalenderwoche wird ein Optimum des Verbrauchers $N_{TT, VB}^*$ bestimmt, was als schwarze, horizontale Linie eingezeichnet ist.

Im zweiten Schritt kann nun das Optimum des Verbrauchers $N_{TT, VB}^*$ bestimmt werden. Damit diese Bestimmung robust gegen Ausreißer innerhalb der $N_{TT, KW}^*$ ist, wird $N_{TT, VB}^*$ mittels Median aus allen $N_{TT, KW}^*$ (ganzes Jahr) ermittelt. Dieses Optimum ist in Abbildung 5.4a bzw. 5.4b als schwarze, horizontale Linie eingezeichnet und liegt bei 9 bzw. 13.

Im dritten Schritt wird mittels Median aus diesen sieben lokalen Optima je Verbraucher ein globales Optimum aller Verbraucher N_{TT}^* bestimmt. Für die Arbeitswoche liegen die Optima je Verbraucher im Bereich der bereits bestimmten Optima von VB1 (9) und VB6 (13), weshalb ein globales Optimum aller Verbraucher als gute Näherung bestimmt werden kann. Dieses globale Optimum lautet für die Arbeitswoche bzw. das Wochenende:

$$N_{TT, MoFr}^* = 11 \text{ d} \quad \text{bzw.} \quad N_{TT, SaSo}^* = 5 \text{ d.} \quad (5.7)$$

Es wäre auch möglich N_{TT}^* abhängig von der Kalenderwoche variabel zu wählen. So könnten weniger Tage in der Übergangszeit als im Winter sinnvoll sein. Aus den Verläufen in Abbildung 5.4a und 5.4b, sowie den Verläufen für die restlichen Verbraucher ist kein derartiger Trend zu erkennen, der für alle Verbraucher gelten würde. Man könnte auch für jeden Verbraucher einzeln ein N_{TT}^* bestimmen, was jedoch in der praktischen Umsetzung aufwändig wäre. Zusätzlich bewegen sich die Optima für die Kalenderwochen in einem breiten Band, womit ein fixes N_{TT}^* , selbst wenn es für den Verbraucher explizit bestimmt wurde, immer nur ein Kompromiss sein kann. Deshalb wurde auf so eine variable Anzahl an Trainingstagen verzichtet und entsprechend (5.7) für alle Verbraucher und Kalenderwochen fixiert.

5.2.3. Ergebnis

Der Ansatz für den ersten Prädiktor ist definiert und die Parameterbestimmung ist geklärt. Nun kann auf die Qualität der Prädiktion (Prognose) näher eingegangen werden. Die Prädiktionsergebnisse wurden zwar schon verwendet um mittels minimalem MAE die optimale Anzahl an Trainingstagen zu fixieren, doch wurde diese Prädiktion noch nicht genauer erläutert. Zuerst wird anhand einer Beispielwoche die Prädiktion erklärt und dann dieser erste Prädiktor für alle Verbraucher evaluiert.

Als Beispielwoche wurde eine gesamte Arbeitswoche im Frühling des VB1 gewählt, da hier ein interessanter Effekt auftritt. Für die Parameterbestimmung werden entsprechend (5.7) die vorangegangenen 11 Arbeitstage betrachtet. Die verwendeten Trainingsdaten sind in Abbildung 5.1 zu sehen, wobei die Expertenregel 1 in diesem Beispiel noch nicht angewandt wurde.

Angenommen die Parameter sind für alle Stunden bestimmt, dann kann entsprechend (5.1) mit einer prognostizierten Außentemperatur der zukünftige Wärmestrom prädiziert werden. Da diese Auswertung auf Messdaten der Vergangenheit basieren sind Messungen der Außentemperatur vorhanden und werden für die Evaluierung eingesetzt. Somit ist das Ergebnis der Lastprognose unabhängig von der eingesetzten Wetterprognosemethode, was für eine unabhängige Bewertung der Lastprognose notwendig ist. Unter Verwendung der realen Außentemperatur kann nun eine gesamte Arbeitswoche prädiziert werden. Hierfür werden mittels der 24 Regressionsmodelle Stützstellen zu jeder vollen Stunde berechnet und zwischen diesen Stützstellen linear interpoliert. Diese Prädiktion wird in Abbildung 5.5 mit dem realen Wärmestrom verglichen. Hierbei ist zu beachten, dass die Parameter $\alpha[n]$ und $\beta[n]$ für $n = 0, \dots, 23$ vor Montag bestimmt wurden und dann innerhalb der Arbeitswoche nicht aktualisiert werden. Dieser Prädiktionshorizont von 5 Tagen ist aufgrund der guten Darstellbarkeit und Auswertung gewählt; es könnten methodisch aber auch mehr bzw. weniger Stunden oder Tage prädiziert werden.

Betrachtet man Abbildung 5.5, so stimmt der prädizierte Wärmestrom \hat{Q}_P erstaunlich gut mit dem realen Wärmestrom \dot{Q} überein, insbesondere wenn man die Einfachheit der Methode berücksichtigt. In Zahlen entspricht dieser Verlauf einem MAE von 1,63 kW, einem MPE von $-26,4\%$ und einem MAPE von $40,9\%$. Wie in Beispiel 2.1 wird hier für die Bestimmung des MPE bzw. MAPE eine Datenschanke von einem Zehntel des

5. Prognosemethode

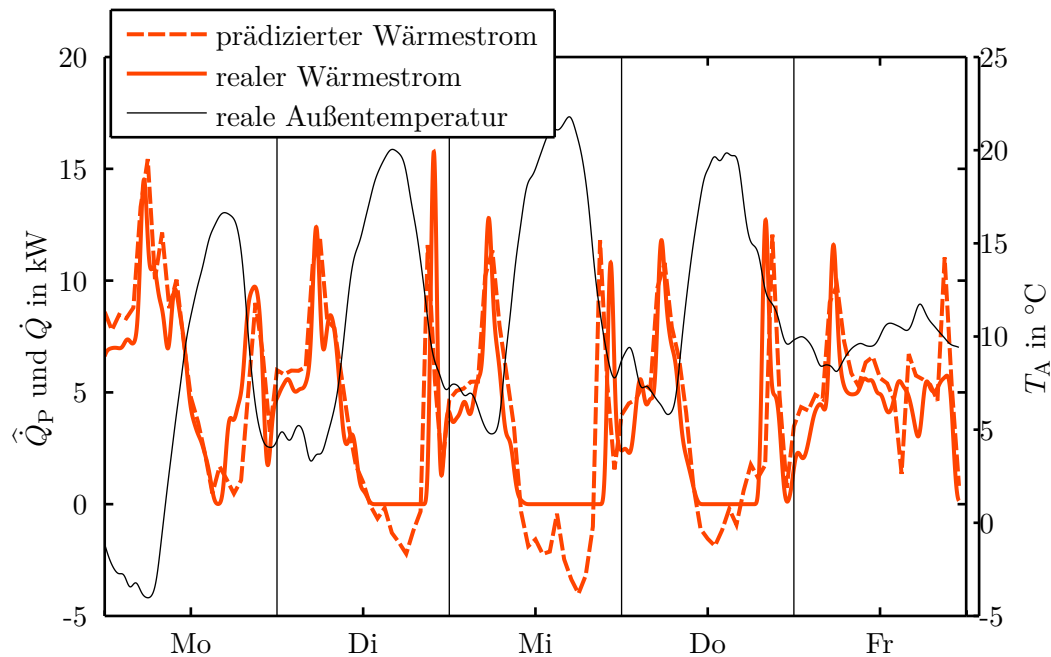


Abbildung 5.5.: VB2 – Vergleich zwischen prädiziertem und realem Wärmestrom für eine Arbeitswoche. Die dazugehörige reale Außentemperatur dient als Eingangsgröße des Prädiktors und ist daher ebenfalls eingezeichnet.

mittleren, realen Wärmestroms eingeführt. Alle Daten unter dieser Schranke werden für die Berechnung des MPE bzw. MAPE nicht berücksichtigt. Dies erklärt auch den stark negativen MPE, was gemäß (2.10) einem zu großen prädizierten Wärmestrom entspricht, da die großen Abweichungen bei $\dot{Q} = 0$ nicht in diese Berechnung eingehen.

Diese großen Abweichungen bei $\dot{Q} = 0$ sind durch die Gebäudeautomation zu erklären. Die Grundlast des Wärmestroms bildet die Heizung, die ab einer gewissen Außentemperatur nicht mehr heizen muss und abgeschaltet wird. Das lineare Regressionsmodell interpoliert hier natürlich auch in negative Bereiche des Wärmestroms hinein, weshalb hier eine weitere Expertenregel erforderlich ist.

Expertenregel 2. Falls ein negativer Wärmestrom prädiziert wird, so ist die Prädiktion für diesen Zeitraum auf Null zu setzen.

Diese Expertenregel darf nicht angewandt werden, wenn der Verbraucher auch in der Lage ist Wärmeenergie in den Pufferspeicher zurückzuspeisen. Solch ein Verbraucher wird Prosumer genannt, was sich aus Consumer und Producer zusammensetzt, wird im Zuge dieser Arbeit aber nicht näher betrachtet.

Die Messpunkte bei $\dot{Q} = 0$ werden für die Parameterbestimmung aber weiterhin verwendet. So ist in Abbildung 5.1 gut zu sehen wie durch diese Messpunkte für 15:00 eine nahezu horizontale Gerade erzeugt wird, was den Zusammenhang für diese Stunde gut beschreibt. Würden diese Messpunkt nicht in Parameterbestimmung eingehen, so

würden durch die wenigen verbleibenden Messpunkte nicht robuste Regressionsmodelle (Geraden) erzeugt werden, die den Zusammenhang für diese Stunde unter Umständen nur schlecht beschreiben.

Nachdem die Prädiktion anhand einer Arbeitswoche beschrieben wurde, wird dieser erste Prädiktor für alle Verbraucher evaluiert. In dieser Evaluierung wird immer eine gesamte Arbeitswoche bzw. ein gesamtes Wochenende mit der realen Außentemperatur prädiziert, und mit MPE bzw. MAPE bewertet. Da ein Jahr 52 oder 53 Kalenderwochen enthält, wären für die sieben Verbraucher mindestens $52 \cdot 7 = 364$ Bewertungen erforderlich, um die Qualität der Prognosemethode für ein Jahr zu beschreiben. Um dies übersichtlicher zu gestalten, wird das Jahr gemäß Tabelle 5.1 in Jahreszeiten unterteilt und die Qualitätsmaße werden für diese Zeiträume gemittelt. Sind bei einem Verbraucher in einer Jahreszeit Daten aus mehreren Jahren vorhanden, so wird dies ebenfalls durch Mittelwertbildung berücksichtigt. Die Evaluierung wird für die Arbeitswoche bzw. das Wochenende getrennt durchgeführt, da sich die Prognosemethode hier in der Anzahl der Trainingstage gemäß (5.7) und in der Auswahl der Trainingsdaten für Parameterbestimmung unterscheidet.

Das Ergebnis der Evaluierung ist für die Arbeitswoche in Tabelle 5.2 bzw. für das Wochenende in Tabelle 5.3 zu sehen. Diese Ergebnisse sind auch graphisch in Abbildung 5.6 bzw. in Abbildung 5.7 aufbereitet, um eine leichtere Interpretation zu ermöglichen.

Jahr	Kalenderwochen der Jahreszeit			
	Frühling	Sommer	Herbst	Winter
2012	KW12 – KW24	KW25 – KW38	KW39 – KW51	KW52 – KW11
2013	KW12 – KW25	KW26 – KW38	KW39 – KW51	KW52 – KW11
2014	KW12 – KW25	KW26 – KW38	KW39 – KW51	KW52 – KW11

Tabelle 5.1.: Kalenderwochen der Jahreszeiten als Evaluierungszeiträume.

Für die folgende Interpretation der Ergebnisse wird daran erinnert, dass sich beim MPE positive und negative Fehler aufheben, womit er systematische Fehler aufzeigen kann. Der MAPE berücksichtigt den Absolutwert des Fehler relativ zum realen Wärmestrom, und ist somit ein übliches Qualitätsmaß. Da beides relative Größen sind, kann die Qualität zwischen den einzelnen Jahreszeiten bzw. Verbrauchern verglichen werden, was mit MAE nicht möglich wäre.

Der MPE in Tabelle 5.2 und 5.3 ist zum Großteil negativ, was gemäß (2.10) einem zu großen prädizierten Wärmestrom entspricht. Die großen negativen Werte im Frühling resultieren aus der stetig sinkenden Heizlast. Somit wird das Regressionsmodell durch die Trainingsdaten an einen zu hohen Wärmestromverlauf angepasst, als dieser dann in der Prognosewoche (Bewertungswoche) auftritt. Für den Sommer resultiert dies, aus einem Wärmestrom nahe der Nulllinie, wobei durch Expertenregel 2 negative Prädiktionen verhindert werden. Für den Herbst nimmt nun, entgegen dem Frühling, der Wärmestrom im Schnitt zu wodurch zu wenig prädiziert wird, was sich einem positiven MPE äußert.

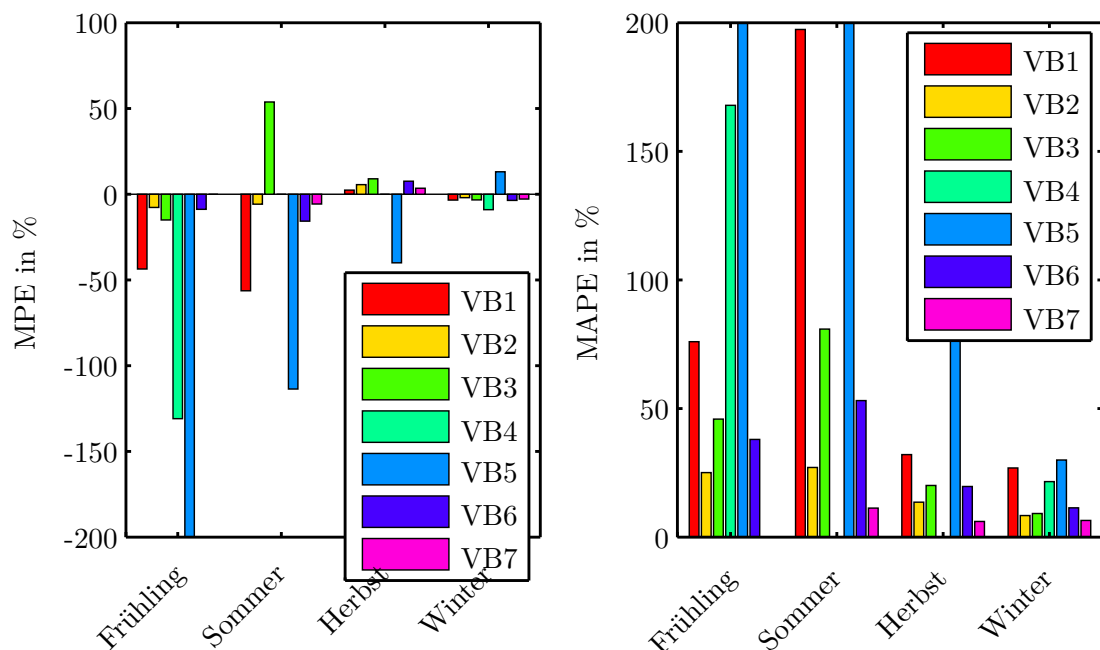
5. Prognosemethode

Die MPE-Werte in Winter sind eine Mischung der Effekte von Frühling und Herbst zu erklären. Insgesamt ist der MPE in Herbst und Winter deutlich niedriger als in Frühling und Sommer, da in den kalten Jahreszeiten die Heizung eine beständige Grundlast liefert. Dieser systematische Fehler der Prognosemethode ist nur schwer allgemein zu behandeln, da es zum Beispiel im Frühling auch Tage gibt an denen es wieder kälter wird. Daher bleibt dieser systematische Fehler unberücksichtigt.

Der MAPE beschreibt die Qualität der Prädiktion besser als der MPE da sich positive und negative Fehler nicht gegenseitig aufheben und ist ein übliches Bewertungsmaß in der Literatur. In den Jahreszeiten mit einer höheren Heizlast ist die Prädiktion besser, da sich diese Lastkomponente stark an der Außentemperatur orientiert, was mit dem Regressionsmodell abgebildet ist. Im Sommer wird die Last vom Benutzerverhalten dominiert, das sichtlich schwieriger zu präzisieren ist. Vergleicht man die Verbraucher untereinander so liefert die Prädiktion mit steigender Anschlussleistung (siehe Tabelle 4.1) bessere Ergebnisse. Eine Ausnahme ist hier das Mikrowärmenetz VB5, welches durch sein unregelmäßiges Verhalten große MAPE-Werte bewirkt. Besonders bei den Wärmenetzen (VB5, VB6 und VB7), die der Größe nach geordnet sind, führt ein Zunehmen der Anschlussleistung zu deutlich besseren Ergebnissen. Außerdem ist ein regelmäßiges Benutzerverhalten wie beim Bürogebäude (VB2) und dem Bürogebäude mit Gewerbe (VB3) deutlich einfacher zu präzisieren, als ein unregelmäßiges Benutzerverhalten wie beim Wohnhaus (VB1) und dem Doppelhaus (VB4). Schlussendlich liefert die Arbeitswoche im Schnitt auch bessere Ergebnisse als das Wochenende, was auf die zeitliche Verteilung der Trainingsdaten und das sich am Wochenende stärker ändernde Benutzerverhalten zurückzuführen ist.

Bez.	MPE in %				MAPE in %			
	Frühling	Sommer	Herbst	Winter	Frühling	Sommer	Herbst	Winter
VB1	-43,6	-56,3	2,4	-3,4	76,0	197,4	32,1	26,9
VB2	-7,7	-5,8	5,6	-2,0	25,1	27,1	13,6	8,4
VB3	-15,0	53,8	9,0	-3,3	45,9	80,9	20,1	9,2
VB4	-130,9	–	–	-9,0	167,9	–	–	21,6
VB5	-333,2	-113,6	-40,0	13,1	407,8	201,0	108,4	30,0
VB6	-8,8	-15,7	7,6	-3,6	38,0	53,1	19,7	11,4
VB7	–	-5,7	3,5	-2,8	–	11,3	6,1	6,5

Tabelle 5.2.: Evaluierung des ersten Prädiktors für die *Arbeitswoche*. Bei den Feldern die mit „–“ gekennzeichnet sind, stehen keine Daten zur Verfügung.



(a) MPE – Die y-Achse wurde auf -200% beschnitten.

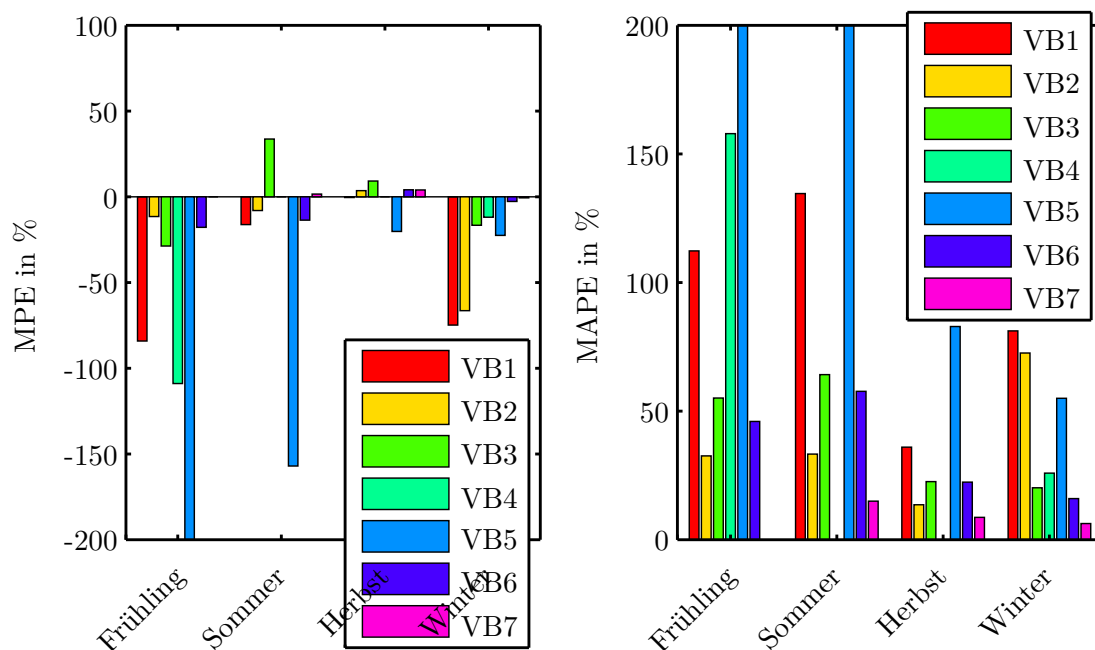
(b) MAPE – Die y-Achse wurde auf 200% beschnitten. VB5 ist im Herbst verdeckt und entspricht 108%.

Abbildung 5.6.: Evaluierung des ersten Prädiktors für die *Arbeitswoche*. Die exakten Werte sind Tabelle 5.2 zu entnehmen.

5. Prognosemethode

Bez.	MPE in %				MAPE in %			
	Frühling	Sommer	Herbst	Winter	Frühling	Sommer	Herbst	Winter
VB1	-84,1	-16,2	-0,4	-74,8	112,3	134,6	36,0	81,2
VB2	-11,5	-8,0	3,6	-66,4	32,6	33,3	13,6	72,6
VB3	-28,7	33,7	9,2	-16,6	55,1	64,2	22,6	20,2
VB4	-108,9	–	–	-11,9	157,9	–	–	25,9
VB5	-274,5	-157,0	-20,2	-22,5	347,9	225,5	82,9	55,0
VB6	-17,8	-13,6	4,1	-2,7	46,0	57,7	22,4	16,0
VB7	–	1,6	4,0	-0,6	–	15,0	8,7	6,3

Tabelle 5.3.: Evaluierung des ersten Prädiktors für das *Wochenende*. Bei den Feldern die mit „–“ gekennzeichnet sind, stehen keine Daten zur Verfügung.



(a) MPE – Die y-Achse wurde auf –200 % beschnitten.

(b) MAPE – Die y-Achse wurde auf 200 % beschnitten.

Abbildung 5.7.: Evaluierung des ersten Prädiktors für das *Wochenende*. Die exakten Werte sind Tabelle 5.3 zu entnehmen.

5.3. Erweiterung des Prädiktors

In diesem Abschnitt wird der Prädiktor erweitert, wobei das bereits erfolgreiche erste Konzept als Basis dient. So bleibt die Trennung in Arbeitswoche und Wochenende, sowie die Aufteilung des Tages in Stunden erhalten. Es werden allerdings zusätzliche vielversprechende Eingangsgrößen berücksichtigt und deren Auswirkung auf das Prädiktionsergebnis untersucht.

5.3.1. Ansatz

Der Ansatz des ersten Prädiktors (5.1) soll um additive Terme erweitert werden, was einer multiplen linearen Regression entspricht. Hierfür werden Erweiterungsfunktionen f_{ij} spezifiziert, hinzugefügt und bezüglich ihres Verbesserungspotentials bewertet:

$$\hat{Q}_P[n] = \underbrace{\alpha[n] + \beta[n]\hat{T}_A[n]}_{\text{erster Ansatz}} + \gamma[n]f_{ij}, \quad n = 0, 1, \dots, 23, \quad (5.8)$$

mit

$$\begin{aligned} f_{00} &= 0, & f_{20} &= \hat{I}_G[n], \\ f_{11} &= \hat{T}_A[n-1], & f_{21} &= \hat{I}_G[n-1], \\ f_{12} &= \hat{T}_A[n+1], & f_{22} &= \hat{I}_G[n+1], \\ f_{13} &= (\hat{T}_A[n+1] - \hat{T}_A[n-1]), & f_{23} &= (\hat{I}_G[n+1] - \hat{I}_G[n-1]). \end{aligned}$$

Durch f_{00} kann der erste Ansatz unverändert in die Evaluierung der Erweiterungen integriert werden. Die Erweiterungsfunktionen f_{1j} repräsentieren Erweiterungen um die prognostizierte Außentemperatur \hat{T}_A und f_{2j} repräsentieren Erweiterungen um die prognostizierte Globalstrahlung \hat{I}_G . Bei f_{11} wird die jeweilige Wetterprognose der vorherigen Stunde und bei f_{12} die jeweilige Wetterprognose der kommenden Stunde berücksichtigt. Die Funktionen f_{i3} entsprechen dem mittleren Differenzenquotient, durch den die Ableitung zum Zeitpunkt n approximiert wird. Der hier übliche Faktor $\frac{1}{2}$ wurde nicht mitgezogen, da dies nur eine Skalierung von γ bedeuten würde.

5.3.2. Parameterbestimmung

Die Parameterbestimmung wird analog zur Bestimmung der Regressionsparameter des ersten Prädiktors (siehe Abschnitt 5.2) durchgeführt. Nachdem der Ansatz nur eine Erweiterung des ersten Ansatzes ist, bleibt die Anzahl der Parametersätze bei 48; 24 für Arbeitstage und 24 für Tage des Wochenendes. Diese Parametersätze können analog zu

5. Prognosemethode

(5.2) als Spaltenvektoren \mathbf{x}_i einer Parametermatrix \mathbf{X} angeschrieben werden

$$\begin{aligned} \mathbf{X} &= \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_{24} & \mathbf{x}_{25} & \mathbf{x}_{26} & \dots & \mathbf{x}_{48} \end{bmatrix} \\ &= \begin{bmatrix} \alpha_{\text{MoFr}}[0] & \alpha_{\text{MoFr}}[1] & \dots & \alpha_{\text{MoFr}}[23] & \alpha_{\text{SaSo}}[0] & \alpha_{\text{SaSo}}[1] & \dots & \alpha_{\text{SaSo}}[23] \\ \beta_{\text{MoFr}}[0] & \beta_{\text{MoFr}}[1] & \dots & \beta_{\text{MoFr}}[23] & \beta_{\text{SaSo}}[0] & \beta_{\text{SaSo}}[1] & \dots & \beta_{\text{SaSo}}[23] \\ \gamma_{\text{MoFr}}[0] & \gamma_{\text{MoFr}}[1] & \dots & \gamma_{\text{MoFr}}[23] & \gamma_{\text{SaSo}}[0] & \gamma_{\text{SaSo}}[1] & \dots & \gamma_{\text{SaSo}}[23] \end{bmatrix}, \end{aligned} \quad (5.9)$$

wobei die 48 Parametersätze insgesamt $3 \cdot 48 = 144$ Unbekannte beinhalten.

Die Parameterbestimmung wird wieder anhand der Arbeitswoche erklärt; beim Wochenende ist analog zu verfahren. Das überbestimmte Gleichungssystem hat die gleiche Gestalt wie (5.3) nur mit erweiterter Matrix \mathbf{A} und erweitertem Parametervektor $\mathbf{x} = [\alpha \ \beta \ \gamma]^\top$:

$$\mathbf{y}[n] = \mathbf{A}[n]\mathbf{x}[n] + \mathbf{e}[n], \quad n = 0, 1, \dots, 23, \quad (5.10)$$

mit

$$\begin{aligned} \mathbf{A}[n] &= \begin{bmatrix} \mathbf{1} & \mathbf{u}_1[n] & \mathbf{u}_2[n] \end{bmatrix} \\ &= \begin{bmatrix} 1 & T_{\text{A,MoFr}}[n - N_{\text{TT,MoFr}} \cdot 24] & f_{ij}[n - N_{\text{TT,MoFr}} \cdot 24] \\ 1 & T_{\text{A,MoFr}}[n - (N_{\text{TT,MoFr}} - 1) \cdot 24] & f_{ij}[n - (N_{\text{TT,MoFr}} - 1) \cdot 24] \\ \vdots & \vdots & \vdots \\ 1 & T_{\text{A,MoFr}}[n - 48] & f_{ij}[n - 48] \\ 1 & T_{\text{A,MoFr}}[n - 24] & f_{ij}[n - 24] \end{bmatrix}. \end{aligned} \quad (5.11)$$

Wieder wird durch Minimierung des Residuums gemäß der LAV-Kostenfunktion der optimale Parametervektor bestimmt. Hierbei ist die Spalte \mathbf{u}_2 für f_{00} immer Null, was im Linearen Programm zu einem $\gamma = 0$ führt. Somit ist sichergestellt, dass man für f_{00} das gleiche Ergebnis wie beim ersten Prädiktor erhält.

Da diese Prognosemethode nun von zwei unabhängigen Variablen (u_1 und u_2) abhängt, ist sie nicht mehr so einfach graphisch interpretierbar wie der erste Prädiktor in Abbildung 5.1. Es soll aber der stündliche Einfluss der Globalstrahlung näher betrachtet werden, da diese Wetterkomponente nun, trotz Ausscheiden in der Messdatenauswertung in Abschnitt 4.4, berücksichtigt wird. In der Messdatenauswertung wurde versucht, die Messdaten einer gesamten Woche durch lineare Regression zu modellieren. Da der hier verwendete Ansatz (5.8) den Tag in Stunden n aufteilt und zusätzlich zwischen Arbeitswoche und Wochenende unterscheidet, wird dies differenzierter berücksichtigt. Für die nun folgende Betrachtung wird ein lineares Regressionsmodell angesetzt, in dem der Einfluss der Außentemperatur nicht berücksichtigt wird:

$$\hat{Q}_P[n] = \alpha[n] + \gamma[n]\hat{I}_G[n]. \quad (5.12)$$

Anhand dieser 24 Modelle werden nun die Messdaten der Arbeitswoche von Abbildung 4.5 erneut analysiert. Hierfür wird in Abbildung 5.8 der Wärmestrom \dot{Q} über die

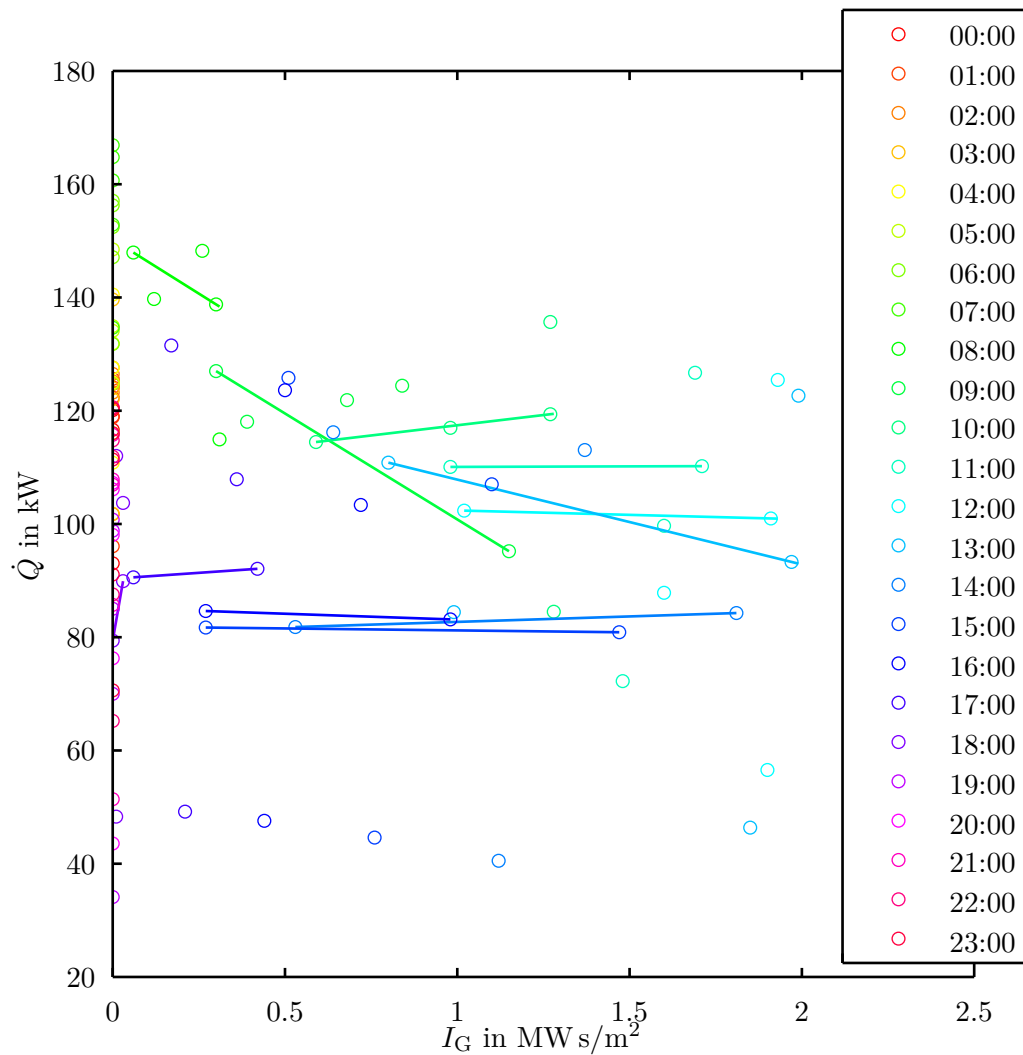


Abbildung 5.8.: VB2 – Wärmestrom in Abhängigkeit der Globalstrahlung

Globalstrahlung I_G aufgetragen, aber nun werden die LAV-Schätzungen für jede Stunde n einzeln bestimmt.

Die erste Erkenntnis ist hier, dass für die Nacht keine Regressionsmodelle bestimmt werden können. Daher gilt bei dem erweiterten Ansatz (5.8) für f_{2j} in der Nacht der erste Ansatz, womit die Erweiterung durch Globalstrahlung nur am Tag zum Einsatz kommt. Die zweite Erkenntnis gilt nur für diese Messdaten und betrifft die Geraden, welche die linearen Regressionsmodelle repräsentieren. Die Optimierung mit LAV liefert für einige Stunden nahezu horizontale Geraden, was in (5.12) einem $\gamma = 0$ entspricht. In anderen Worten, die Optimierung findet für diese Stunden keinen Zusammenhang zwischen Wärmestrom und Globalstrahlung. Die Globalstrahlung wird trotzdem in die

5. Prognosemethode

gesamte Evaluierung der Erweiterungen (5.8) aufgenommen, da für bestimmte Verbraucher und/oder Jahreszeiten ein anderer Zusammenhang gelten kann.

5.3.3. Ergebnis

Das Ergebnis der Erweiterung ist eine Erweiterungsfunktion f_{ij} gemäß (5.8), welche für alle Verbraucher die beste Prädiktion im Sinne des MAE liefert. Hierfür werden die Verbraucher getrennt betrachtet, wobei VB6 nicht evaluiert werden kann, da hier keine Globalstrahlungsdaten vorhanden sind. Je Verbraucher wird nach Jahreszeiten getrennt evaluiert, da gewisse Erweiterungen auch nur für bestimmte Jahreszeiten eine Verbesserung liefern könnten.

Für diese Evaluierung wird, analog zum ersten Prädiktor, jeweils eine gesamte Arbeitswoche bzw. ein gesamtes Wochenende evaluiert, ohne dass innerhalb dieser 5 bzw. 2 Tage die Parameter aktualisiert werden. Die Prädiktion jeder Kalenderwoche wird mittels MAE evaluiert und auf eine Jahreszeit gemittelt. Somit wird für jeden Verbraucher je Jahreszeit und Erweiterungsfunktion ein aussagekräftiger, gut vergleichbarer MAE-Wert bestimmt.

Dieser Vergleich ist in Balkendiagrammen aufbereitet, die für die Arbeitswoche in Abbildung 5.9a bis 5.10b, und für das Wochenende in Abbildung 5.10c bis 5.11d zu sehen sind. Laut dieser Evaluierung ist keine Erweiterungsfunktion f_{ij} für alle Verbraucher besser als der erste Prädiktor aus Abschnitt 5.2 (mit f_{00} gekennzeichnet). Es lässt sich auch für keine einzelne Jahreszeit eine Erweiterungsfunktion finden, die für alle Verbraucher besser ist. Jedoch für einzelne Jahreszeiten einzelner Verbraucher liefern gewisse f_{ij} eine Verbesserung, wie in Abbildung 5.9a für f_{11} in Sommer und Winter zu sehen ist.

Hierbei fällt auf, dass sich das Prognoseergebnis durch die Erweiterungsfunktionen auch verschlechtern kann. Zwar eliminiert die LAV-Parameteroptimierung die Erweiterungsfunktion durch ein $\gamma = 0$, falls die Erweiterung keinen Zusammenhang zum Wärmestrom aufweist. Dies ist für ein vereinfachtes Regressionsmodell (5.12) für einige Stunden in Abbildung 5.8 anhand der horizontalen Linien zu sehen. Jedoch können für die Trainingstage (scheinbare) Zusammenhänge gelten, welche für die Prognostage nicht mehr gelten, womit Verschlechterungen möglich sind. Würde man die Prognose unsinnigerweise für die Trainingstage durchführen, so wäre keine Verschlechterung möglich, da die Prognose mit dem MAE bewertet wird, welcher der LAV-Kostenfunktion entspricht.

Da die gesuchte Prognosemethode ohne manuelle Parametrierung für möglichst viele Verbraucher eingesetzt werden soll, wird keine Erweiterung hinzugenommen. Somit bleibt es beim ersten Prädiktor, der weiterhin als f_{00} bezeichnet wird.

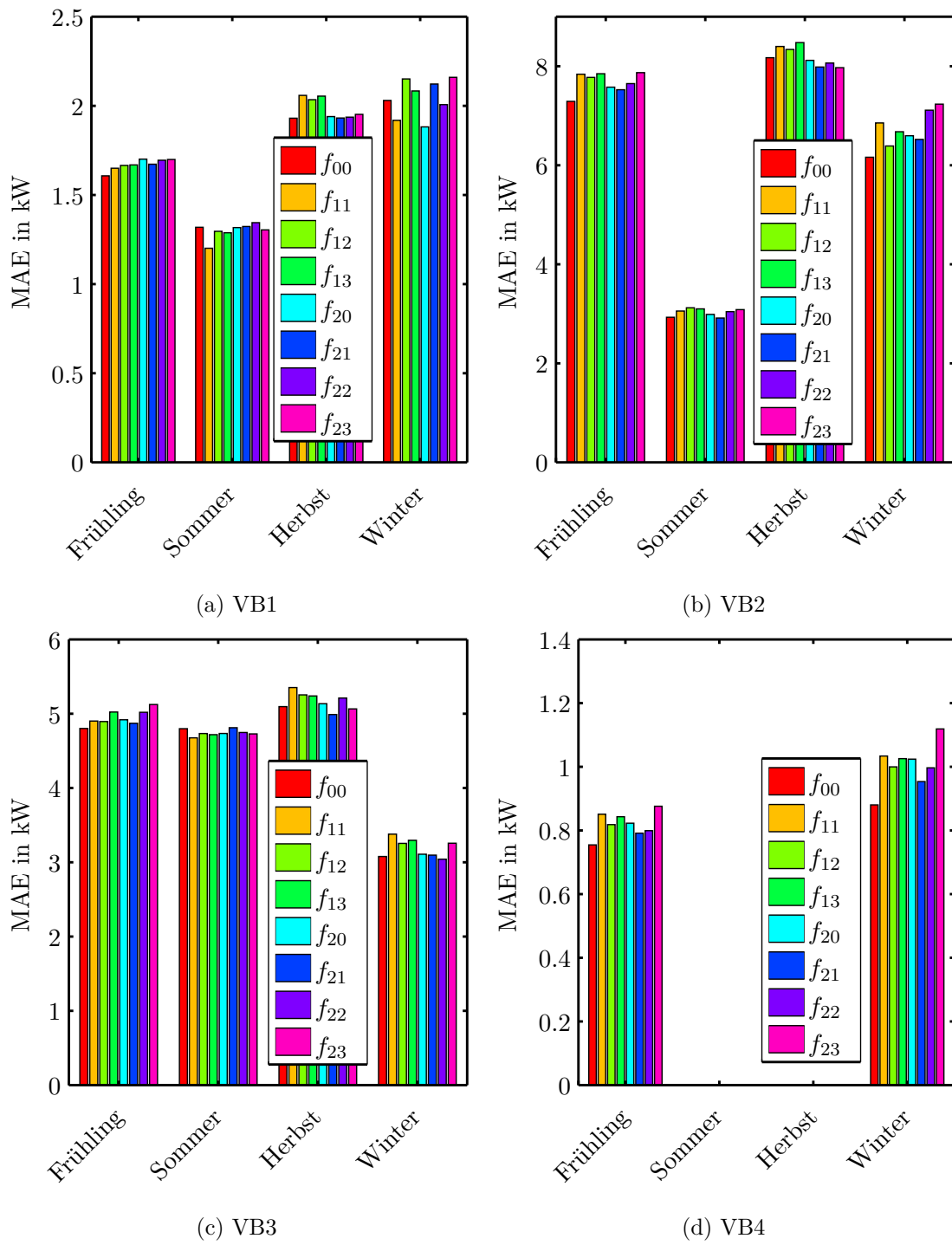
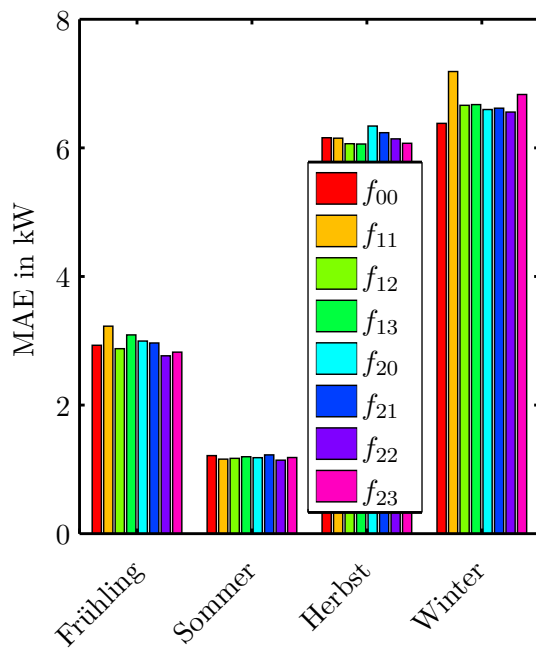
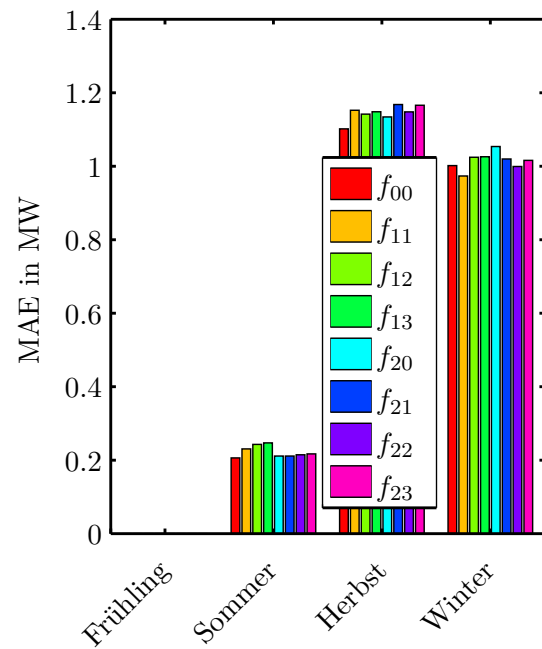


Abbildung 5.9.: Evaluierung der Erweiterungen f_{ij} gemäß (5.8) für die *Arbeitswoche*.

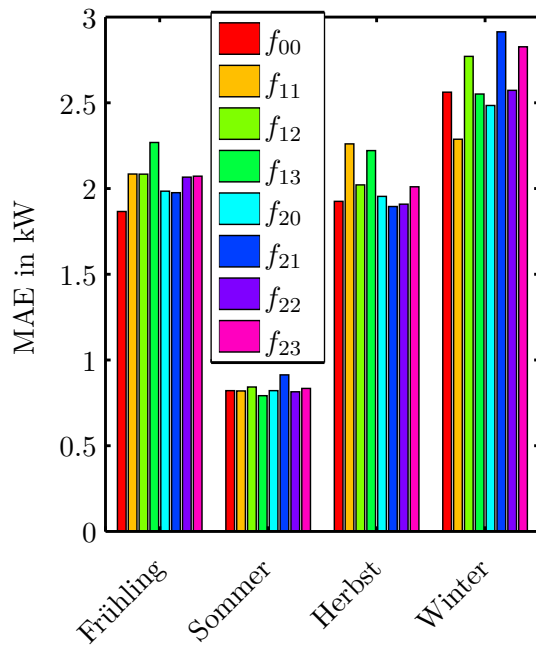
5. Prognosemethode



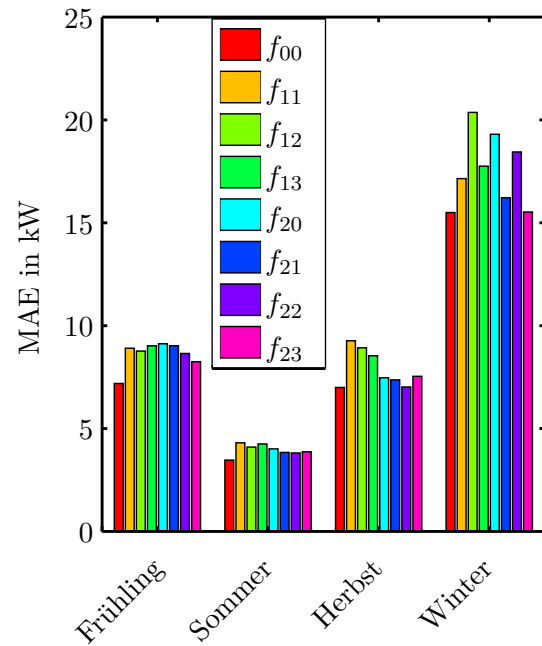
(a) VB5 – Arbeitswoche.



(b) VB7 – Arbeitswoche.



(c) VB1 – Wochenende.



(d) VB2 – Wochenende.

Abbildung 5.10.: Evaluierung der Erweiterungen f_{ij} gemäß (5.8) für die *Arbeitswoche* bzw. für das *Wochenende*.

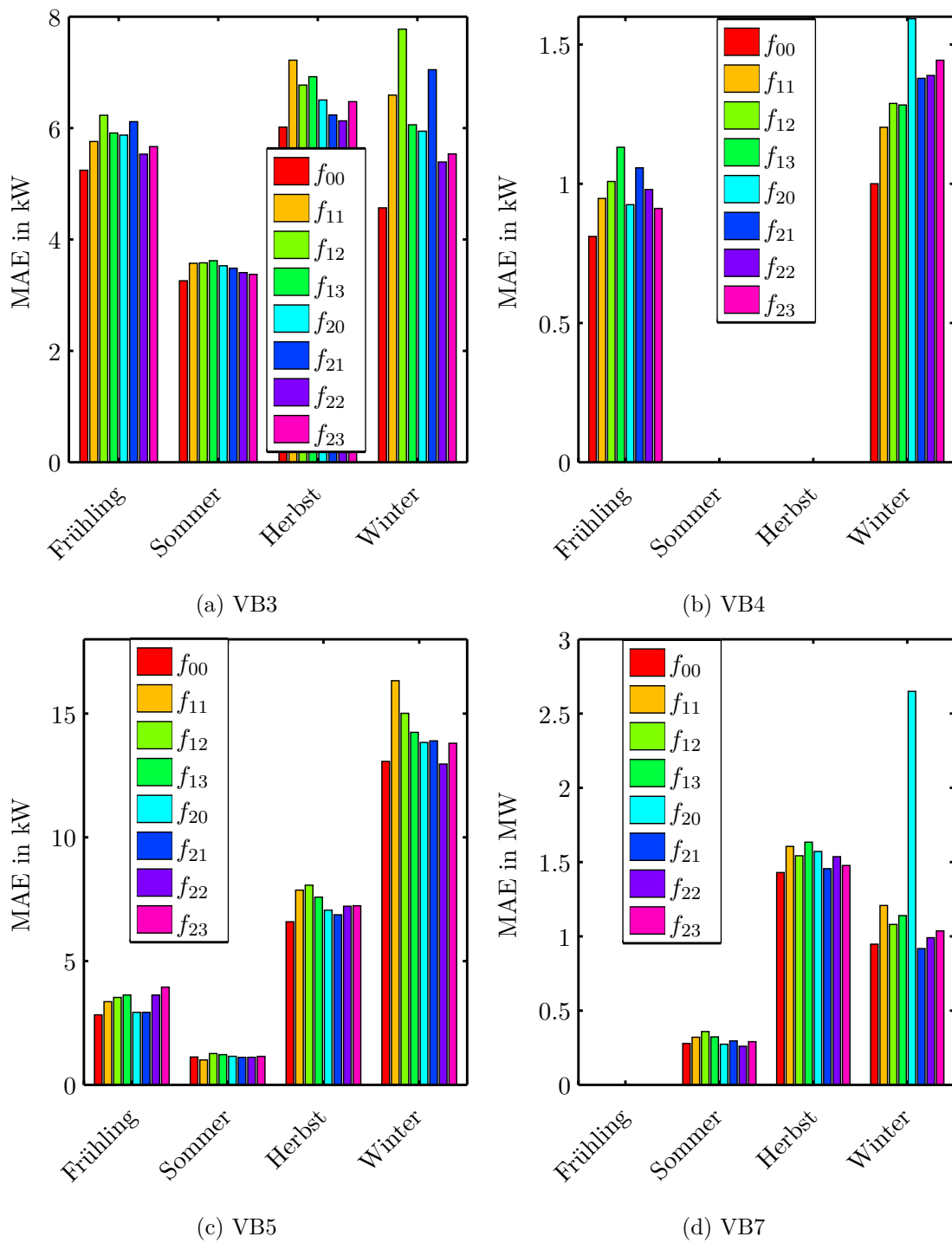


Abbildung 5.11.: Evaluierung der Erweiterungen f_{ij} gemäß (5.8) für das *Wochenende*.

5.4. Erster Korrektor

Der erste Korrektor entspricht einer Erweiterung des Prädiktors, hat aber aufgrund seiner gänzlich anderen Struktur einen eigenen Namen. Der Prädiktor besteht aus 24 statischen, linearen Regressionsmodellen, mit eigenen Konfigurationen für Arbeitstage bzw. für Tage des Wochenendes. Diese 24 Modelle unterscheiden sich nur durch ihre Parametersätze, die aus den stündlichen Messwerten der vergangenen Tagen bestimmt werden. Zum Beispiel soll der Arbeitstag 12 (00:00 bis 23:00) prädiziert werden, wofür Trainingsdaten von Arbeitstag 1 00:00 bis Arbeitstag 11 23:00 herangezogen werden. Hier könnte die Bestimmung der 24 Parametersätze am Arbeitstag 11 im Zeitraum von 23:01 bis 23:59 erfolgen. Wichtig ist nun, dass eine erneute Berechnung der Parametersätze innerhalb des Tages 12 die Prädiktion für diesen Tag nicht verbessern würde. Dies ist ein Resultat des Ansatzes (5.1) der die einzelnen, stündlichen Prädiktionen nicht miteinander verknüpft.

Nun ist aber zur aktuellen Stunde n_A bereits eine Messung des Wärmestroms $\dot{Q}[n_A]$ vorhanden, welche für die Prognose für $n > n_A$ verwendet werden könnte. Dies ist die Idee für einen Korrektor, der die Prädiktion der nahen Zukunft verbessern soll.

5.4.1. Ansatz

Der erste Korrektor wird als additive Korrektur der Prädiktion der nächsten Stunde $\hat{Q}_P[n_A + 1]$ angesetzt:

$$\hat{Q}_K[n_A + 1] = \hat{Q}_P[n_A + 1] + f_{K,i}(e_P[n_A - m]), \quad m = 0, 1, \dots, \quad (5.13)$$

mit dem Prädiktionsfehler

$$e_P[n] = \dot{Q}[n] - \hat{Q}_P[n]. \quad (5.14)$$

Die Korrekturfunktion $f_{K,i}$ verwendet den Prädiktionsfehler e_P der m vergangenen Stunden als einzige Variable.

Analog zu den Erweiterungsfunktionen aus Abschnitt 5.3 werden verschiedene Korrekturfunktionen $f_{K,i}$ betrachtet:

$$\begin{aligned} f_{K,0} &= 0, \\ f_{K,1} &= k_1 e_P[n_A] && \text{mit } 0 \leq k_1 \leq 1, \\ f_{K,2} &= k_2 \frac{e_P[n_A] + e_P[n_A - 1]}{2} && \text{mit } 0 \leq k_2 \leq 1. \end{aligned} \quad (5.15)$$

Hier ist wieder eine Korrekturfunktion $f_{K,0}$ enthalten, mit der die Erweiterung ignoriert wird. Bei $f_{K,1}$ wird der gewichtete Prädiktionsfehler der vergangenen Stunde addiert, womit bleibende Abweichungen reduziert werden:

- $e_P[n_A] < 0 \rightarrow$ Prädiktion zu groß $\rightarrow f_{K,1}$ wird subtrahiert,
- $e_P[n_A] > 0 \rightarrow$ Prädiktion zu klein $\rightarrow f_{K,1}$ wird addiert,
- $e_P[n_A] = 0 \rightarrow$ Prädiktion passt \rightarrow keine Korrektur.

Die Korrekturfunktion $f_{K,2}$ ist durch die Mittelwertbildung eine robustere Variante von $f_{K,1}$ und beruht auf den gleichen Überlegungen zur Reduktion des Prognosefehlers.

5.4.2. Parameterbestimmung

Analog zur Bestimmung der optimalen Anzahl an Trainingstagen wird der Korrektorparameter k_i mit $i = 1, 2$ für alle Verbraucher und Kalenderwochen bestimmt. Hier ist wieder das Ziel, die Komplexität der Prognosemethode möglichst gering zu halten.

Die Bestimmung des optimalen Korrektorparameters k_i^* erfolgt mit den gleichen drei Schritten, die schon bei der optimalen Anzahl an Trainingstagen in Abschnitt 5.2 beschrieben wurden. Diese drei Schritte werden hier nur kurz angeführt und sind ausführlich in Abschnitt 5.2 nachzulesen.

Schritt 1: Optima je Kalenderwoche eines Verbrauchers: $k_{i,KW}^*$.

Schritt 2: Optima der Verbraucher: $k_{i,VB}^*$.

Schritt 3: Optimum aller Verbraucher: k_i^* .

So wird im ersten Schritt das Optimum je Kalenderwoche $k_{i,KW}^*$ eines Verbrauchers bestimmt. Hierfür wird, für jede Kalenderwoche extra, der Korrektorparameter entsprechend (5.15) von 0 bis 1 in 0,1er-Schritten erhöht und daraus eine Lastprognose für die nächste Kalenderwoche ermittelt. Man beachte, dass ein Wert von $k_i = 0$ einer reinen Prädiktion (ohne Korrektor) entspricht. Die Qualität dieser Lastprognose wird anhand des MAE für die gesamte Kalenderwoche bewertet. Das Ergebnis dieser Vorgangsweise ist für k_1 für VB1 in Abbildung 5.12 für alle Kalenderwochen beispielhaft dargestellt. Die parabelförmigen Verläufe in dieser Abbildung lassen die Minima gut erkennen, welche in Abbildung 5.13 für die jeweilige Kalenderwoche aufgetragen sind.

Jetzt wird je Kalenderwoche der Korrektorparameter, welcher zu dem minimalen MAE gehört, bestimmt, womit ein Optimum für jede Kalenderwoche gefunden wurde. Wieder werden das Wohnhaus (VB1) in Abbildung 5.14a und das Nahwärmenetz (VB6) in Abbildung 5.14b als Beispiele für diese einzelnen Optima verwendet. Die große Streuung des Korrektorparameters über die Kalenderwochen für VB1 ist eine Ausnahme von dem üblichen Wertebereich der restlichen untersuchten Verbraucher von 0,5 bis 0,8.

Im zweiten Schritt wird nun mittels Median aus diesen einzelnen Optima je Kalenderwoche, das Optimum des Verbrauchers $k_{i,VB}^*$ bestimmt.

Im dritten Schritt wird aus diesen sieben Optima je Verbraucher mittels Median ein Optimum aller Verbraucher bestimmt:

$$\begin{aligned} k_{1,MoFr}^* &= 0,6 & \text{bzw.} & & k_{1,SaSo}^* &= 0,7, \\ k_{2,MoFr}^* &= 0,6 & \text{bzw.} & & k_{2,SaSo}^* &= 0,6. \end{aligned} \tag{5.16}$$

Die vorgestellte Methode ist nur eine von vielen Methoden, den Korrektorparameter zu bestimmen, liefert aber konstante Parameter, die in der Recheneinheit nicht mehr neu bestimmt werden müssen. Man hätte auch den Korrektorparameter in dem Parametervektor \mathbf{x} inkludieren können, wodurch die Trennung in Prädiktor und Korrektor hinfällig wäre. Dies hätte das Optimierungsproblem aber drastisch verkompliziert, weshalb darauf verzichtet wurde. Der Hauptgrund für die durchgeführte Methode ist der schmale Wertebereich in dem sich der Korrektorparameter bewegt, wie in Abbildung 5.14b als übliche

5. Prognosemethode

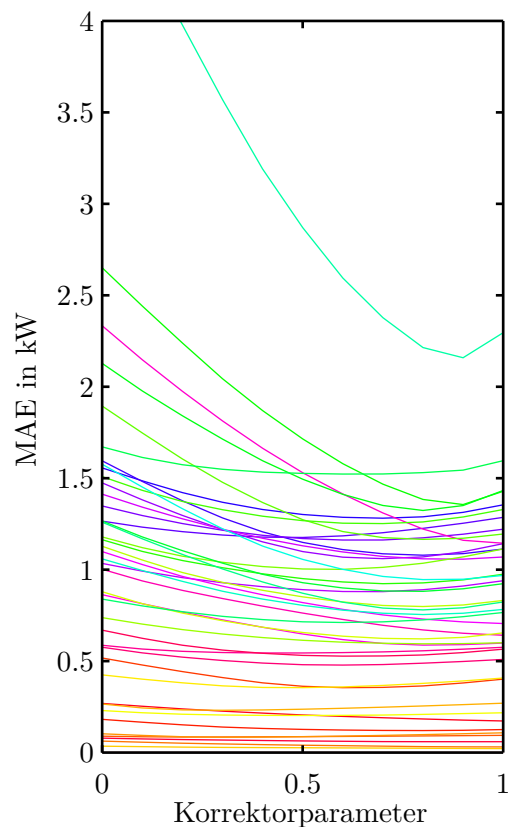


Abbildung 5.12.: VB1 – Qualität der Prognose (mittels MAE bewertet) in Abhängigkeit des Korrektorparameters k_1 für die Arbeitswoche.

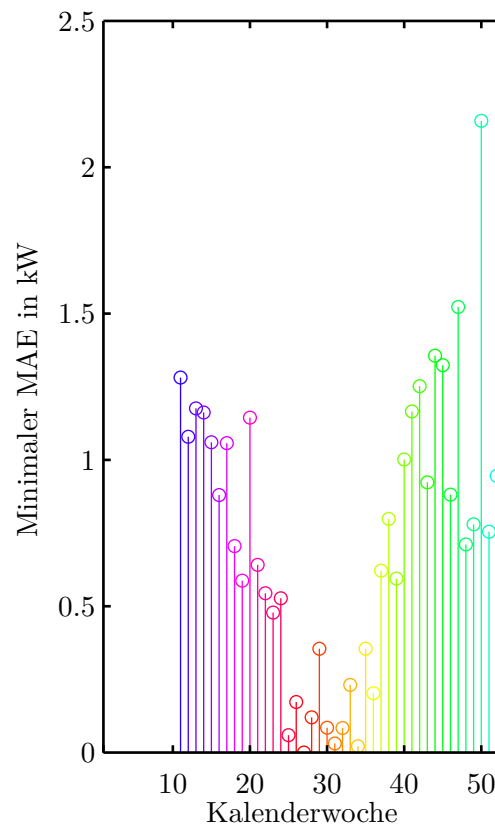


Abbildung 5.13.: VB1 – Minimum der einzelnen Kurvenverläufe aus Abbildung 5.12. Am Beginn des Jahres fehlen für diesen Verbraucher Daten.

Verteilung zu sehen ist. Dies ist der entscheidende Unterschied zu den Parametern α und β , die sehr stark variieren können, wie zum Beispiel aus Abbildung 5.1 abgeleitet werden kann.

5.4.3. Ergebnis

In der Folge wird zuerst auf die Bestimmung der Lastprognose mit dem Zusammenspiel von Prädiktor und Korrektor eingegangen. Dann wird, wie schon bei der Erweiterung des Prädiktors in Abschnitt 5.3, der Einfluss der Korrektorfunktionen $f_{K,i}$ auf die Prognosequalität untersucht. Ziel ist eine Korrektorfunktion zu finden, die für alle Verbraucher eine Verbesserung der Prognose im Sinne des MAE liefert. Danach wird die gefundene Prädiktor-Korrektor-Prognosemethode analog zu Abschnitt 5.2 mittels MPE und MAPE anhand aller betrachteten Verbraucher evaluiert. Hierbei sei an Abschnitt 2.4 erinnert,

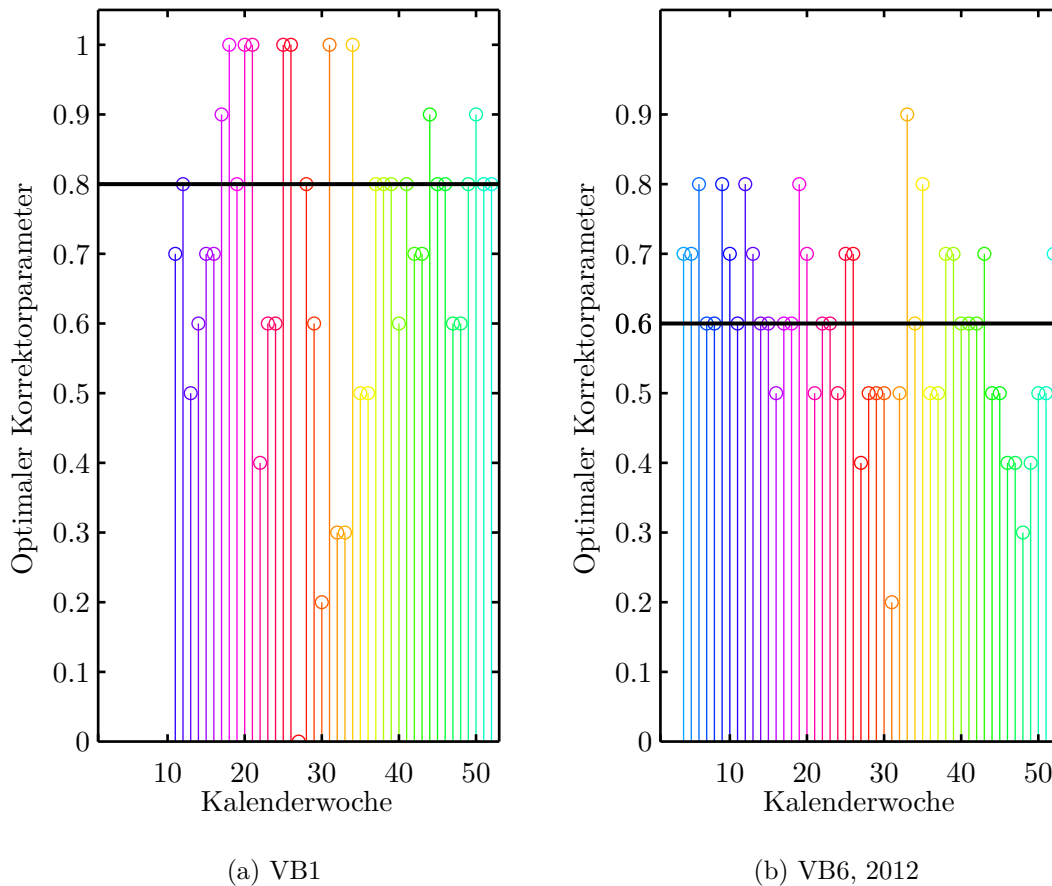


Abbildung 5.14.: Optimaler Korrektorparameter je Kalenderwoche für die Arbeitswoche inklusive Optimum des Verbrauchers (Schwarz).

wonach der MAE für die Auswahl und Entwicklung der Prognosemethode-Varianten, der MPE zur Erkennung systematischer Fehler und der MAPE zur anschaulichen Darstellung der Ergebnisse verwendet wird.

Bei der Bestimmung der optimalen Korrektorparameter k_i^* wurde bereits eine Lastprognose verwendet, aber noch nicht näher erklärt, was nun nachgeholt wird.

Sind die 24 Parametersätze des Prädiktors bestimmt, so kann für einen beliebig langen Prädiktionshorizont prädiziert werden. In den bisherigen Abschnitten entsprach dieser Prädiktionshorizont einer gesamten Arbeitswoche bzw. eines gesamten Wochenendes, was jeweils mit Kalenderwoche bezeichnet wurde. Je kürzer der Prädiktionshorizont desto besser die Prädiktion, wobei ein Tag die untere Grenze darstellt. Da für jeden Tag 24 unterschiedliche Parametersätze je Stunde verwendet werden, verbessert ein Prädiktionshorizont von 12h die Prädiktion nicht weiter; es werden einfach nur 12 der 24

5. Prognosemethode

Parametersätze verwendet.

Der erste Korrektor korrigiert gemäß (5.13) nur die nächste Stunde $n_A + 1$ und lässt alle zukünftigen Stunden unverändert, weshalb von einem Korrektorhorizont $h_K = 1$ h gesprochen werden kann. Für die Evaluierung dieses ersten Korrektors ist somit nur die Lastprognose für $n_A + 1$ relevant, da alle zukünftigen Prognosewert allein von der Prädiktion abhängen.

Die Erzeugung einer zusammenhängenden Lastprognose wird nun anhand einer Arbeitswoche erklärt. Zuerst wird durch tägliche Aktualisierung der Parametersätze die bestmögliche Prädiktion erzeugt. Dann werden stündlich Einzelprognosen mit einem Prognosehorizont von einer Stunde mittels Korrektor korrigiert. Somit ist der Prognosehorizont gleich dem Korrektorhorizont. Nun können diese $5 \cdot 24 = 120$ Einzelprognosen durch Aneinanderreihung zu einer zusammenhängenden Lastprognose für die gesamte Arbeitswoche verbunden werden, welche gut evaluiert werden kann und ausschließlich die Qualität der Korrektur bewertet. Es sei aber noch einmal darauf hingewiesen, dass jeder Wert dieser Prognose erst eine Stunde davor zur Verfügung steht.

Bestimmung der besten Korrektorfunktion

In der Folge wird die beste der hier betrachteten Korrektorfunktion bestimmt, wobei die Verbraucher wieder getrennt betrachtet und die Jahreszeiten je Verbraucher getrennt evaluiert werden. Da die Erweiterungen des Prädiktors f_{ij} keine Verbesserung brachten (siehe Abschnitt 5.3), ist der erste Prädiktor f_{00} die Referenz dieser Evaluierung. Bei f_{00} wurde immer zu Beginn einer Arbeitswoche bzw. eines Wochenendes die Parametersätze bestimmt und damit die gesamte Arbeitswoche bzw. das gesamte Wochenende prädiziert. Nun wird aber die bestmögliche Prädiktion bestimmt und mit $f_{K,0}$ bezeichnet, womit auch der Unterschied in der Prädiktion bewertet werden kann.

Die Prognose jeder Kalenderwoche wird mittels MAE evaluiert und in weiterer Folge werden die zugehörigen MAE-Werte über die jeweilige Jahreszeit gemittelt. Das Ergebnis dieser Evaluierung ist für die Arbeitswoche in Abbildung 5.15 und 5.16 bzw. für das Wochenende in Abbildung 5.17 und 5.18 zu sehen. Die für die betrachteten Korrektorfunktionen erhaltenen MAE-Werte der Arbeitswoche, haben mit Ausnahme weniger Sommer folgendes Verhältnis zueinander: $\text{MAE}(f_{K,1}) < \text{MAE}(f_{K,2}) < \text{MAE}(f_{K,0}) < \text{MAE}(f_{00})$. Somit ist die Korrektorfunktion $f_{K,1}$ optimal im Sinne des MAE für alle Verbraucher und Jahreszeiten. Für das Wochenende gilt immer $\text{MAE}(f_{K,1}) < \text{MAE}(f_{K,2}) < \text{MAE}(f_{K,0})$; jedoch ist für manche Verbraucher und Jahreszeiten $\text{MAE}(f_{00}) < \text{MAE}(f_{K,0})$. In diesen Fällen bringt die erneute Berechnung der Parametersätze keine Verbesserung, was auf den Unterschied zwischen Samstag und Sonntag zurückzuführen ist.

Da die Korrektorfunktion $f_{K,1}$ für alle Verbraucher und Jahreszeiten die beste im Sinne des MAE ist, wird sie für die finale Prognosemethode ausgewählt.

Man beachte die Analogie des somit gewählten „ersten Korrektors“ $f_{K,1}$ zum asymptotischen Beobachter (siehe z.B. [40, S. 444]). Für ein lineares, zeitdiskretes SISO-System

$$\mathbf{x}[n + 1] = \mathbf{A}\mathbf{x}[n] + \mathbf{b}u[n], \quad y[n] = \mathbf{c}^\top \mathbf{x}[n], \quad (5.17)$$

wird der asymptotische Beobachter angesetzt als

$$\hat{\mathbf{x}}[n+1] = \underbrace{\mathbf{A}\hat{\mathbf{x}}[n] + \mathbf{b}u[n]}_{\text{Prädiktion}} + \underbrace{\hat{\mathbf{b}}(y[n] - \hat{y}[n])}_{\text{Korrektur}}. \quad (5.18)$$

Bezeichnet man die Prädiktion als $\hat{\mathbf{x}}_P$ und die Abweichung der Ausgänge als $e_P = y - \hat{y}$, so kann dies geschrieben werden als

$$\hat{\mathbf{x}}[n+1] = \hat{\mathbf{x}}_P[n+1] + \hat{\mathbf{b}}e_P[n]. \quad (5.19)$$

Nun ist die Analogie zum ersten Korrektor

$$\hat{Q}_K[n_A+1] = \hat{Q}_P[n_A+1] + k_1e_P[n_A], \quad (5.20)$$

deutlich sichtbar. Da es sich bei $\hat{Q}_P[n_A+1]$ um einen statischen Zusammenhang und kein dynamisches System (5.17) handelt, können die Methoden des asymptotischen Beobachters allerdings nicht angewandt werden.

5. Prognosemethode

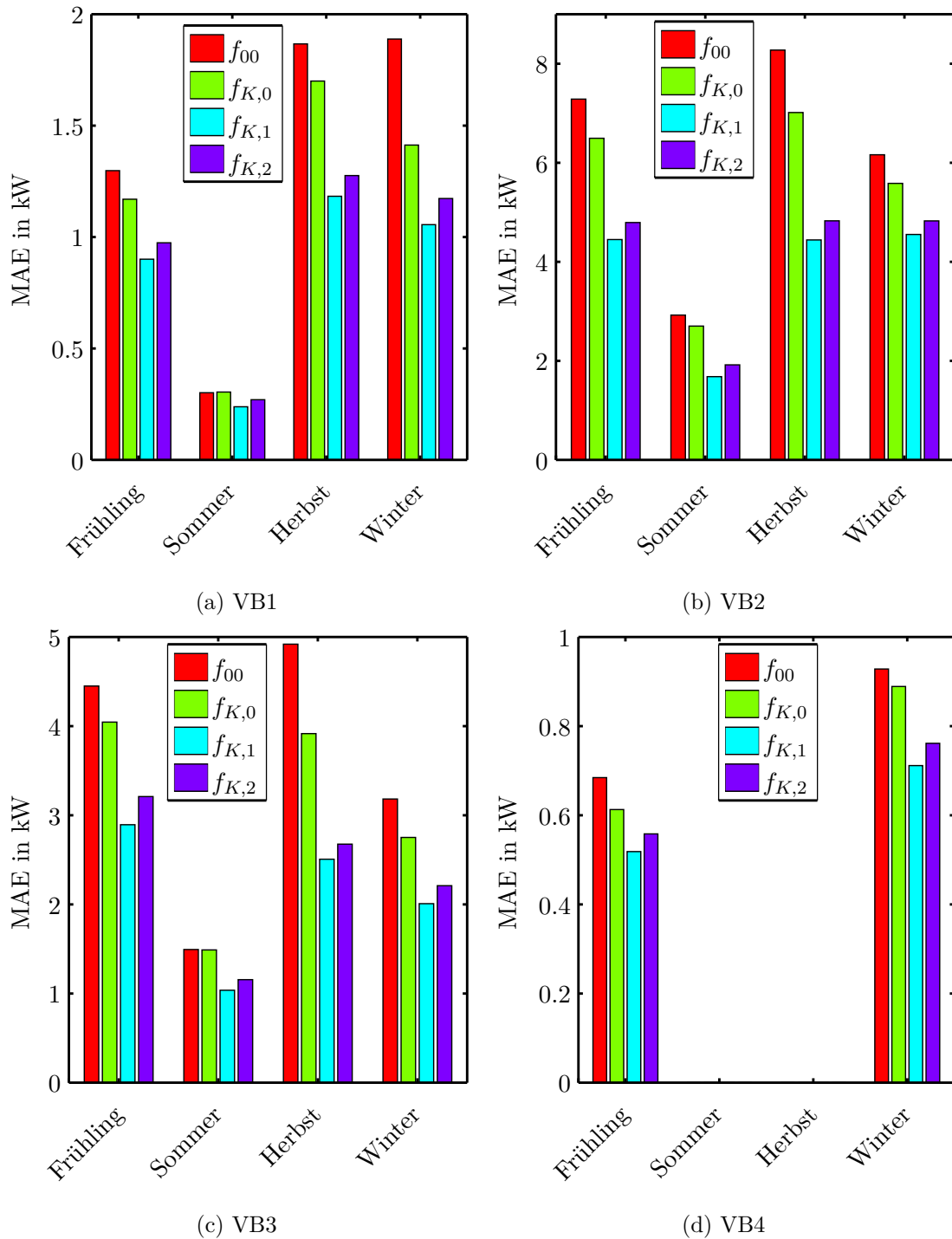


Abbildung 5.15.: Evaluierung des Prädiktors und der Korrektorfunktionen für die *Arbeitswoche*. Der erste Prädiktor ohne Erweiterungen wird gemäß (5.8) als f_{00} bezeichnet, wobei hier die Parametersätze einmalig am Beginn des jeweiligen Prognosehorizonts bestimmt wurden. Diese reine Prädiktion mit täglich aktualisierten Parametersätzen wird als $f_{K,0}$ bezeichnet. Die Korrektorfunktionen $f_{K,1}$ und $f_{K,2}$ gemäß (5.15) entsprechen nun tatsächliche Anwendungen des Korrektors.

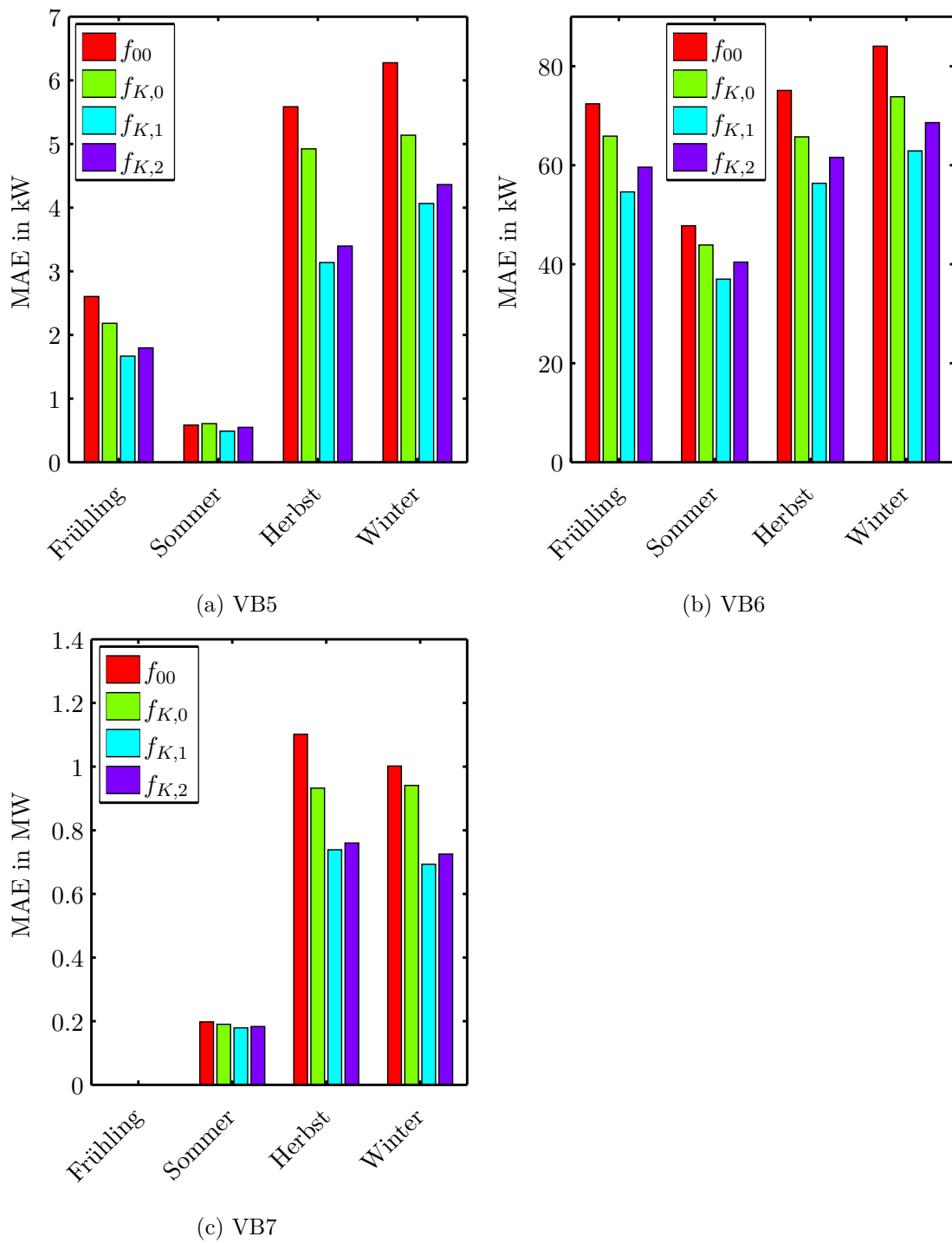


Abbildung 5.16.: Evaluierung des Prädiktors und der Korrektorfunktionen für die *Arbeitswoche*. Für eine Erklärung der Legende siehe Abbildung 5.15.

5. Prognosemethode

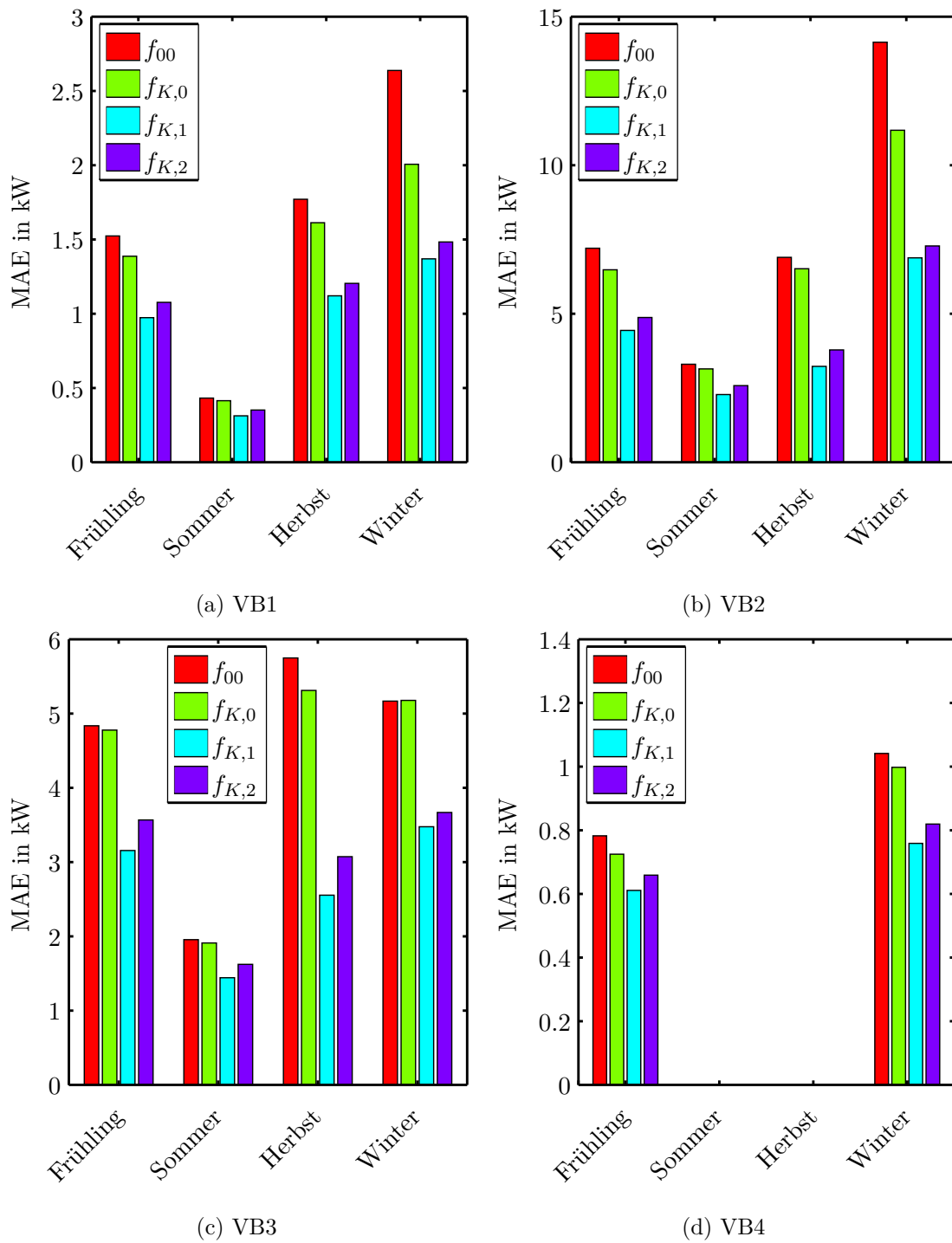


Abbildung 5.17.: Evaluierung des Prädiktors und der Korrektorfunktionen für das *Wochenende*. Für eine Erklärung der Legende siehe Abbildung 5.15.

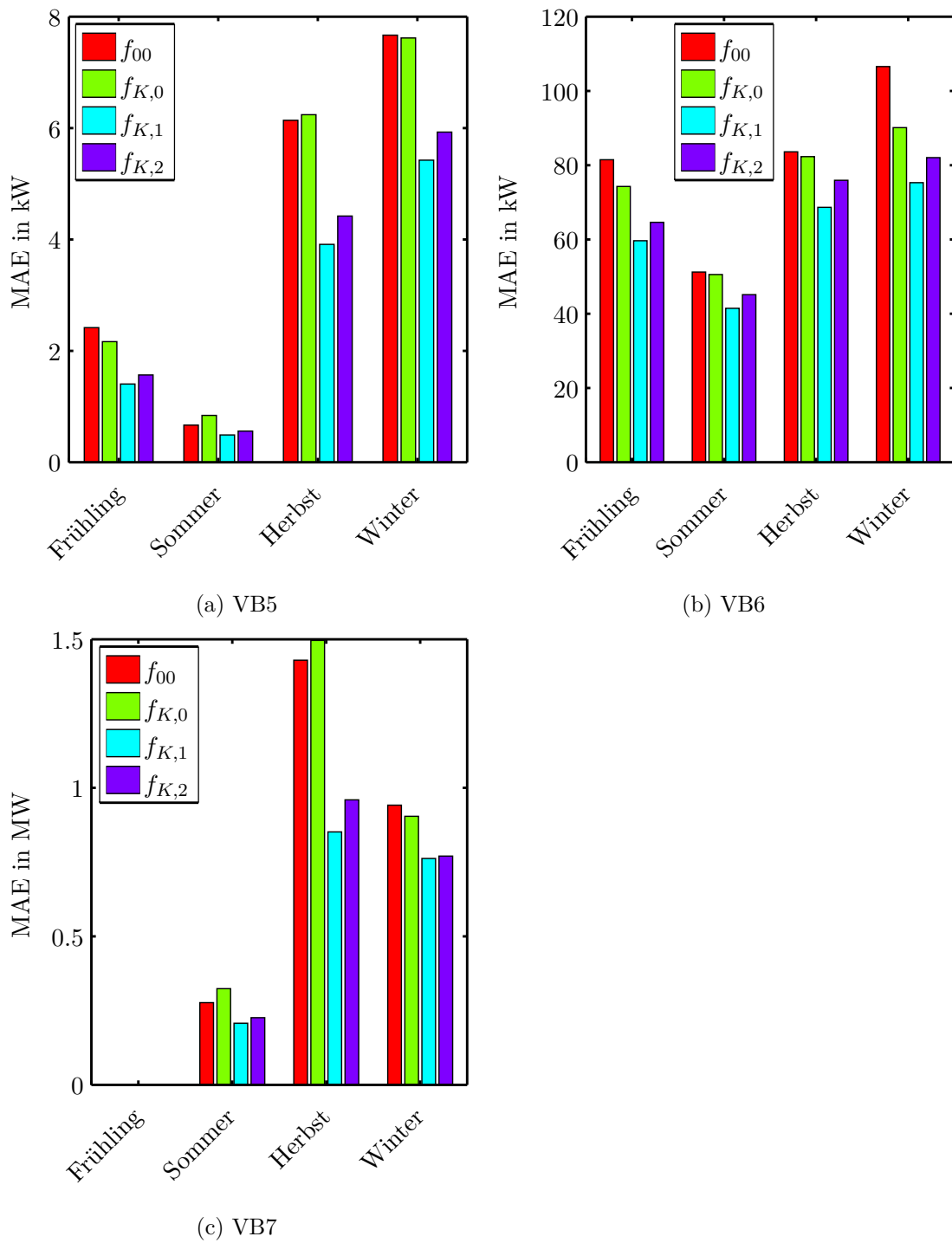


Abbildung 5.18.: Evaluierung des Pr diktors und der Korrektorfunktionen f r das *Wochenende*. F r eine Erkl rung der Legende siehe Abbildung 5.15.

Ergebnis des ersten Korrektors

Nachdem $f_{K,1}$ als die beste der betrachteten Korrektofunktion ermittelt wurde, wird diese analog zu Abschnitt 5.2 mittels MPE und MAPE bewertet. Diese Bewertung soll die Qualität der Prognosemethode beschreiben und mit anderen Quellen vergleichbar machen. An dieser Stelle sei nochmal darauf hingewiesen, dass die hier präsentierten Ergebnisse einer Aneinanderreihung von Einzelprognosen mit einem Prognosehorizont von 1 h entsprechen. Zuerst wird anhand einer Beispielwoche die Qualität veranschaulicht und anschließend wird die Methode für alle Verbraucher evaluiert.

Als Beispielwoche wird erneut die Arbeitswoche im Frühling von VB1 (siehe Abbildung 5.5) betrachtet, um die Prognoseergebnisse für reine Prädiktion und korrigierte Einzelprognosen vergleichen zu können. Zum Unterschied von Abbildung 5.5 werden nun die Parametersätze täglich neu bestimmt, die Prognose wird korrigiert und die Expertenregel 2 wird sowohl auf Prädiktor als auch Korrektor angewandt.

Wieder wird die reale (gemessene), statt einer prognostizierten Außentemperatur verwendet, um Einflüsse der Wetterprognose zu eliminieren. Das Ergebnis der Prädiktor-Korrektor-Prognosemethode ist in Abbildung 5.19 mit dem realen Wärmestrom verglichen.

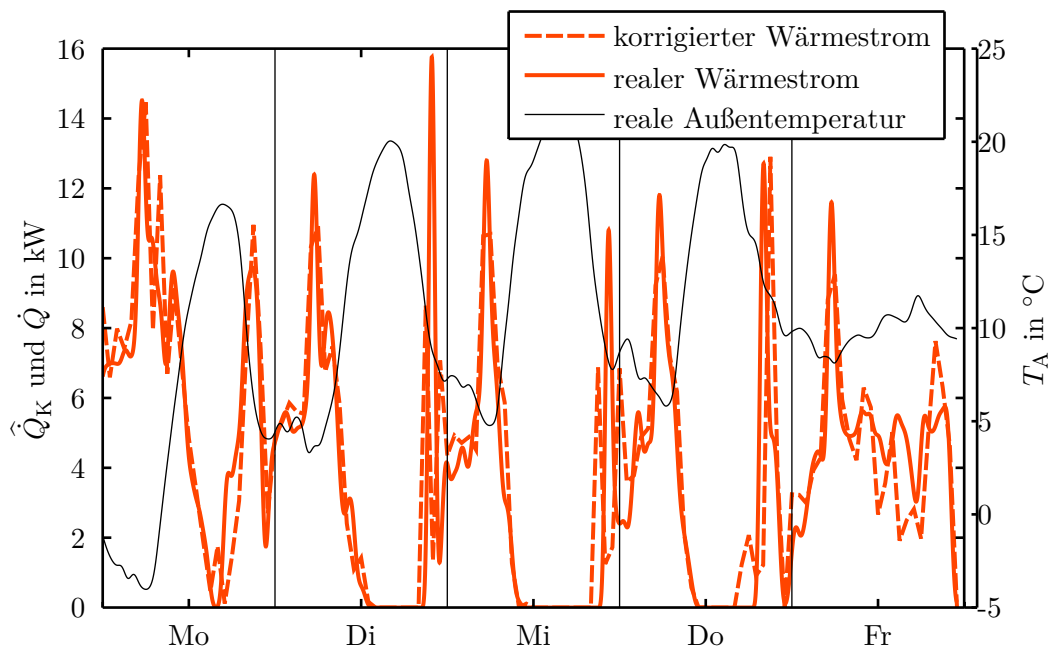


Abbildung 5.19.: VB2 – Vergleich zwischen korrigiertem und realem Wärmestrom für eine Arbeitswoche. Die dazugehörige reale Außentemperatur dient als Eingangsgröße der Prädiktor-Korrektor-Prognosemethode und ist daher ebenfalls eingezeichnet.

Vergleicht man Abbildung 5.5 mit Abbildung 5.19 so hat sich der Verlauf verbessert, was in Zahlen einer Reduktion des MAE von 1,63 kW auf 1,14 kW, des MPE von $-26,4\%$ auf $-11,2\%$, und des MAPE von $40,9\%$ auf $33,8\%$ entspricht. Eine deutliche Verbesserung ist beim MPE zu erkennen, da der Korrektor zum Beispiel eine zu kleine Prognose zur Stunde n_A in der darauffolgenden Stunde $n_A + 1$ durch eine Addition einer positiven Zahl korrigiert.

Der Verlauf in Abbildung 5.19 ist durch seine hochfrequenten Anteile denkbar ungünstig für die Anwendung des Korrektors. So muss sich eine Abweichung zur Stunde n_A in der nächsten Stunde $n_A + 1$ wiederholen, damit der Korrektor das Ergebnis deutlich verbessert. Trotzdem ist die Verbesserung durch den Korrektor sichtbar, wobei hier auch bereits die täglich aktualisierte Prädiktion hineinspielt.

Nach diesem Beispielverlauf werden nun alle Verbraucher für die vier Jahreszeiten evaluiert. Das Ergebnis ist analog zum ersten Prädiktor in Abschnitt 5.2 aufbereitet und die Diskussion bezieht sich auf den Vergleich dieses ersten Prädiktors mit diesem ersten Korrektor. Das Ergebnis ist wieder in Arbeitswoche und Wochenende aufgeteilt, wobei es für die Arbeitswoche in Tabelle 5.4 und Abbildung 5.20, und für das Wochenende in Tabelle 5.5 und Abbildung 5.21 zu sehen ist.

Vergleicht man die Arbeitswoche (Tabelle 5.4 mit 5.2) so fällt auf, dass sich die Prognose für den Sommer von VB1 verschlechtert hat. Ansonsten hat sich die Prognose für alle Verbraucher in allen Jahreszeiten ca. um ein Drittel verbessert.

Vergleicht man das Wochenende (Tabelle 5.5 mit 5.3), so ist ein ähnlicher Effekt zu sehen. So verschlechtert sich der MPE bei VB1 in Sommer und Herbst, und bei VB6 im Sommer. Der MAPE, welcher die Qualität des Prognose besser beschreibt, verbessert sich für alle Verbraucher.

Die Interpretation von MPE und MAPE bleibt unverändert, wobei die negativen MPE-Werte wieder durch Expertenregel 2 und durch die Saison erklärt werden können. Die teilweise Verschlechterungen im Sommer sind kaum relevant, da der Pufferspeicher immer für die hohe Last im Winter ausgelegt wird. Somit können Prognosefehler im Sommer durch den Pufferspeicher in der Regel problemlos ausgeglichen werden.

Dieses Ergebnis zeigt das hohe Potential welches hinter diesem Prädiktor-Korrektor-Ansatz steckt, wobei diese Prognosen aus einstündigen Einzelprognosen zusammengesetzt wurden.

5. Prognosemethode

Bez.	MPE in %				MAPE in %			
	Frühling	Sommer	Herbst	Winter	Frühling	Sommer	Herbst	Winter
VB1	-26,2	-163,9	-5,6	-2,1	52,4	239,0	24,2	14,3
VB2	-4,7	-3,9	0,6	-0,3	16,8	15,3	7,4	6,3
VB3	-11,5	-15,0	0,8	-1,3	32,5	74,3	11,5	5,8
VB4	-57,0	–	–	-3,5	92,4	–	–	14,9
VB5	-151,3	-140,9	-18,7	-1,5	189,9	180,5	49,2	20,4
VB6	-8,9	-16,3	-0,5	-1,4	28,1	43,8	15,4	8,5
VB7	–	-1,1	0,9	-0,8	–	9,5	4,1	4,5

Tabelle 5.4.: Evaluierung des ersten Korrektors für die *Arbeitswoche*. Bei den Feldern die mit „–“ gekennzeichnet sind, stehen keine Daten zur Verfügung.

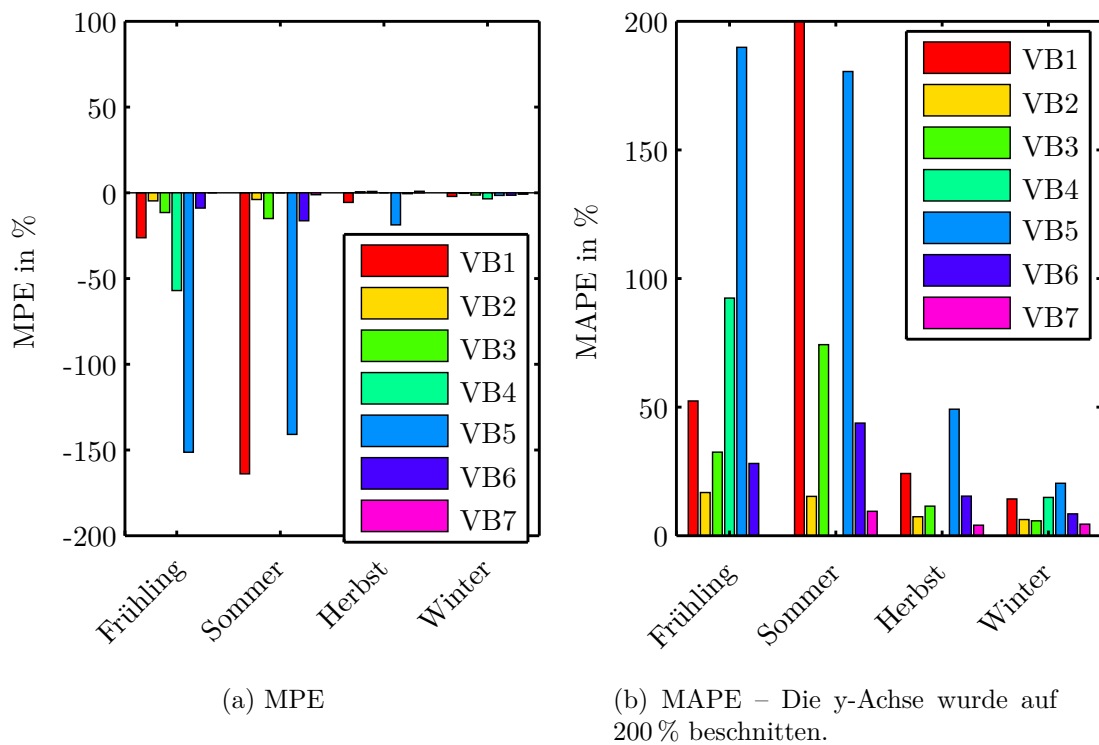


Abbildung 5.20.: Evaluierung des ersten Korrektors für die *Arbeitswoche*. Die exakten Werte sind Tabelle 5.4 zu entnehmen.

Bez.	MPE in %				MAPE in %			
	Frühling	Sommer	Herbst	Winter	Frühling	Sommer	Herbst	Winter
VB1	-37,4	-27,9	-6,1	-14,3	65,2	103,6	22,8	25,3
VB2	-7,1	-6,9	0,8	-14,2	22,2	22,1	6,1	24,2
VB3	-16,2	4,8	1,4	-4,3	39,5	56,0	10,6	13,3
VB4	-81,8	–	–	-3,9	118,2	–	–	15,8
VB5	-157,6	-126,7	-12,7	1,4	192,4	160,3	44,1	23,7
VB6	-10,9	-16,9	-1,3	-1,6	32,2	47,1	18,3	10,8
VB7	–	1,3	0,8	0,2	–	10,9	5,0	4,7

Tabelle 5.5.: Evaluierung des ersten Korrektors für das *Wochenende*. Bei den Feldern die mit „–“ gekennzeichnet sind, stehen keine Daten zur Verfügung.

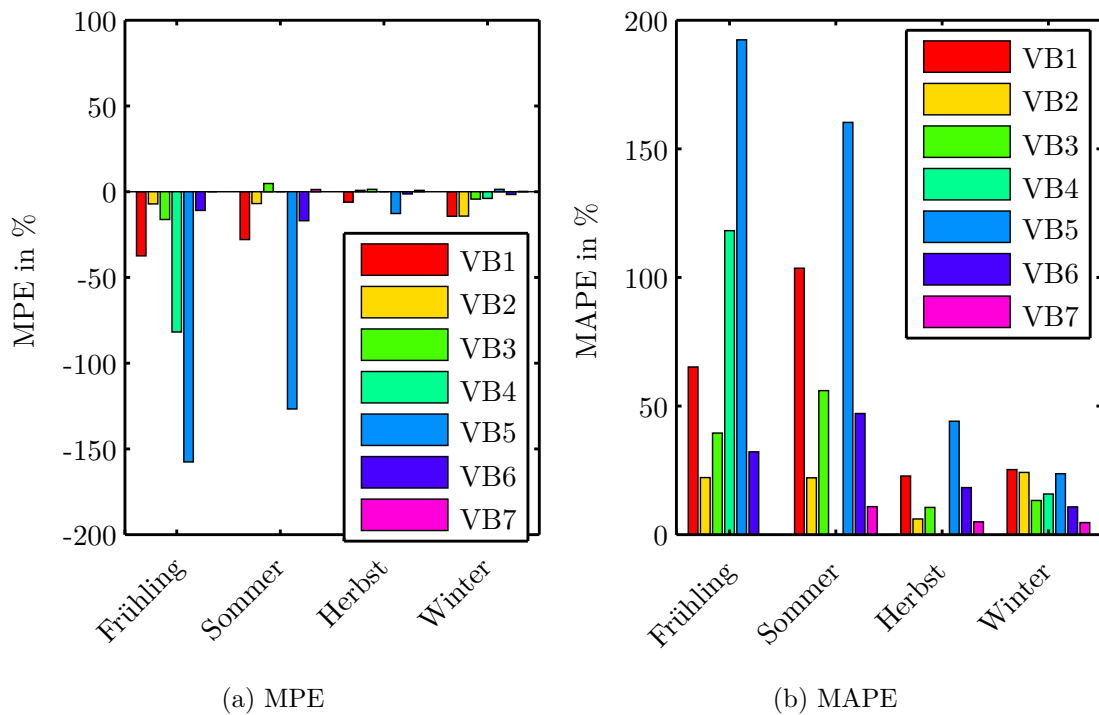


Abbildung 5.21.: Evaluierung des ersten Korrektors für das *Wochenende*. Die exakten Werte sind Tabelle 5.5 zu entnehmen.

5.5. Erweiterung des Korrektors

In diesem Abschnitt wird der „erste Korrektor“ $f_{K,1}$ auf eine praxistaugliche Prognosemethode erweitert. Bis jetzt wurde der Prädiktionsfehler der aktuellen Stunde n_A verwendet, um gemäß (5.13) die Prädiktion der nächsten Stunde $n_A + 1$ zu korrigieren. Bei einer Prognose zur Laufzeit würde somit nur die nächste Stunde korrigiert und für alle nachfolgenden Stunden würde die reine Prädiktion verwendet werden. Diese Methode wäre möglich, jedoch sollte diese Korrektur auf die Zukunft ausgeweitet werden, um die Prognose zu verbessern und eine Quasistetigkeit zu gewährleisten. Da es sich hier um diskrete Verläufe handelt, wird von Quasistetigkeit statt Stetigkeit gesprochen. Diese Ausdehnung des Korrektors auf die Zukunft wird in diesem Abschnitt behandelt.

Für die weiteren Betrachtungen wird der bereits erwähnte Korrektorhorizont h_K exakt definiert, der die zeitliche Ausdehnung der Korrektur beschreibt. So entspricht der erste Korrektor einem Korrektorhorizont $h_K = 1$ h, da nur die Prädiktion zur nächsten Stunde $n_A + 1$ korrigiert wird. Bei einem $h_K = m$ h, wird die Korrektur auf weitere Stunden bis $n_A + m$ ausgedehnt.

Im vorherigen Abschnitt 5.4 wurde gezeigt, dass, basierend auf dem Prädiktionsfehler zur aktuellen Stunde n_A , eine Korrektur für die Stunde $n_A + 1$ ermittelt werden kann, wodurch die Prognose für die Stunde $n_A + 1$ verbessert wird. Es kann daraus allerdings nicht geschlossen werden, dass diese Korrektur für alle zukünftigen Stunden des Prognosehorizonts (oder Prädiktorhorizonts) auch eine Verbesserung liefert. Daher sollte diese Ausdehnung so erfolgen, dass die Korrektur am Ende des Korrektorhorizonts verschwindet, zu Null wird. Diese Korrektur der Prädiktion für die Stunde $n_A + 1$ und deren Ausdehnung auf die nahe Zukunft wird in einem Beispiel veranschaulicht.

Beispiel 5.1. In diesem Beispiel wird eine fiktive Prädiktion durch eine Korrektur mit einem Korrektorhorizont $h_K = 1$ h bzw. mit $h_K = 7$ h verbessert. Dies ist in Abbildung 5.22 für beide Korrektorhorizonte dargestellt.

In diesem Beispiel ist die aktuelle Stunde $n_A = 3$, d.h. die Prognose wird in der Zeit von 03:00 ($n = 3$) bis 04:00 ($n = 4$) erstellt. Die dabei erzielte Prädiktion entspricht dem verschobenen Verlauf des realen Wärmestroms. Dieser konstante Prädiktionsfehler wird in beiden Abbildungen für $n_A = 3$ mittels erstem Korrektor korrigiert. In Abbildung 5.22a wird diese Korrektur nicht weiter ausgedehnt, womit der Prognosefehler für $n > n_A + 1$ gleich groß bleibt und die Prognose einen starken Sprung aufweist. Man beachte, dass hier die Stunden 0 bis 3 der Vergangenheit angehören und die Prognose erst mit Stunde 4 beginnt. In Abbildung 5.22b wird die Korrektur auf die nächsten 6 Stunden ausgedehnt, was einem $h_K = 6$ entspricht. Die Korrektur der nahen Zukunft wird hier durch lineares Abklingen des Korrekturwerts $f_{K,1}$ realisiert.

Anhand dieses Beispiels wurde veranschaulicht, dass eine Ausdehnung der Korrektur auf die nahe Zukunft das Prognoseergebnis verbessern kann. Im Beispiel wurde von einem konstanten Prädiktionsfehler für alle Stunden ausgegangen, womit die Korrektur

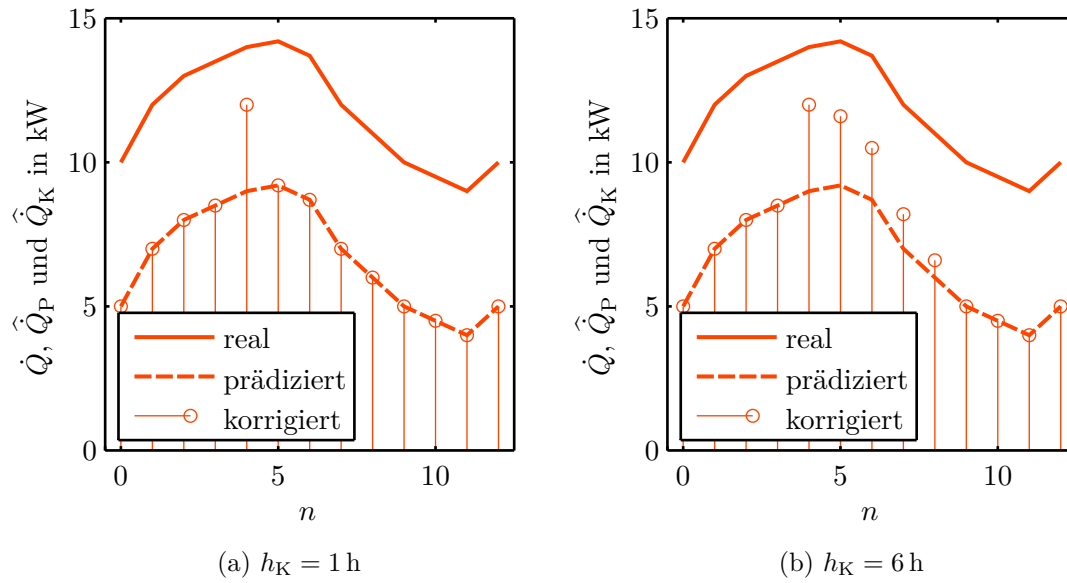


Abbildung 5.22.: Beispiel 5.1 – Fiktive Wärmeströme mit unterschiedlichem Korrektorhorizont h_K .

hier über den Korrektorhorizont nicht abklingen müsste. Da solch ein konstanter Prädiktionsfehler aber eine Ausnahme darstellt, muss die Korrektur innerhalb eines noch zu bestimmenden Korrektorhorizont abklingen.

5.5.1. Ansatz

Als Grundlage dieses Ansatzes dient die Korrektorfunktion $f_{K,1}$ (5.15) mit dem optimalen Korrektorparameter k_1^* (5.16), womit die nächste Stunde $n_A + 1$ korrigiert wird. Diese Korrektur wird nun durch lineares Abklingen auf den Korrektorhorizont h_K erweitert:

$$\hat{Q}_K[n] = \hat{Q}_P[n] + f_{K,h}[n], \quad (5.21)$$

mit

$$f_{K,h}[n] = \begin{cases} f_{K,1}[n_A] - \frac{f_{K,1}[n_A]}{h_K - 1}(n - n_A - 1) & n_A < n \leq n_A + h_K, \\ 0 & \text{sonst.} \end{cases} \quad (5.22)$$

Diese Funktion ist für Beispiel 5.1 in Abbildung 5.23 dargestellt, wobei $f_{K,1} = 3 \text{ kW}$, $n_A = 3$ und $h_K = 6 \text{ h}$ sind.

5.5.2. Parameterbestimmung

Der Parameter h_K soll wieder für alle Verbraucher und Kalenderwochen bestimmt werden, wobei analog zu der Bestimmung der Anzahl an Trainingstagen bzw. des Korrektorparameters vorgegangen werden soll.

5. Prognosemethode

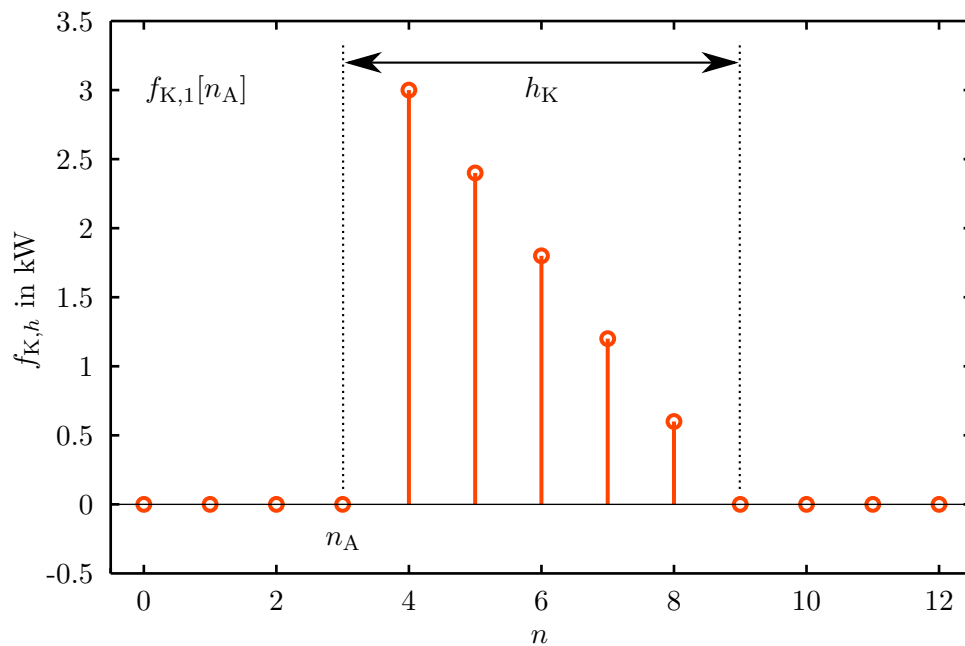


Abbildung 5.23.: Beispiel 5.1 – Korrekturfunktion (5.22) zu den fiktiven Wärmeströmen.

Jedoch kann für den erweiterten Korrektor keine abgeschlossene Prognose einer Kalenderwoche (getrennt in Arbeitswoche und Wochenende) erzeugt werden. So wurde bei der Anzahl an Trainingstagen für die Prädiktion mit einem fixen Parametersatz die gesamte Kalenderwoche prädiziert. Beim Korrektorparameter wurde immer eine einzelne Stunde korrigiert, wodurch diese stündlichen Einzelprognosen zu einer Prognose für die Kalenderwoche zusammengefasst werden konnten. Jetzt wird durch den Korrektorhorizont eine Einzelprognose erzeugt, die länger als eine Stunde dauert. Da diese Einzelprognosen nicht einfach aneinandergereiht werden können, wird für jede Stunde der Kalenderwoche eine eigene Prognose bestimmt und mittels MAE bewertet. Diese einzelnen MAE-Werte werden gemittelt, womit wieder ein Qualitätsmaß für diese Kalenderwoche für die Optimierung des Parameters h_K zur Verfügung steht.

Nun soll wieder anhand der bereits bekannten drei Schritte von Abschnitt 5.2 vorgegangen werden:

Schritt 1: Optima je Kalenderwoche eines Verbrauchers: $h_{K,KW}^*$.

Schritt 2: Optima der Verbraucher: $h_{K,VB}^*$.

Schritt 3: Optimum aller Verbraucher: h_K^* .

Im ersten Schritt wird das Optimum für jede Kalenderwoche jedes Verbrauchers einzeln bestimmt. Hierfür wird der Korrektorhorizont von 1 bis 24 h erhöht, wobei 1 h dem ersten Korrektor entspricht. Die Obergrenze von 24 h wurde empirisch ermittelt und

ermöglicht eine klare Bewertungsmethode. Wie bereits beschrieben, werden Einzelprognosen für jede Stunde der Kalenderwoche bestimmt und zunächst einzeln mittels MAE bewertet. Nun besteht die Kalenderwoche für die Arbeitswoche aus 5 Tagen und für das Wochenende aus 2 Tagen. Da der Korrektorhorizont maximal 24 h lang werden kann, reicht dieser zum Teil in den nächsten Tag hinein. Für die letzten Tage des Zeitraums (Freitag bzw. Sonntag) steht aber kein nächster Tag zu Verfügung, weshalb hier keine Prognose erzeugt bzw. bewertet werden kann. Daher werden nur Einzelprognosen für Montag bis Donnerstag bzw. Samstag bestimmt und bewertet.

Um die Bestimmung der Einzelprognosen numerisch zu beschleunigen, wird für die gesamte Kalenderwoche eine täglich aktualisierte Prädiktion bestimmt. Für diese täglich aktualisierte Prädiktion werden die Parametersätze des Prädiktors in jeder Nacht der Kalenderwoche neu bestimmt und daraus der nächste Tag prädiziert. Nun wird bei festgehaltenem Korrektorhorizont jede Stunde der Kalenderwoche ohne den letzten Tag (Freitag oder Sonntag) korrigiert und diese Einzelprognosen mittels MAE bewertet. Diese einzelnen MAE-Werte werden gemittelt, womit man einen MAE-Wert je Korrektorhorizontwert je Kalenderwoche und je Verbraucher erhält. Dieser Zusammenhang ist für VB1 für die Arbeitswoche in Abbildung 5.24 dargestellt.

Die Kurven in dieser Abbildung scheinen nahezu horizontal zu verlaufen, sind jedoch gekrümmt und können anhand der Methode erklärt werden. So wird zuerst die gesamte Kalenderwoche mit aktuellsten Parametersätzen prädiziert. Diese Prädiktion bleibt für diese Kalenderwoche konstant und liefert einen Basis-MAE der für alle Korrektorhorizonte gleich bleibt. Nun werden nur einzelne Stunden korrigiert, wodurch sich die Prognose nur geringfügig verändert. Die dazugehörigen, geringfügig unterschiedlichen MAE-Werte werden nun gemittelt, womit sich je Korrektorhorizont ein MAE ergibt, der sich kaum von dem der reinen Prädiktion unterscheidet. Somit ist der Offset (Basis-MAE) durch den Prädiktionsfehler der Kalenderwoche bei jedem MAE dabei, wodurch diese MAE-Werte in einen Wertebereich gehoben werden, in dem die geringfügigen Unterschiede mit freiem Auge schwer zu erkennen sind. Da dieser Offset aber für jede Kalenderwoche konstant ist, kann eine Minimum des MAE je Kalenderwoche bestimmt werden, dass eine direkte Aussage auf den Korrektorhorizont beinhaltet. Diese minimalen MAE sind für VB1 in Abbildung 5.25 dargestellt.

Nun kann der Korrektorhorizont bestimmt werden, welcher zum jeweiligen minimalen MAE gehört, womit ein Optimum für jede Kalenderwoche $h_{K,KW}^*$ gefunden wurde. Erneut werden das Wohnhaus (VB1) in Abbildung 5.26a und das Nahwärmenetz (VB6) in Abbildung 5.26b als Beispiele für diese einzelnen Optima betrachtet. Die Korrektorhorizonte $h_{K,KW}^*$ verteilen sich, laut diesen Abbildungen und den Ergebnissen für die restlichen Verbraucher, großteils auf die vorgegebenen Grenzen für $h_{K,KW}$. Dies ist problematisch, da somit ein einziger Korrektorhorizont für den jeweiligen Verbraucher $h_{K,VB}^*$ die tatsächlichen Optima nur schlecht wiedergibt. Da $h_{K,KW}^*$ aber auch zur Laufzeit nicht im Vorhinein bestimmt werden kann, wird hier trotzdem im nächsten Schritt ein Kompromiss gebildet.

Im zweiten Schritt wird aus diesen Optima je Kalenderwoche mittels Median ein Optimum des jeweiligen Verbrauchers $h_{K,VB}^*$ bestimmt. Es ist zu beachten, dass für einige Kalenderwochen 24 h optimal wären, womit die Obergrenze des Optimierungsspielraums

5. Prognosemethode

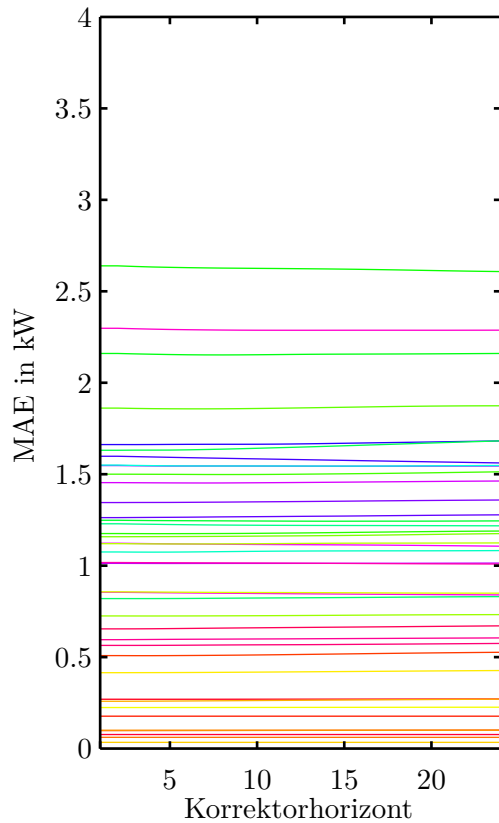


Abbildung 5.24.: VB1 – Qualität der Prognose (mittels MAE bewertet) in Abhängigkeit des Korrektorhorizonts für die Arbeitswoche.

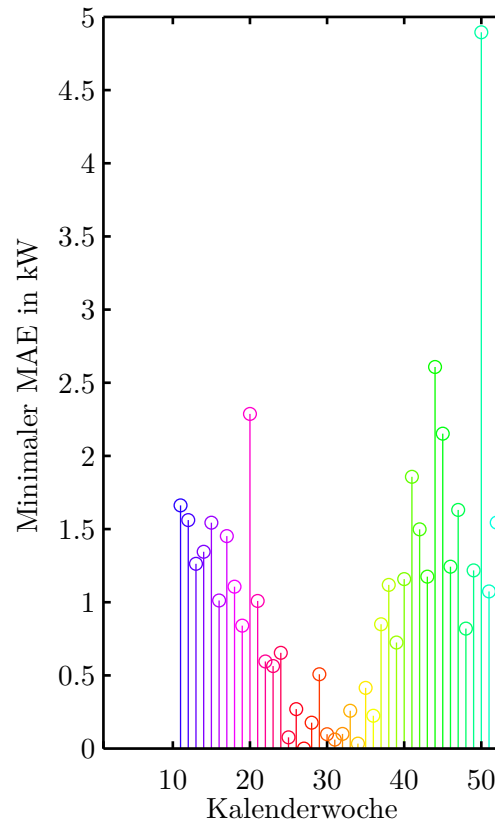


Abbildung 5.25.: VB1 – Minimum der einzelnen Kurvenverläufe aus Abbildung 5.24. Am Beginn des Jahres fehlen für diesen Verbraucher Daten.

erreicht ist. Somit müsste dieser Obergrenze erhöht werden, um das wahre Optimum der Kalenderwoche zu finden. Da aber ohnehin der Median aller Kalenderwochen gebildet wird, würde eine Erhöhung der Obergrenze das Optimum je Verbraucher nicht verändern.

Im dritten Schritt wird aus diesen sieben lokalen Optima je Verbraucher mittels Median ein Optimum aller Verbraucher h_K^* für die Arbeitswoche bzw. das Wochenende bestimmt. Hierbei ergibt sich $h_{K,MoFr}^* = 5,5$ h bzw. $h_{K,SaSo}^* = 7,5$ h, was bei der gewählten stündlichen Auflösung der Prognosemethode noch gerundet werden muss. Da die jeweilige Kommazahl durch den Mittelwert der beiden Pseudomedianwerte berechnet wird (vgl. Abschnitt 2.1.2), ist nicht eindeutig ob auf- oder abgerundet werden soll. Nachdem die Optima der Kalenderwochen auch zum Teil bei der Untergrenze von $h_K = 1$ h liegen (vgl. Abbildung 5.26a und 5.26b) wird abgerundet um eine Prädiktion nicht zu verschlechtern.

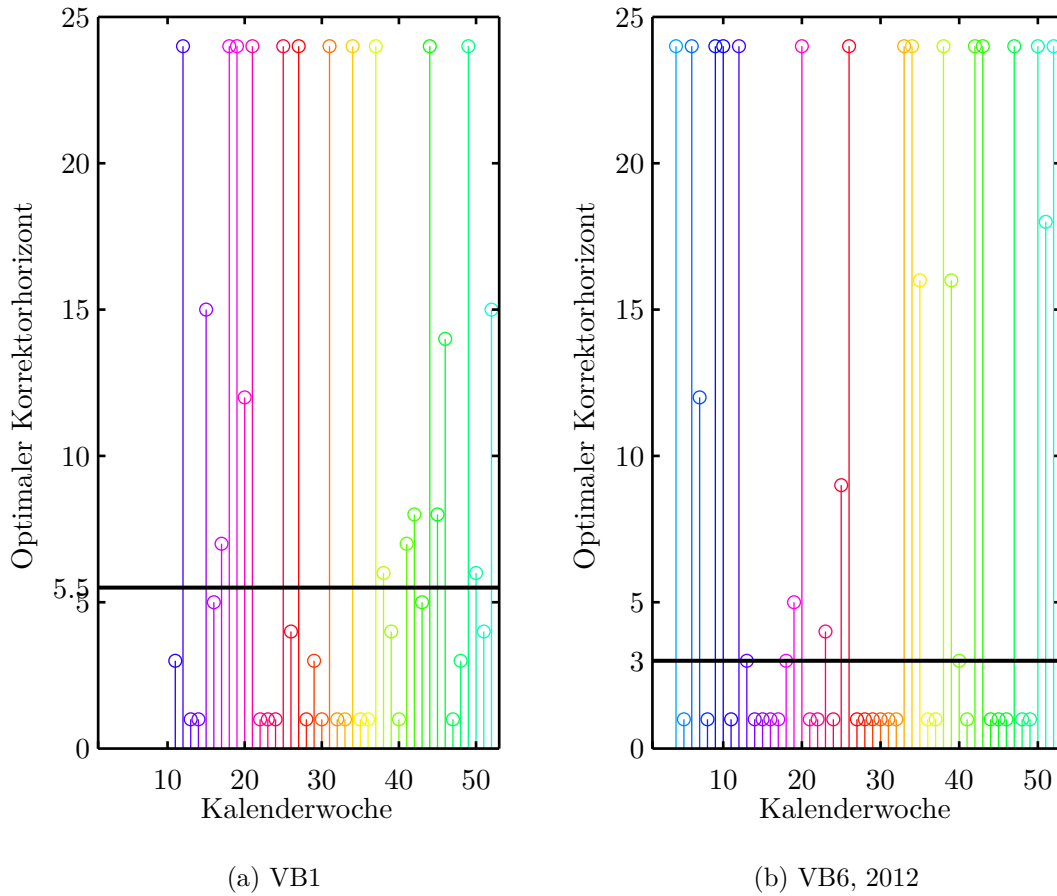


Abbildung 5.26.: Optimaler Korrektorparameter je Kalenderwoche für die Arbeitswoche inklusive Optimum des Verbrauchers (Schwarz).

Somit ergeben sich die Korrektorhorizonte für Arbeitswoche bzw. Wochenende als

$$h_{K,MoFr}^* = 5 \text{ h} \quad \text{bzw.} \quad h_{K,SaSo}^* = 7 \text{ h.} \quad (5.23)$$

Man hätte hier auch aufrunden können und dies mit dem Verbesserungspotential durch den Korrektor argumentieren können. Da die optimalen Korrektorhorizonte für die Kalenderwochen $h_{K,KW}^*$ stark gestreut sind, ist ein Korrektorhorizont für alle Kalenderwochen und alle Verbraucher in jedem Fall nur ein beschränkt aussagekräftiger Kompromiss. Somit könnte für die Implementierung auch der Mittelwert von (5.23) $h_K^* = 6 \text{ h}$ verwendet werden, da die beiden Zahlenwerte nahe bei einander liegen.

5.5.3. Ergebnis

Für diesen Abschnitt kann kein eindeutiges Ergebnis bestimmt werden, da dieses maßgeblich vom Prognosehorizont abhängt. Dies wird anhand einer Arbeitswoche erläutert und es werden obere und untere Grenzen für das Ergebnis genannt.

Beim reinen Prädiktor kann für jede Stunde eines jeden Tages des Jahres eine Prädiktion bestimmt werden. Hierbei ist nur zu beachten, dass die Parametersätze regelmäßig erneuert werden. Je aktueller die Parametersätze, desto besser die Prognose, wie in Abbildungen 5.15 bis 5.18 anhand von f_{00} und $f_{K,0}$ gezeigt wurde. Bei einer laufenden Aktualisierung der Parametersätze wie in $f_{K,0}$ gibt es für jede Stunde nur eine aktuellste Prädiktion, womit hier ein eindeutiges Ergebnis bestimmt und evaluiert werden kann.

Beim ersten Korrektor wird immer nur eine Stunde mittels des vergangenen Prädiktionsfehlers korrigiert. Diese einzelne Korrektur kann als Einzelprognose mit einem Prognosehorizont von einer Stunde betrachtet werden. Diese Einzelprognosen wurden je Kalenderwoche zusammengefasst, was zwar nicht der Prognose zur Laufzeit entspricht, aber einfach bewertet werden kann.

Beim erweiterten Korrektor wird die Korrektur auf die Zukunft entsprechend des optimalen Korrektorhorizonts h_K^* von (5.23) ausgedehnt. Somit können die Einzelprognosen nicht wie beim ersten Korrektor zusammengefasst werden, da sie sich überlappen würden. Dies wurde zwar für die Bestimmung h_K^* gemacht, jedoch nur um einen MAE je Kalenderwoche zu erhalten und nicht um Prognoseergebnisse zu präsentieren.

Nun könnte man analog zur Bestimmung von h_K^* Einzelprognosen generieren und diese mitteln, das Ergebnis würde aber maßgeblich von dem Prognosehorizont abhängen. Diese Abhängigkeit wird anhand von VB2 für eine Arbeitswoche im Frühling für einen langen und den kürzest möglichen Prognosehorizont erläutert.

In Abbildung 5.27 ist eine Prognose für die gesamte Arbeitswoche zu sehen, was einem langen Prognosehorizont von 5 Tagen entspricht. Die Prognose wird am Montag zwischen der aktuellen Stunde (n_A) 00:00 und der nächsten Stunde ($n_A + 1$) 01:00 erzeugt, wobei hier die Parametersätze des Prädiktors natürlich nur in dieser Zeit bestimmt und nicht weiter aktualisiert werden. Man sieht gut, wie der Korrektor mit dem Prädiktionsfehler um 00:00 die nächsten 5 Stunden korrigiert, was laut $h_{K,MoFr}^*$ von (5.23) dem optimalen Korrektorhorizont für die Arbeitswoche entspricht. Bei diesem Verbraucher wäre für diese Kalenderwoche ein längerer Korrektorhorizont besser, da großteils ein konstanter Prädiktionsfehler vorliegt. Diese sichtbare Verbesserung für den Korrektorhorizont ist jedoch aufgrund des langen Prognosehorizonts im MAE kaum sichtbar. So entspricht die reine Prädiktion einem MAE von 9,27 kW und der Korrektor verbessert nur auf einen MAE von 9,11 kW. Daher ist bei einem langen Prognosehorizont die Auswirkung des Korrektors auf den MAE vernachlässigbar.

In Abbildung 5.28 ist die gleiche Arbeitswoche, wie von Abbildung 5.27 zu sehen. Jedoch wird jetzt für den kürzest möglichen Prognosehorizont von einer Stunde prognostiziert, womit der Einfluss des Korrektors auf den MAE deutlich zunimmt. Nun entspricht die reine Prädiktion einem MAE von 12,06 kW und der Korrektor verbessert auf einen MAE von 4,59 kW.

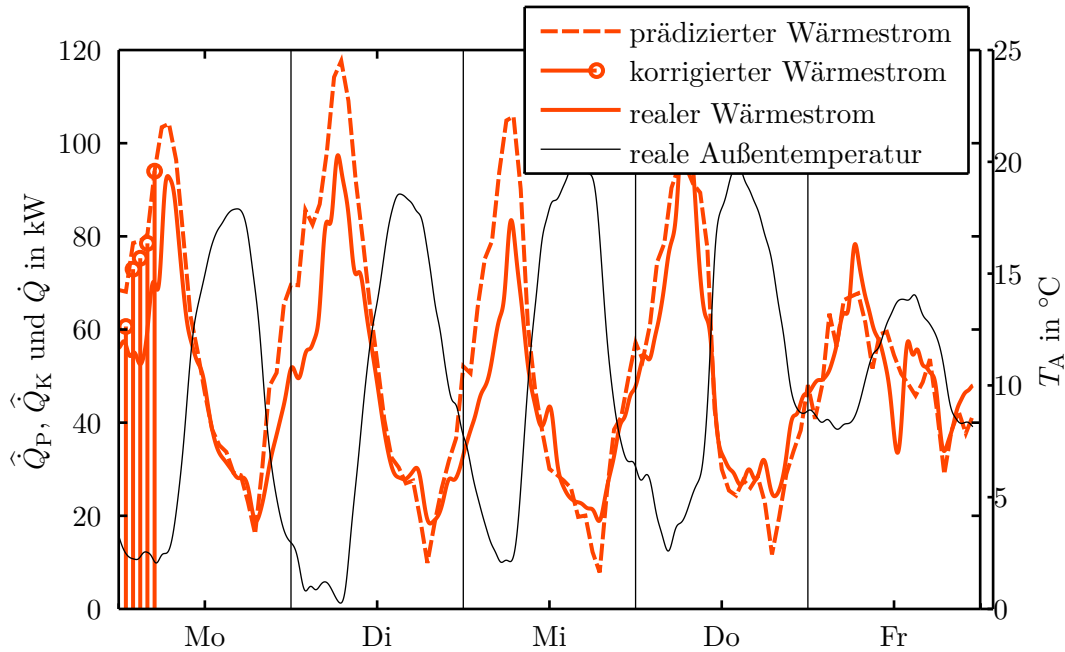


Abbildung 5.27.: VB2 – Prognose für einen langen Prognosehorizont; sie gilt von Montag 01:00 bis Freitag 23:00.

Anhand dieser beiden Extrembeispiele ist klar, dass der Einfluss des Korrektors auf das Prognoseergebnis maßgeblich vom Prognosehorizont abhängt. Da kein Prognosehorizont fixiert ist, kann der erweiterte Korrektor nicht aussagekräftig evaluiert werden. Zusätzlich ist für manche Anwendungen die Prognosequalität für die nahe Zukunft deutlich wichtiger, was eine aussagekräftige Evaluierung weiter erschwert.

Um aber eine Abschätzung der Qualität der Prädiktor-Korrektor-Prognosemethode mit erweitertem Korrektor machen zu können, wird diese Qualität nach oben und unten begrenzt. Wird nun zum einen der Prognosehorizont lang gewählt, so ist die Auswirkung des Korrektors auf die Qualität der Gesamtprognose vernachlässigbar. Wird zum anderen der Prognosehorizont sehr kurz gewählt, so hat der Korrektor einen massiven Einfluss. Beim Extrem des besonders langen Prognosehorizonts kann der Korrektor vernachlässigt werden, womit es sich um eine reine Prädiktion handelt. Diese reine Prädiktion wurde in Abschnitt 5.2 bereits für den Prognosehorizont von 5 Tagen für die Arbeitswoche und 2 Tagen für das Wochenende evaluiert. Beim Extrem des kürzest möglichen Prognosehorizonts von einer Stunde handelt es sich aber um den ersten Korrektor der in Abschnitt 5.4 evaluiert wurde. Das Ergebnis des erweiterten Korrektors liegt zwischen diesen beiden Extrema, womit die Qualität der Prognose nach oben und unten abgeschätzt werden kann. Hierbei entspricht das bestmögliche Ergebnis dem des ersten Korrektors, aber das schlechteste mögliche Ergebnis kann nur für die maximal 5 Tage der reinen Prädiktion abgeschätzt werden, da keine obere Schranke für den Prognosehorizont existiert.

5. Prognosemethode

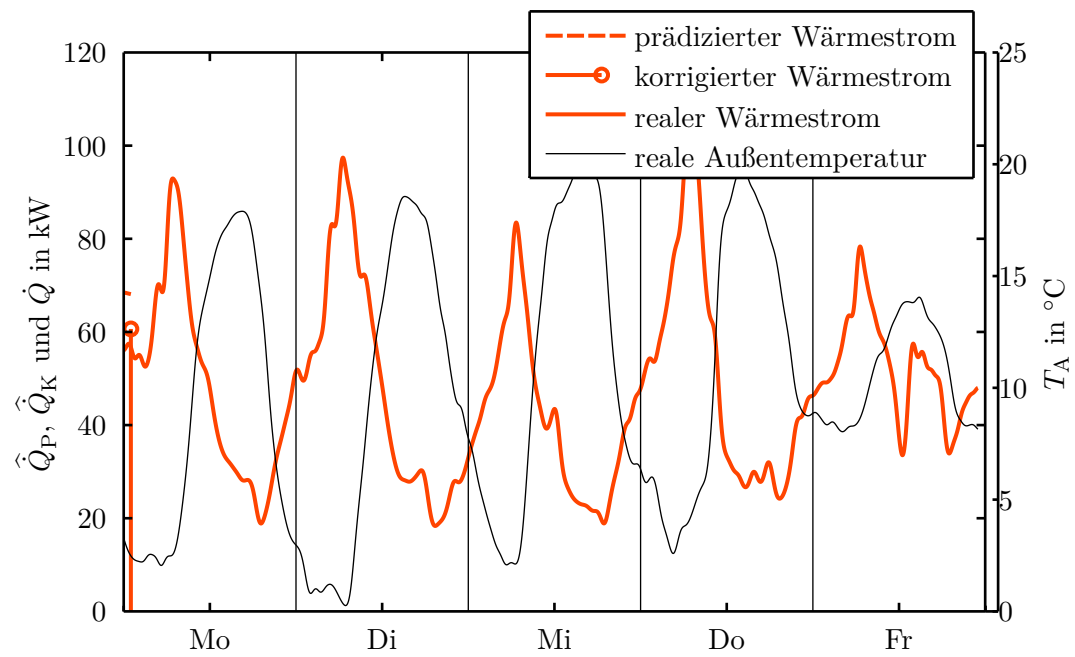


Abbildung 5.28.: VB2 – Prognose für einen kurzen Prognosehorizont; sie gilt nur für Montag 01:00.

6. Finale Methode und deren Implementierung

In diesem Kapitel wird die finale Methode zusammengefasst und auf mögliche Implementierungen eingegangen, womit es der praktischen Umsetzung der entwickelten Prognosemethode gewidmet ist. So werden, entgegen der bisherigen Vorgehensweise, Formeln und Parameterwerte erneut angeschrieben, um eine übersichtliche Zusammenfassung der Prognosemethode zu bieten.

6.1. Finale Methode

Die finale Methode entspricht einer Prädiktor-Korrektor-Prognosemethode, wobei der Prädiktor in Abschnitt 5.2 und der Korrektor in Abschnitt 5.5 beschrieben wurden. Da im voran gegangenen Kapitel 5 viele Evaluierungen und mögliche Kombinationen beschrieben wurden, wird die finale Methode hier übersichtlich zusammengefasst. Die Prognosemethode wird anhand eines Prognosehorizonts von 48 h beschrieben, was gegenüber einer allgemeinen Nomenklatur die Lesbarkeit und Anschaulichkeit erhöht.

Im Zuge dieser Arbeit wurden Parameter bestimmt, die für alle Verbraucher angewandt werden können und zur Laufzeit nicht aktualisiert werden müssen. Diese Parameter sind in Tabelle 6.1 gelistet, wobei jeweils ein Zahlenwert für die Arbeitswoche (MoFr) und das Wochenende (SaSo) bestimmt wurde. Die Werte für k und h_K unterscheiden sich für Arbeitswoche und Wochenende nur gering, weshalb hier auch jeweils der Mittelwert aus MoFr und SaSo für alle Tage verwendet werden könnte.

Parameter	Wert		Bezeichnung	Referenz
	MoFr	SaSo		
N_{TT}	11 d	5 d	Anzahl Trainingstage	(5.7)
k	0,6	0,7	Korrektorparameter	k_1 in (5.16)
h_K	5 h	7 h	Korrektorhorizont	(5.23)

Tabelle 6.1.: Optimal bestimmte Parameter aller Verbraucher.

Zur aktuellen Stunde n_A müssen Messungen des Wärmestroms und der Außentemperatur der vergangenen N_{TT} Tage, sowie eine Wetterprognose der Außentemperatur für den gesamten Prognosehorizont von 48 h, vorliegen. Diese Messdaten sollten akausal gefiltert (siehe Abschnitt 4.3) und anschließend in Arbeitswoche und Wochenende

6. Finale Methode und deren Implementierung

getrennt abgelegt werden, um die Parameterbestimmung des Prädiktors zu erleichtern. Zusätzlich muss für den Korrektor der prädizierte Wärmestrom der aktuellen Stunde $\hat{Q}_P[n_A]$ verfügbar sein.

Für den Prädiktor werden mit den Messungen (Trainingsdaten) entsprechend (5.5) die notwendigen Parametersätze $\mathbf{x}[n] = [\alpha[n] \ \beta[n]]^\top$ bestimmt:

$$\mathbf{y}[n] = \mathbf{A}[n]\mathbf{x}[n] + \mathbf{e}[n], \quad n = n_A + 1, n_A + 2, \dots, n_A + 24, \quad (6.1)$$

mit

$$\mathbf{A}[n] = \begin{bmatrix} 1 & T_A[n - N_{TT} \cdot 24] \\ 1 & T_A[n - (N_{TT} - 1) \cdot 24] \\ \vdots & \vdots \\ 1 & T_A[n - 48] \\ 1 & T_A[n - 24] \end{bmatrix} \quad \text{und} \quad \mathbf{y}[n] = \begin{bmatrix} \dot{Q}[n - N_{TT} \cdot 24] \\ \dot{Q}[n - (N_{TT} - 1) \cdot 24] \\ \vdots \\ \dot{Q}[n - 48] \\ \dot{Q}[n - 24] \end{bmatrix}. \quad (6.2)$$

Hierbei wurde der Index $MoFr$ bzw. $SaSo$ für eine bessere Lesbarkeit weggelassen. Jedoch müssen, je nachdem in welchen Zeitraum die Stunde n fällt, die Trainingsdaten der Arbeitswoche bzw. des Wochenendes verwendet werden. Fallen alle 48 h in einen dieser beiden Zeiträume, so sind 24 Parametersätze für die Prädiktion der 48 h ($n = n_A + 1, \dots, n_A + 48$) ausreichend. Fallen aber zum Beispiel 36 h Stunden in die Arbeitswoche und die restlichen 12 h in das Wochenende so müssen zunächst 24 Parametersätze für $n = n_A + 1, \dots, n_A + 36$ mit Trainingsdaten der Arbeitswoche, und zusätzlich 12 Parametersätze für $n = n_A + 37, \dots, n_A + 48$ mit Trainingsdaten des Wochenendes bestimmt werden. Hierbei ist auch darauf zu achten, dass N_{TT} entsprechend Tabelle 6.1 für die Arbeitswoche bzw. das Wochenende deutlich unterschiedlich ist.

Diese Parametersätze sind in Abbildung 6.1 für beide angeführten Fälle dargestellt und mit Nummern \textcircled{x} bezeichnet. Im ersten Fall liegt der gesamte Prognosehorizont in der Arbeitswoche, weshalb 24 Parametersätze $\textcircled{1}$ ausreichen. Im zweiten Fall fallen die letzten 12 h ins Wochenende, weshalb die Parametersätze $\textcircled{1b}$ nicht erneut verwendet werden können, sondern zusätzliche 12 Parametersätze $\textcircled{2a}$ benötigt werden. Somit sind im ersten Fall 24 Parametersätze $\textcircled{1}$ ausreichend, wohingegen im zweiten Fall $24 + 12 = 36$ Parametersätze $\textcircled{1a}$, $\textcircled{1b}$ und $\textcircled{2a}$ benötigt werden.

Nachdem diese Parametersätze entsprechend der LAV-Kostenfunktion bestimmt wurden, muss mittels Expertenregel 1 von Seite 46 ein steigend linearer Zusammenhang zwischen Wärmestrom und Außentemperatur verhindert werden. Dann kann gemäß (5.1):

$$\hat{Q}_P[n] = \alpha[n] + \beta[n]\hat{T}_A[n], \quad n = n_A + 1, \dots, n_A + 48, \quad (6.3)$$

mit der prognostizierten Außentemperatur der Wärmestrom für die gesamten 48 h prädiziert werden. Das Ergebnis ist entsprechend Expertenregel 2 von Seite 52 auf positive Wärmeströme zu beschränken.

Dieser prädizierte Verlauf wird nun für die nächsten h_K Stunden ($n = n_A + 1, \dots, n_A + h_K$) mittels Korrektor verbessert. Hierfür wird der Prädiktionsfehler der aktuellen Stunde

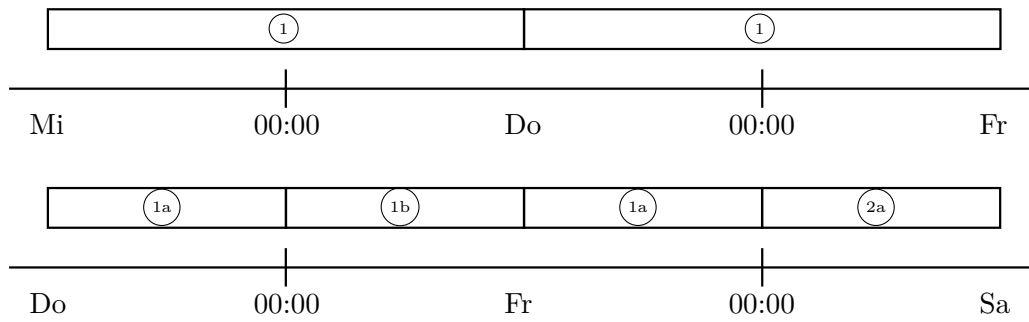


Abbildung 6.1.: Beispiel – Parametersätze für die Prädiktion einer reinen Arbeitswoche und für eine Kombination aus Arbeitswoche und Wochenende.

n_A entsprechend (5.14) bestimmt:

$$e_P[n_A] = \dot{Q}[n_A] - \hat{Q}_P[n_A]. \quad (6.4)$$

Daraus wird die linear abklingende Korrekturfunktion f_K entsprechend (5.22) berechnet:

$$f_K[n] = \begin{cases} ke_P[n_A] - \frac{ke_P[n_A]}{h_K-1}(n - n_A - 1) & n_A < n \leq n_A + h_K, \\ 0 & \text{sonst.} \end{cases} \quad (6.5)$$

mit der die Prädiktion korrigiert wird (5.21):

$$\hat{Q}_K[n] = \hat{Q}_P[n] + f_K[n]. \quad (6.6)$$

Dieser korrigierte Wärmestrom \hat{Q}_K entspricht der Prognose für die nächsten 48 h Stunden zum Zeitpunkt n_A , wobei erneut mittels Expertenregel 2 nur positive Wärmeströme zugelassen werden.

Die hier beschriebene Vorgehensweise ist stündlich zu wiederholen, womit jede Stunde eine aktuelle Prognose zur Verfügung steht. Für die Bestimmung der Parametersätze des Prädiktors kommen hierbei die Messwerte T_A und \dot{Q} der vergangenen Stunde hinzu und die am weitesten zurückliegenden Messwerte dieser Stunde werden nicht mehr berücksichtigt. Somit ist es für diese Iteration ausreichend nur die Parametersätze dieser vergangenen Stunde durch einmaliges Lösen von (6.1) zu aktualisieren. Die anderen Parametersätze können nicht aktualisiert werden, da für diese Stunden noch keine aktuelleren Messdaten zur Verfügung stehen.

6.2. Implementierung

In diesem Abschnitt wird die Auswirkung durch die Verwendung von Wetterprognosedaten und eine ressourcenschonenderen Kostenfunktion untersucht. So stehen bei einer realen Implementierung für die Zukunft nur Wetterprognosedaten zur Verfügung, deren

6. Finale Methode und deren Implementierung

Qualität das Ergebnis maßgeblich beeinflussen. Ein weiterer wichtiger Aspekt bei der Implementierung ist der Ressourcenbedarf und insbesondere der Implementierungsaufwand der Berechnungen. Da die Parametersätze des Prädiktors als einzige Parameter zur Laufzeit bestimmt werden müssen, wird hierfür eine ressourcenschonendere Kostenfunktion betrachtet. Diese beiden Implementierungsaspekte (Wetterprognose und Kostenfunktion) werden zunächst im Detail beschrieben und anschließend für alle Verbraucher evaluiert und diskutiert.

6.2.1. Wetterprognosedaten

Für diese Untersuchung werden Wetterprognosedaten eines Wetterinstituts benutzt, welche aus einem räumlich und zeitlich grob aufgelösten Großraumwettermodell stammen. Als Beispiel sind in Abbildung 6.2 drei Verläufe für die Außentemperatur von VB1 dargestellt.

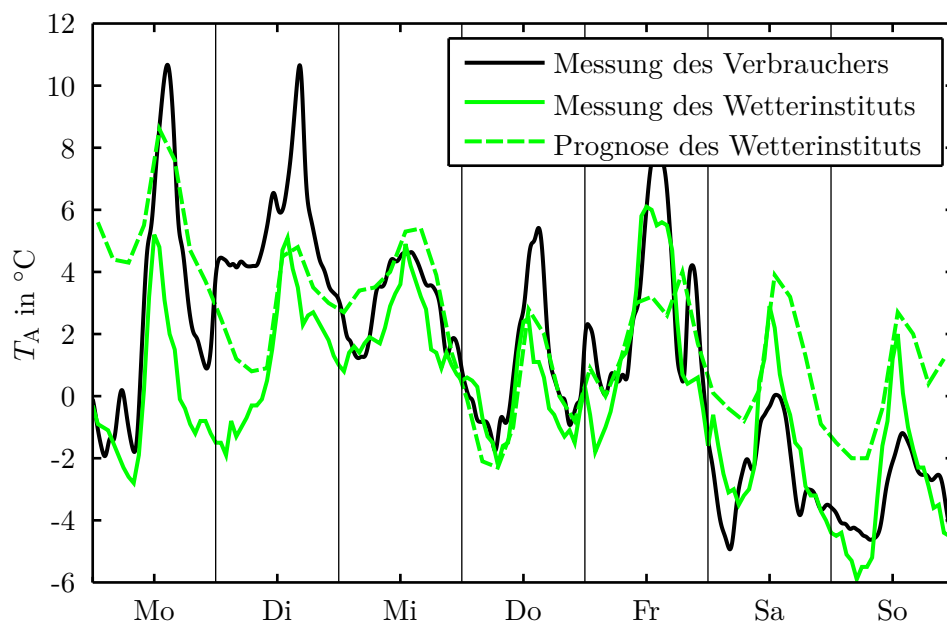


Abbildung 6.2.: VB1 – Vergleich von Messungen des Verbrauchers und des Wetterinstituts inklusive Wetterprognose des Wetterinstituts.

Es sind zwei Messungen eingezeichnet, da sich die Wetterstation üblicherweise an einer anderen geographischen Position als der Verbraucher befindet, womit hier bereits eine deutliche Abweichung festzustellen ist. Somit würde selbst bei perfekter Wetterprognose die tatsächliche Außentemperatur beim Verbraucher nicht erreicht werden. Diese Abweichung wird durch die grobe Wetterprognose noch verschlechtert bzw. in manchen Fällen auch zufällig verbessert.

Um diesen systematischen Fehler zu kompensieren, könnten als Trainingsdaten für den Prädiktor die Wetterprognosedaten des Wetterinstituts verwendet werden. Solch eine Fehlerkorrektur ist besonders wirksam, falls die Temperaturdifferenz zwischen der Messung des Verbrauchers und der Wetterprognose des Wetterinstituts konstant ist.

6.2.2. Ressourcenschonende Kostenfunktion

Bis jetzt wurden die Parametersätze des Prädiktors unter Minimierung der LAV-Kostenfunktion bestimmt, welche nicht geschlossen lösbar ist. Daher wird die Auswirkung der Verwendung der OLS- bzw. der WLS-Kostenfunktion untersucht, da diese Optimierungsaufgaben mittels Pseudoinverser geschlossen lösbar sind. Dadurch würde zum einen der Implementierungsaufwand und die benötigte Rechenzeit drastisch reduziert werden. Aber zum anderen werden bei OLS und WLS die einzelnen Fehler (Residuen) quadriert, womit große Fehler unerwünscht skaliert werden (vgl. Abschnitt 2.2). Die Auswirkung dieser Skalierung auf die Parameter des Prädiktors, und somit auf das Prognoseergebnis, ist das Entscheidungskriterium für oder gegen diese effizienteren Kostenfunktionen OLS bzw. WLS.

Bei WLS sind laut (2.16) Gewichte w_i zu wählen mit denen die einzelnen Fehler e_i gewichtet werden. Hierbei ist es sinnvoll die aktuelleren Messwerte stärker zu gewichten als jene die schon älter sind. Wird der Fehlervektor \mathbf{e} entsprechend der Matrizen und Vektoren in (6.2) berechnet, so kann die zeitliche Abhängigkeit der Gewichte in der Diagonalmatrix \mathbf{W} geschrieben werden als

$$\begin{aligned} \mathbf{W}[n] &= \begin{bmatrix} w_1 & & & & & & & & \\ & w_2 & & & & & & & \\ & & \ddots & & & & & & \\ & & & & w_{N_{\text{TT}}-1} & & & & \\ & & & & & & w_{N_{\text{TT}}} & & \\ & & & & & & & & \end{bmatrix} \\ &= \begin{bmatrix} w[n - N_{\text{TT}} \cdot 24] & & & & & & & & \\ & w[n - (N_{\text{TT}} - 1) \cdot 24] & & & & & & & \\ & & \ddots & & & & & & \\ & & & & & & w[n - 48] & & \\ & & & & & & & & w[n - 24] \end{bmatrix}. \end{aligned} \quad (6.7)$$

Um nun aktuellere Messwerte stärker zu gewichten wird $w_1 = 1$ fixiert und die restlichen Gewichte steigen bis zu einem definierten $w_{N_{\text{TT}}} \geq 1$ linear an. Somit ist $w_{N_{\text{TT}}}$ ein Parameter der wieder für alle Verbraucher bestimmt werden soll. Diese Bestimmung wird analog zu der Bestimmung der global-optimalen Parameter aus Tabelle 6.1 in den bereits bekannten drei Schritten durchgeführt (vgl. Abschnitt 5.2).

Bei dieser Evaluierung hat sich ergeben, dass der Parameter $w_{N_{\text{TT}}}$ je nach Kalenderwoche und Verbraucher entweder gleich Eins oder entsprechend des oberen Optimierungslimits gewählt werden sollte. Diese Ergebnisse entsprechen „keiner Gewichtung“ bzw. „maximaler Gewichtung“, womit hier kein globales Optimum bestimmt werden konnte, da es ausschließlich von der Wahl des oberen Optimierungslimits abhängt. Aus diesem

6. Finale Methode und deren Implementierung

Grund wird die Auswirkung der Parameterbestimmung mittels OLS-Kostenfunktion untersucht, was ja für viele Kalenderwochen dem Optimum „keiner Gewichtung“ entspricht.

6.2.3. Ergebnis

Die bisherigen Lastprognoseergebnisse wurden mittels LAV-Kostenfunktion für die Parameterbestimmung des Prädiktors und der gemessenen Außentemperatur des Verbrauchers T_A als Eingangsgröße des Prädiktors ermittelt. Nun wird die Auswirkung der OLS-Kostenfunktion und der prognostizierten Außentemperatur \hat{T}_A auf die Lastprognose für alle Verbraucher evaluiert und diskutiert.

Wie bereits in Abschnitt 5.5 beschrieben hängt das Ergebnis maßgeblich vom Prognosehorizont ab, weshalb nur die Extrema der reinen Prädiktion und des ersten Korrektors mit einem Prognosehorizont von einer Stunde evaluiert werden. Das tatsächliche Ergebnis für einen bestimmten Prognosehorizont liegt zwischen dem langen Horizont, bei dem der Einfluss des Korrektors verschwindet (reine Prädiktion), und dem kürzest möglichen Horizont von einer Stunde für den ersten Korrektor. Um eine Aufteilung in Kalenderwochen zu ermöglichen, wird für die reine Prädiktion ein Prognosehorizont von 5 Tagen für die Arbeitswoche bzw. von 2 Tagen für das Wochenende betrachtet. Um die Anzahl der Abbildungen zu begrenzen wird hier nur die Arbeitswoche evaluiert; das Ergebnis für das Wochenende führt aber zu den gleichen Schlussfolgerungen.

Für die Bewertungen werden zunächst einzelne Prognosen für jede Kalenderwoche des Jahres mit den vier möglichen Kombinationen aus LAV bzw. OLS und T_A bzw. \hat{T}_A getrennt bestimmt. Hierbei wurden die Parametersätze für die reine Prädiktion nur einmal zu Beginn der Kalenderwoche bestimmt und im Laufe des Prognosehorizonts von 5 Tagen (Arbeitswoche) nicht aktualisiert. Für die erste Korrektur mit einem Prognosehorizont von einer Stunde wurden die aktuellsten (täglich aktualisierten) Parametersätze für die Prädiktion verwendet. Diese stündlichen Prognosen des ersten Korrektors wurden aneinandergereiht, um eine durchgehende Prognose für die gesamten 5 Tage (Arbeitswoche) zu erzeugen. Diese Prognosen je Kalenderwoche wurden wieder mittels MAE bewertet und diese MAE-Werte je Verbraucher auf Jahreszeiten gemittelt. Somit kann je Verbraucher für jede der vier Kombinationen aus LAV bzw. OLS und T_A bzw. \hat{T}_A ein MAE-Wert je Jahreszeit bestimmt werden und hiermit die Prognosequalität beurteilt werden. Da der mittlere absolute Fehler MAE den Prognosefehler angibt, entspricht ein kleinerer MAE einer besseren Prognose.

Die Ergebnisse für die reine Prädiktion (Prognosehorizont 5 Tage) sind in Abbildungen 6.3 und 6.4, und für den ersten Korrektor (Prognosehorizont eine Stunde) in Abbildungen 6.5 und 6.6 zu sehen. Man beachte, dass für VB4 und VB6 keine Wetterprognosedaten zu Verfügung stehen, weshalb diese nur bezüglich der Kostenfunktionen evaluiert werden.

Die erste Kombination aus LAV und T_A entspricht der bisherigen Vorgehensweise und dient somit als Referenz. Als Trainingsdaten wurde immer die Messung des Verbrauchers verwendet.

Die Auswahl der Kostenfunktion zeigt einen vergleichsweise geringen Einfluss auf das Prognoseergebnis. Allerdings hat die Wetterprognose einen massiven Einfluss, welche eine zum Teil starke Verschlechterung bewirkt (siehe VB2 und VB7). Für VB5 liefert die Wetterprognose bessere Ergebnisse als die Messung vor Ort, was auf zufällige Abweichungen zurückzuführen ist.

Zusammengefasst zeigte die Verwendung der OLS-Kostenfunktion keine wesentlichen Nachteile gegenüber der Verwendung der LAV-Kostenfunktion und kann deshalb der Einfachheit halber für die Implementierung eingesetzt werden. Die Verwendung der Wetterprognosedaten anstatt der Messungen des Verbrauchers verschlechtert das Ergebnis zum Teil stark, und ist somit maßgeblich für die Qualität der Lastprognose verantwortlich. Diese Verschlechterung könnte durch eine bessere Wetterprognose und die Verwendung von Wetterprognosedaten des Wetterinstituts als Trainingsdaten reduziert werden, was im Zuge dieser Arbeit nicht mehr untersucht werden konnte. Diese beiden Punkte sollten in weiteren Arbeiten näher untersucht werden.

6. Finale Methode und deren Implementierung

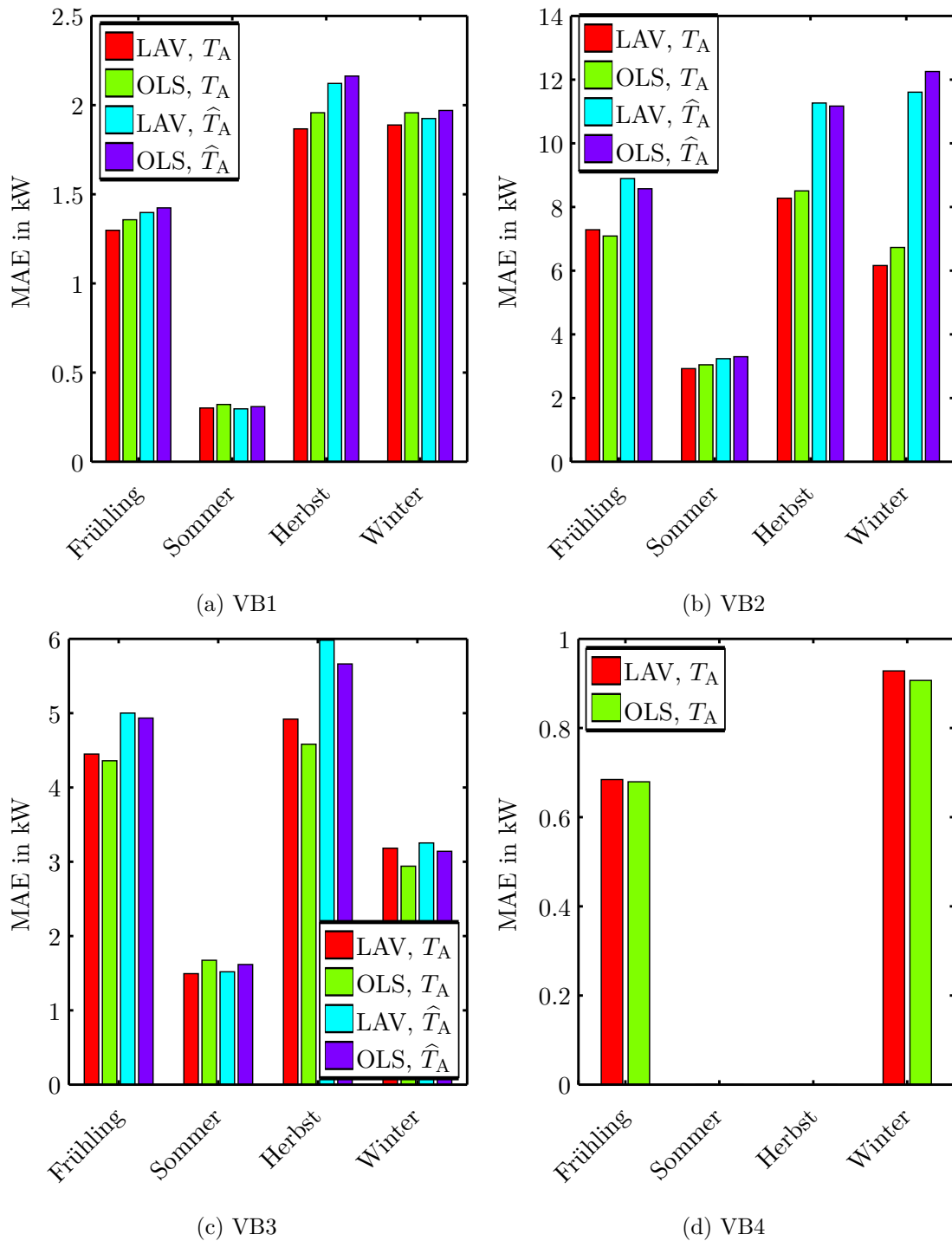
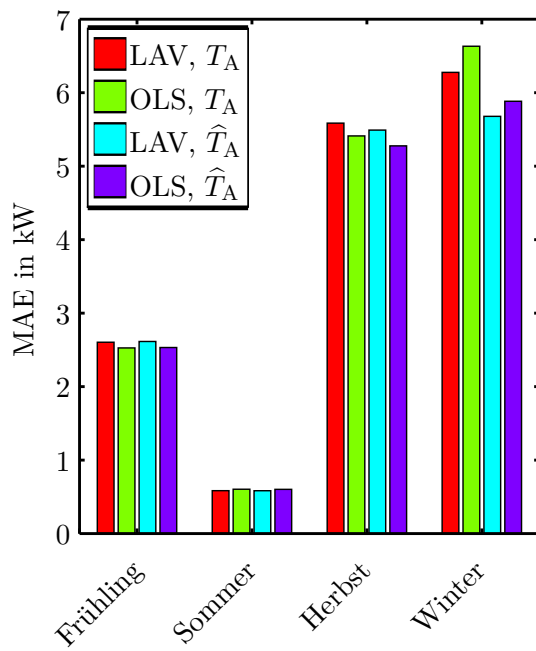
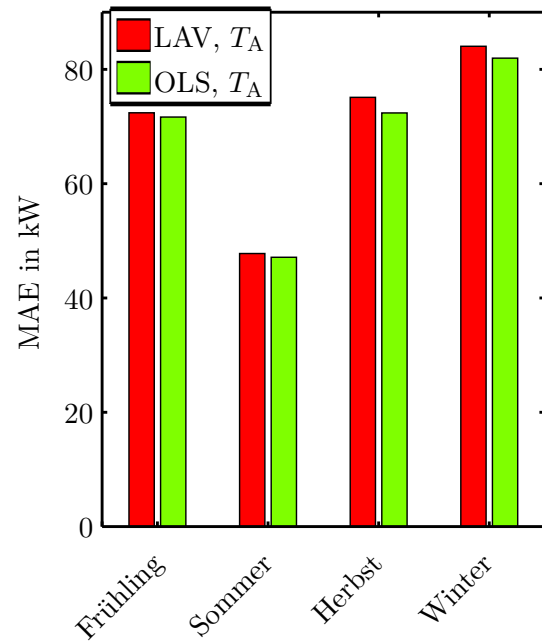


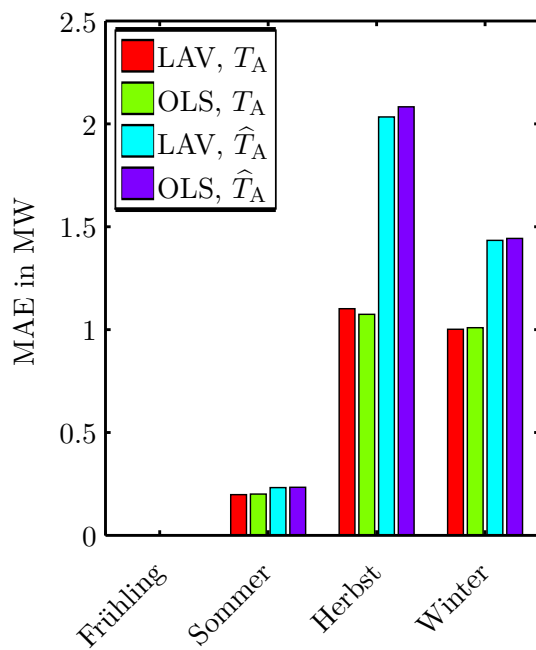
Abbildung 6.3.: Evaluierung der Auswirkung möglicher Implementierungen der reinen Prädiktion für die Arbeitswoche. Mit LAV und OLS werden die beiden betrachteten Kostenfunktionen für die Parameteroptimierung des Prädiktors bezeichnet. Bei T_A werden Messungen des Verbrauchers für die Außentemperatur verwendet und \hat{T}_A entspricht der Verwendung von Wetterprognosedaten.



(a) VB5



(b) VB6



(c) VB7

Abbildung 6.4.: Evaluierung der Auswirkung m glicher Implementierungen der reinen Pr diktionsf r die Arbeitswoche. F r eine Erkl rung der Legende siehe Abbildung 6.3.

6. Finale Methode und deren Implementierung

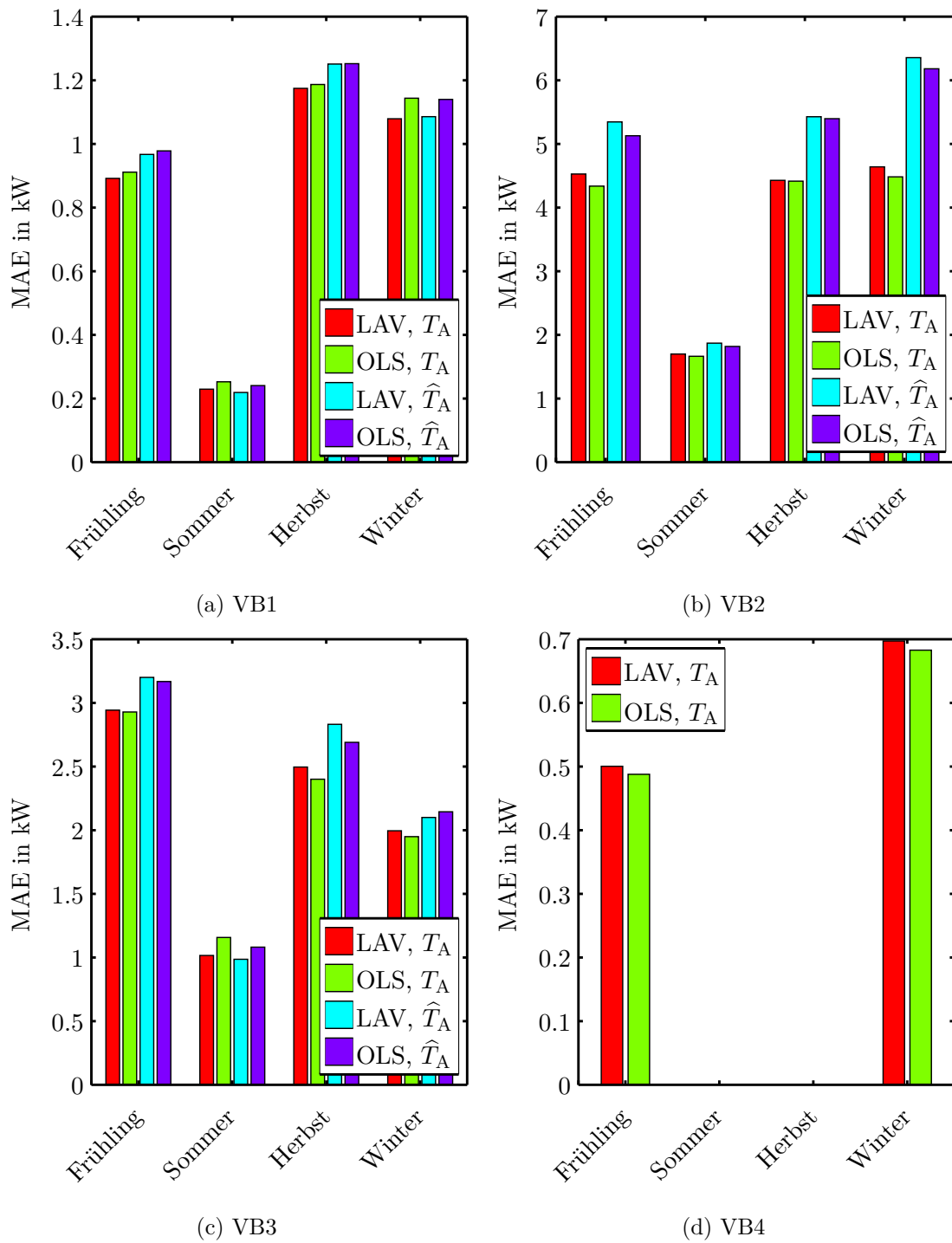
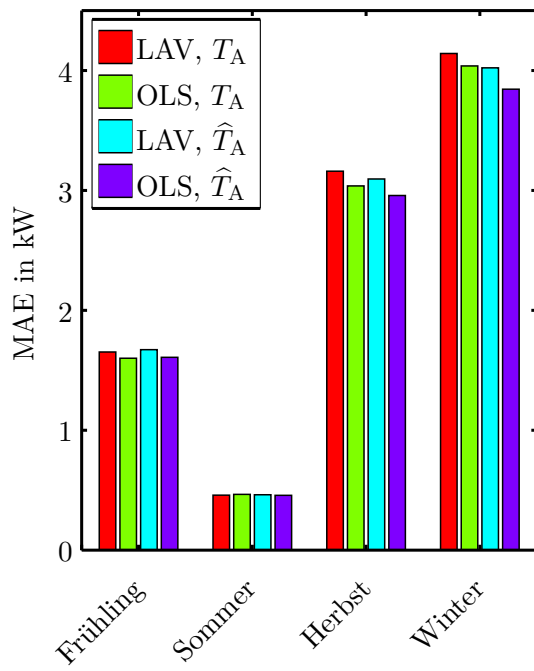
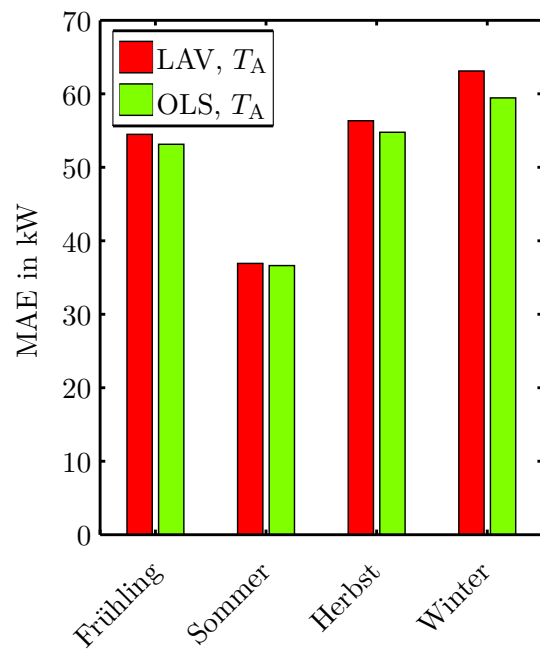


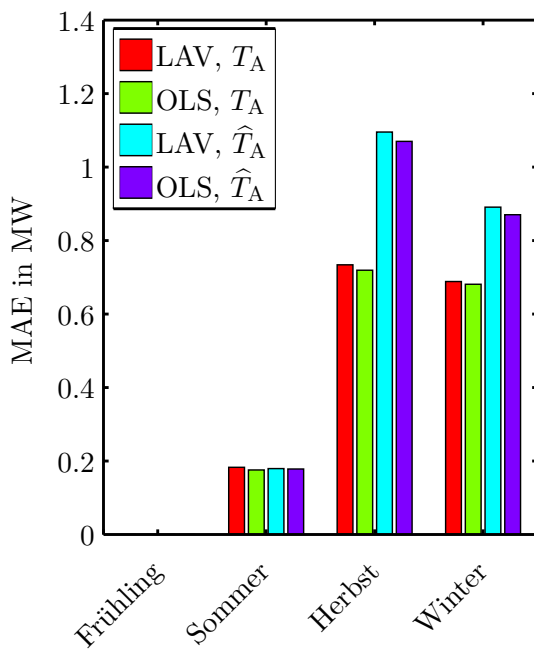
Abbildung 6.5.: Evaluierung der Auswirkung möglicher Implementierungen des ersten Korrektors für die Arbeitswoche. Für eine Erklärung der Legende siehe Abbildung 6.3.



(a) VB5



(b) VB6



(c) VB7

Abbildung 6.6.: Evaluierung der Auswirkung möglicher Implementierungen des ersten Korrektors für die Arbeitswoche. Für eine Erklärung der Legende siehe Abbildung 6.3.

7. Zusammenfassung und Ausblick

In dieser Arbeit wurde eine Methode für die Prognose des Wärmebedarfs von Gebäuden und Wärmenetzen entwickelt. Durch die Verifizierung mittels sieben verschiedener Verbraucher (vom Wohnhaus bis zum Fernwärmenetz) ist diese Prognosemethode keine Speziallösung, sondern ohne händische Parametrierung für eine große Variation an Verbraucherklassen einsetzbar. Allerdings ist diese Lastprognose von der Qualität Wetterprognose (Außentemperatur) abhängig, deren Einfluss noch weiterer Untersuchungen bedarf.

7.1. Zusammenfassung

Die entwickelte Prognosemethode besteht aus einem Prädiktor, der beliebig weit in die Zukunft präzisieren kann, und einem Korrektor, der die Prädiktion der nahen Zukunft verbessert. Hierbei wird die Prädiktion mit zunehmendem Horizont natürlich schlechter. Zusätzlich werden durch zwei Expertenregeln unsinnige Parameterwerte und Prognosen verhindert.

Der Prädiktor besteht aus einem linearen Regressionmodell für jede Stunde des Tages, womit Regelmäßigkeiten in Benutzerverhalten, Gebäudeautomation und Wettereinflüssen (z.B. Sonnenaufgang) berücksichtigt werden. Dieses Regressionsmodell beschreibt für jede Stunde den Zusammenhang zwischen benötigtem Wärmestrom des Verbrauchers und Außentemperatur mit einer Geraden. Die Außentemperatur hat von den Wetterfaktoren den stärksten Einfluss auf den Wärmestrom und enthält indirekt teilweise auch andere Faktoren, wie die Globalstrahlung. Erweiterungen des Prädiktors um zusätzliche Wetterfaktoren brachten keine Verbesserung. Der Prädiktor benötigt eine Wetterprognose für die Außentemperatur, von deren Qualität die Qualität der Lastprognose maßgeblich abhängt. Die Parametersätze des Regressionsmodells werden durch Optimierung aus den Messdaten der vergangenen Tage ermittelt und stündlich aktualisiert. Für diese Optimierung kann die ressourcenschonende OLS-Kostenfunktion verwendet werden. Es hat sich gezeigt, dass die Arbeitswoche und das Wochenende ein zum Teil stark unterschiedliches Verhalten aufweisen. Daher werden diese Zeiträume getrennt behandelt, was in der Parameterbestimmung für den Prädiktor berücksichtigt wird.

Der Korrektor verwendet den Prädiktionsfehler der aktuellen Stunde, um die Prognose für die nahe Zukunft zu verbessern. Hierfür wird dieser Prädiktionsfehler gewichtet zur Prädiktion der nächsten Stunde vorzeichenrichtig addiert, um bleibende Abweichungen zu reduzieren. Dieser Korrektor wird linear abklingend auf die nächsten Stunden ausgeweitet.

Mit der vorgestellten Prognosemethode kann eine stündlich aktualisierte Lastprognose bestimmt werden, wobei nur die Parametersätze des Prädiktors aktualisiert werden

7. Zusammenfassung und Ausblick

müssen. Ansonsten wurden alle Parameter, anhand von Untersuchungen aller sieben Verbraucher, auf Zahlenwerte fixiert. Somit ist der Rechenaufwand je Zeitschritt (Stunde) sehr gering und die Prognosemethode kann auch auf Steuerungen mit begrenzten Ressourcen eingesetzt werden.

Die Prognosequalität ist bereits hoch (vgl. Beispiele der reinen Prädiktion in Anhang A.1) und besonders gut für Zeiten mit einer großen Heizlast (Winter), für Verbraucher mit großer Anschlussleistung (Wärmenetze) und für Verbraucher mit regelmäßigem Benutzerverhalten (Bürogebäude). Diese Prognosequalität ist allerdings von einer guten Wetterprognose abhängig, wozu noch weitere Untersuchungen erforderlich sind.

7.2. Ausblick

Die entwickelte Prognosemethode ist grundsätzlich vollständig entwickelt und für eine große Variation an Verbraucherklassen getestet. Somit wäre der nächste Schritt die Implementierung und der Test an einem Realsystem.

Allerdings ist die Prognosemethode (Prädiktor) von einer guten Wetterprognose abhängig, deren Einfluss noch genauer untersucht werden muss. Für die Untersuchungen in Abschnitt 6.2 wurde die Wetterprognose eines räumlich und zeitlich grob aufgelösten Wettermodells verwendet, die das Lastprognoseergebnis stark verschlechterten. Daher ist hier ein feiner aufgelöstes Wettermodell notwendig, das eventuell auch aktuelle Messungen in Verbrauchernähe für die Verbesserung der Wetterprognose einbezieht. Solche Wetterprognosen sind Stand der Technik und können somit auch für diese Lastprognose ohne weitere Entwicklungen eingesetzt werden. Hier ist allerdings noch zu untersuchen inwiefern das Lastprognoseergebnis, durch die Verwendung von Wetterstationsmessdaten bzw. Wetterprognosedaten als Trainingsdaten für die Parameterbestimmung des Prädiktors, verbessert werden kann.

Die Prognosemethode kann aber auch als solide Basis für verschiedene Weiterentwicklungen im Bereich der Prognosemethoden dienen. Durch die konsequente Trennung in Prädiktor und Korrektor kann zum Beispiel der Korrektor ersetzt werden ohne den Prädiktor zu beeinflussen.

Eine Möglichkeit der Verbesserung steckt in der Berücksichtigung von Benutzereingaben. Wird zum Beispiel im Herbst die Zentralheizung eines Gebäudes eingeschaltet, so könnte dieses Wissen in der Prognosemethode durch eine weitere Expertenregel berücksichtigt werden. Für Verbraucher bei denen die Heizlast vorherrscht (z.B. Bürogebäude) müssen nur Effekte berücksichtigt werden, die sich auch in der Gebäudeautomation widerspiegeln. Wird zum Beispiel während der Betriebsferien die Heizung nicht gedrosselt, so muss dies der Prognosemethode auch nicht mitgeteilt werden. Je stärker sich die Gebäudeautomation an das Benutzerverhalten anpasst (Smart-Home), desto mehr Zusatzinformationen können inkludiert werden, um die Prognosemethode auf individuelle Verbraucher zu spezialisieren.

Eine weitere Möglichkeit der Verbesserung liegt in einer feineren Unterteilung des Tages als die bisherige Stundenbasis. Somit könnten vor allem kurzfristige Spitzen deutlich besser abgebildet werden.

Literaturverzeichnis

- [1] G. Gross und F. Galiana, „Short-term load forecasting,” *Proceedings of the IEEE*, Vol. 75, Nr. 12, S. 1558–1573, Dec 1987.
- [2] L. Sachs und J. Hedderich, *Angewandte Statistik, Methodensammlung mit R*, 12. Aufl. Springer, 2006.
- [3] A. V. Oppenheim und R. W. Schaffer, *Discrete-Time Signal Processing*, 2. Aufl. Prentice-Hall, 1999.
- [4] G. B. Dantzig und M. N. Thapa, *Linear Programming 2, Theory and Extensions*, P. Glynn und S. M. Robinson, Hrsgg. Springer, 2003.
- [5] B. R. Auer und H. Rottmann, *Statistik und Ökonometrie für Wirtschaftswissenschaftler: Eine anwendungsorientierte Einführung*, 2. Aufl. Gabler, 2011.
- [6] R. Schlittgen und B. Streitberg, *Zeitreihenanalyse*, 9. Aufl., Serie Lehr- und Handbücher der Statistik. Oldenbourg Wissenschaftsverlag, 2001.
- [7] R. Isermann, *Mechatronische Systeme*, 2. Aufl. Springer, 2008.
- [8] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. John Wiley & Sons, 1996.
- [9] R. J. Hyndman, „Another look at forecast-accuracy metrics for intermittent demand,” *Foresight: The International Journal of Applied Forecasting*, Vol. 4, Nr. 4, S. 43–46, 2006.
- [10] P. Goodwin und R. Lawton, „On the asymmetry of the symmetric {MAPE},” *International Journal of Forecasting*, Vol. 15, Nr. 4, S. 405 – 408, 1999.
- [11] H. A. Nielsen und H. Madsen, „Predicting the Heat Consumption in District Heating Systems using Meteorological Forecasts,” Ph.D. dissertation, Department of Mathematical Modelling, Technical University of Denmark, Lyngby, 2000.
- [12] D. Feinberg, E. Genethliou, *Applied Mathematics for Restructured Electric Power Systems: Optimization, Control, and Computational Intelligence*. Springer, 2005, ch. 12 Load Forecasting, S. 269–285. [Online]. Available: <http://www.ams.sunysb.edu/~feinberg/public/>
- [13] I. Moghram und S. Rahman, „Analysis and evaluation of five short-term load forecasting techniques,” *Power Systems, IEEE Transactions on*, Vol. 4, Nr. 4, S. 1484–1491, Nov 1989.

- [14] H. A. Nielsen und H. Madsen, „Modelling the heat consumption in district heating systems using a grey-box approach,” *Energy and Buildings*, Vol. 38, Nr. 1, S. 63–71, 2006.
- [15] P. Bacher, H. Madsen, und H. Nielsen, „Online short-term heat load forecasting for single family houses,” in *Industrial Electronics Society, IECON 2013 - 39th Annual Conference of the IEEE*, Nov 2013, S. 5741–5746.
- [16] P. Koponen, „Short-term load forecasting model based on smart metering data: Daily energy prediction using physically based component model structure,” in *Smart Grid Technology, Economics and Policies (SG-TEP), 2012 International Conference on*, Dec 2012, S. 1–4.
- [17] R. Ramanathan, R. Engle, C. W. Granger, F. Vahid-Araghi, und C. Brace, „Short-run forecasts of electricity loads and peaks,” *International Journal of Forecasting*, Vol. 13, Nr. 2, S. 161 – 174, 1997.
- [18] S. Soliman, S. Persaud, K. El-Nagar, und M. El-Hawary, „Application of least absolute value parameter estimation based on linear programming to short-term load forecasting,” *International Journal of Electrical Power & Energy Systems*, Vol. 19, Nr. 3, S. 209 – 216, 1997.
- [19] H. Cho, Y. Goude, X. Brossat, und Q. Yao, „Modeling and forecasting daily electricity load curves: a hybrid approach,” *Journal of the American Statistical Association*, Vol. 108, Nr. 501, S. 7–21, 2013.
- [20] T. Mestekemper, G. Kauermann, und M. S. Smith, „A comparison of periodic autoregressive and dynamic factor models in intraday energy demand forecasting,” *International Journal of Forecasting*, Vol. 29, Nr. 1, S. 1 – 12, 2013.
- [21] E. Dotzauer, „Simple model for prediction of loads in district-heating systems,” *Applied Energy*, Vol. 73, Nr. 3–4, S. 277 – 284, 2002.
- [22] S. Grosswindhager, A. Voigt, und M. Kozek, „Online short-term forecast of system heat load in district heating networks,” in *ISF 2011 - Prague PROCEEDINGS*, Vol. 31, 2011, 1997-41.
- [23] N. Amjady, „Short-term hourly load forecasting using time-series modeling with peak load estimation capability,” *Power Systems, IEEE Transactions on*, Vol. 16, Nr. 4, S. 798–805, Nov 2001.
- [24] S. Bercu und F. Proia, „A SARIMAX coupled modelling applied to individual load curves intraday forecasting,” *Journal of Applied Statistics*, Vol. 40, Nr. 6, S. 1333–1348, 2013.
- [25] W. Charytoniuk, M.-S. Chen, und P. Van Olinda, „Nonparametric regression based short-term load forecasting,” *Power Systems, IEEE Transactions on*, Vol. 13, Nr. 3, S. 725–730, Aug 1998.

- [26] S.-J. Huang und K.-R. Shih, „Short-term load forecasting via ARMA model identification including non-Gaussian process considerations,” *Power Systems, IEEE Transactions on*, Vol. 18, Nr. 2, S. 673–679, May 2003.
- [27] Q. Ding, H. Zhang, T. Huang, und J. Zhang, „A holiday short term load forecasting considering weather information,” in *Power Engineering Conference, 2005. IPEC 2005. The 7th International*, Nov 2005, S. 1–61.
- [28] A. Jain, E. Srinivas, und R. Rauta, „Short term load forecasting using fuzzy adaptive inference and similarity,” in *Nature Biologically Inspired Computing, 2009. NaBIC 2009. World Congress on*, Dec 2009, S. 1743–1748.
- [29] T. Senjyu, P. Mandal, K. Uezato, und T. Funabashi, „Next day load curve forecasting using hybrid correction method,” *Power Systems, IEEE Transactions on*, Vol. 20, Nr. 1, S. 102–109, Feb 2005.
- [30] R.-H. Liang und C.-C. Cheng, „Combined regression-fuzzy approach for short-term load forecasting,” *Generation, Transmission and Distribution, IEE Proceedings-*, Vol. 147, Nr. 4, S. 261–266, Jul 2000.
- [31] M. C. Medeiros und L. J. Soares, „Robust statistical methods for electricity load forecasting,” in *RTE-VT workshop*, Paris, 2006.
- [32] H. Daneshi und A. Daneshi, „Real time load forecast in power system,” in *Electric Utility Deregulation and Restructuring and Power Technologies, 2008. DRPT 2008. Third International Conference on*, April 2008, S. 689–695.
- [33] J. Taylor und R. Buizza, „Neural network load forecasting with weather ensemble predictions,” *Power Systems, IEEE Transactions on*, Vol. 17, Nr. 3, S. 626–632, Aug 2002.
- [34] H. Hippert, C. Pedreira, und R. Souza, „Neural networks for short-term load forecasting: a review and evaluation,” *Power Systems, IEEE Transactions on*, Vol. 16, Nr. 1, S. 44–55, Feb 2001.
- [35] K. M. Powell, A. Sriprasad, W. J. Cole, und T. F. Edgar, „Heating, cooling, and electrical load forecasting for a large-scale district energy system,” *Energy*, Vol. 74, Nr. 0, S. 877 – 885, 2014, International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems – {ECOS} 2013.
- [36] C. Turhan, T. Kazanasmaz, I. E. Uygun, K. E. Ekmen, und G. G. Akkurt, „Comparative Study of a Building Energy Performance Software (KEP-IYTE-ESS) and Ann-Based Building Heat Load Estimation,” *Energy and Buildings*, Nr. 0, S. –, 2014.
- [37] Q. Mu, Y. Wu, X. Pan, L. Huang, und X. Li, „Short-term load forecasting using improved similar days method,” in *Power and Energy Engineering Conference (APPEEC), 2010 Asia-Pacific*. IEEE, 2010, S. 1–4.

- [38] Y. Yang, Y. Meng, Y. Xia, Y. Lu, und H. Yu, „An efficient approach for short term load forecasting,” in *International Multiconference of engineers and computer scientists*, 2011.
- [39] Z. Hou und Z. Lian, „An Application of Support Vector Machines in Cooling Load Prediction,” in *Intelligent Systems and Applications, 2009. ISA 2009. International Workshop on*, May 2009, S. 1–4.
- [40] M. Horn und N. Dourdoumas, *Regelungstechnik - Rechnerunterstützter Entwurf zeitkontinuierlicher und zeitdiskreter Regelkreise*. Pearson Studium, 2004.
- [41] E. F. Cellier und E. Kofman, *Continuous System Simulation*. Springer, 2006.
- [42] E. Raviv, „Prediction bias correction for dynamic term structure models,” *Economics Letters*, Vol. 129, Nr. 0, S. 112 – 115, 2015.
- [43] F. Woodcock und C. Engel, „Operational consensus forecasts,” *Weather and forecasting*, Vol. 20, Nr. 1, S. 101–111, 2005.
- [44] D. J. Stensrud und J. A. Skindlov, „Gridpoint predictions of high temperature from a mesoscale model,” *Weather and forecasting*, Vol. 11, Nr. 1, S. 103–110, 1996.
- [45] Verein Deutscher Ingenieure, VDI-Gesellschaft Verfahrenstechnik und Chemieingenieurwesen (GVC), *VDI-Wärmeatlas*, 7. Aufl. VDI Verlag, 1994.
- [46] W. Streicher, *Vorlesungsskriptum Sonnenenergienutzung*, Institut für Wärmetechnik Std., 2005.
- [47] B. Chramcov, „Forecast of heat demand according the Box-Jenkins methodology for specific locality,” *Latest Trends on Systems*, Vol. 1, S. 252–256, 2010.
- [48] H. Al-Hamadi und S. Soliman, „Short-term electric load forecasting based on Kalman filtering algorithm with moving window weather and load model,” *Electric Power Systems Research*, Vol. 68, Nr. 1, S. 47 – 59, 2004.
- [49] D. R. Palacio, „Simple model for load forecast weather sensitive,” in *Electricity Distribution, 2001. Part 1: Contributions. CIRED. 16th International Conference and Exhibition on (IEE Conf. Publ No. 482)*, Vol. Summaries, 2001, S. 288–288.
- [50] G. B. Dantzig und M. N. Thapa, *Linear Programming 1, Introduction*, P. Glynn, Hrsg. Springer, 1997.
- [51] S. Boyd und L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

A. Zusatzmaterial

A.1. Verbraucherdaten und Prognosebeispiele

In diesem Abschnitt des Anhangs werden zunächst die Zeiträume angegeben in denen Verbraucherdaten vorhanden sind. Diese Auflistung ist entsprechend des Jahres in Tabellen A.1, A.1 und A.2 aufgeteilt, wobei vorhandene Daten mit einem Punkt gekennzeichnet sind.

Um einen tieferen Einblick in die Qualität der Prognose zu ermöglichen, werden Prognosebeispiele für alle sieben Verbraucher (siehe Tabelle 4.1) dargestellt. Als Prognose wird hier die reine Prädiktion mit den aktuellsten Parametersätzen (Neuermittlung der Parameter alle 24 h) verwendet. Die Beschränkung auf reine Prädiktion erfolgt wegen der guten Darstellbarkeit und der einfachen Qualitätsbewertung, da für jede Stunde nur eine aktuellste Prädiktion existiert. Würde der Korrektor inkludiert werden, so müsste für jede Stunde des betrachteten Zeitraums eine eigene Einzelprognose erstellt werden, wobei hier auch noch der Prognosehorizont zu wählen wäre.

Die Prognosebeispiele (eigentlich Prädiktionsbeispiele) sind in Abbildung A.1 bis A.14 zu sehen. Hierbei wurde für jeden Verbraucher eine willkürliche Kalenderwoche (KW) gewählt, und die Arbeitswoche bzw. das Wochenende wurden getrennt prädiziert. Die Qualitätsmaße des Zeitraums sind in der Beschreibung der Abbildung inkludiert, damit die Zahlenwerte in der restlichen Arbeit mit Beispielen unterfüttert werden.

A. Zusatzmaterial

2012	VB1	VB2	VB3	VB4	VB5	VB6	VB7
Jänner							
KW1						•	
KW2						•	
KW3						•	
KW4						•	
Februar							
KW5						•	
KW6						•	
KW7						•	
KW8	•	•	•			•	
März							
KW9	•	•	•			•	
KW10	•	•	•			•	
KW11	•	•	•			•	
KW12	•	•	•			•	
KW13	•	•	•			•	
April							
KW14	•	•	•			•	
KW15	•	•	•			•	
KW16	•	•	•			•	
KW17	•	•	•			•	
Mai							
KW18	•	•	•			•	
KW19	•	•	•			•	
KW20	•	•	•			•	
KW21	•	•	•			•	
KW22	•	•	•			•	
Juni							
KW23	•	•	•			•	
KW24	•	•	•			•	
KW25	•	•	•			•	
KW26	•	•	•			•	
Juli							
KW27	•	•	•			•	
KW28	•	•	•			•	
KW29	•	•	•			•	
KW30	•	•	•			•	
August							
KW31	•	•	•			•	
KW32	•	•	•			•	
KW33	•	•	•			•	
KW34	•	•	•			•	
KW35	•	•	•			•	
September							
KW36	•	•	•			•	
KW37	•	•	•			•	
KW38	•	•	•			•	
KW39	•	•	•			•	
Oktober							
KW40	•	•	•			•	
KW41	•	•	•			•	
KW42	•	•	•			•	
KW43	•	•	•			•	
November							
KW44	•	•	•			•	
KW45	•	•	•			•	
KW46	•	•	•			•	
KW47	•	•	•			•	
KW48	•	•	•			•	
Dezember							
KW49	•	•	•			•	
KW50	•	•	•			•	
KW51	•	•	•			•	
KW52	•	•	•			•	

Tabelle A.1.: Betrachtete Daten der Verbraucher für 2012.

A.1. Verbraucherdaten und Prognosebeispiele

2013	VB1	VB2	VB3	VB4	VB5	VB6	VB7
Jänner							
KW1				•			•
KW2				•			•
KW3				•			•
KW4				•			•
KW5				•			•
Februar							
KW6				•			•
KW7				•			•
KW8				•			•
KW9				•			•
März							
KW10				•	•		•
KW11				•	•		•
KW12				•	•		•
KW13				•	•		•
April							
KW14				•	•		•
KW15				•	•		•
KW16				•	•		•
KW17				•	•		•
Mai							
KW18				•	•		•
KW19				•	•		•
KW20				•	•		•
KW21					•		•
KW22					•		•
Juni							
KW23					•		•
KW24					•		•
KW25					•		•
KW26					•		•
Juli							
KW27					•	•	•
KW28					•	•	•
KW29					•	•	•
KW30					•	•	•
August							
KW31					•		•
KW32					•		•
KW33					•		•
KW34					•		•
KW35					•		•
September							
KW36					•		•
KW37					•		•
KW38					•		•
KW39					•		•
Oktober							
KW40					•		•
KW41					•		•
KW42					•		•
KW43					•		•
KW44					•		•
November							
KW45					•	•	•
KW46					•		•
KW47					•		•
KW48					•		•
Dezember							
KW49					•		•
KW50					•		•
KW51					•		•
KW52					•		•

Tabelle A.2.: Betrachtete Daten der Verbraucher für 2013.

A. Zusatzmaterial

2014	VB1	VB2	VB3	VB4	VB5	VB6	VB7
Jänner							
KW1					•		•
KW2					•		•
KW3					•		•
KW4					•		•
KW5					•		•
Februar							
KW6							•
KW7							•
KW8							•
KW9							•
März							
KW10							
KW11							
KW12							
KW13							
April							
KW14							
KW15							
KW16							
KW17							
Mai							
KW18							
KW19							
KW20							
KW21							
KW22							
Juni							
KW23							
KW24							
KW25							
KW26							
Juli							
KW27							
KW28							
KW29							
KW30							
KW31							
August							
KW32							
KW33							
KW34							
KW35							
September							
KW36							
KW37							
KW38							
KW39							
Oktober							
KW40							
KW41							
KW42							
KW43							
KW44							
November							
KW45							
KW46							
KW47							
KW48							
Dezember							
KW49							
KW50							
KW51							
KW52							

Tabelle A.3.: Betrachtete Daten der Verbraucher für 2014.

A.1. Verbraucherdaten und Prognosebeispiele

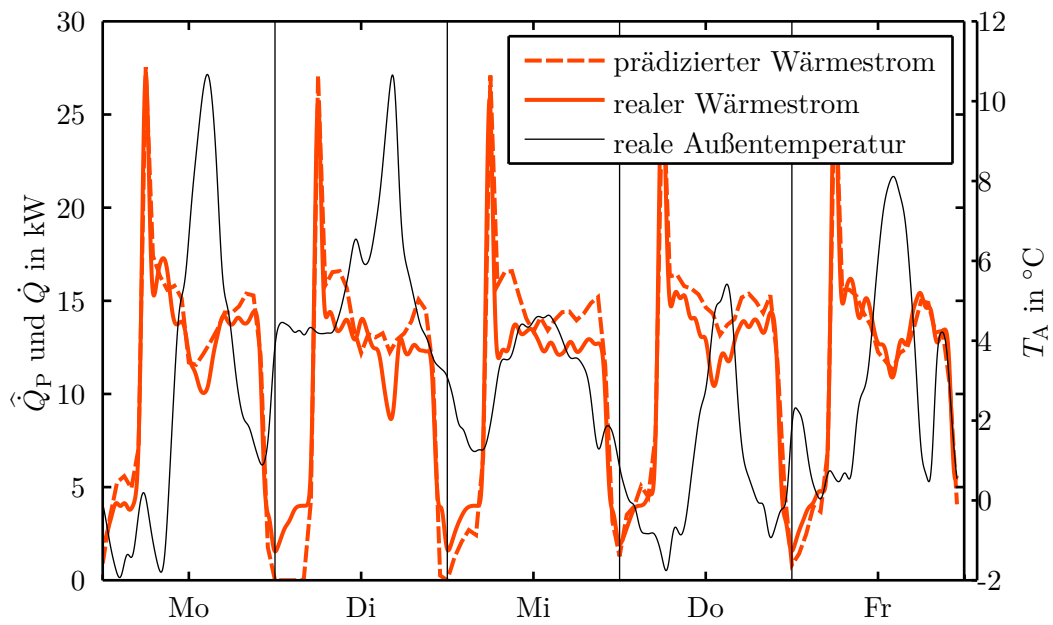


Abbildung A.1.: VB1 – KW52, MAE = 1,33 kW, MPE = 0,4 %, MAPE = 18,3 %.

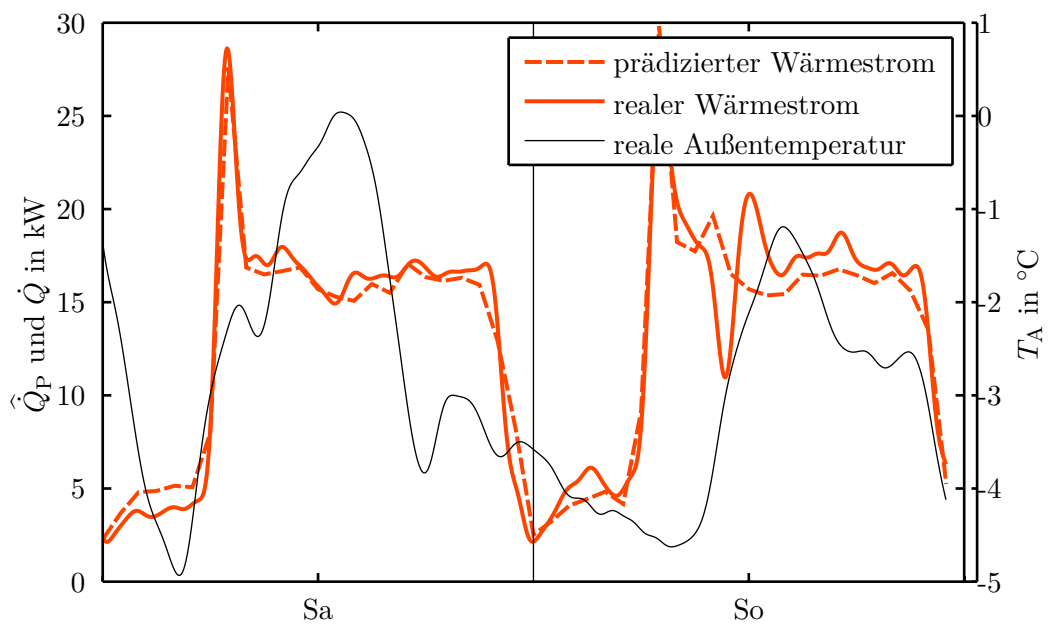


Abbildung A.2.: VB1 – KW52, MAE = 1,16 kW, MPE = -2,8 %, MAPE = 12,5 %.

A. Zusatzmaterial

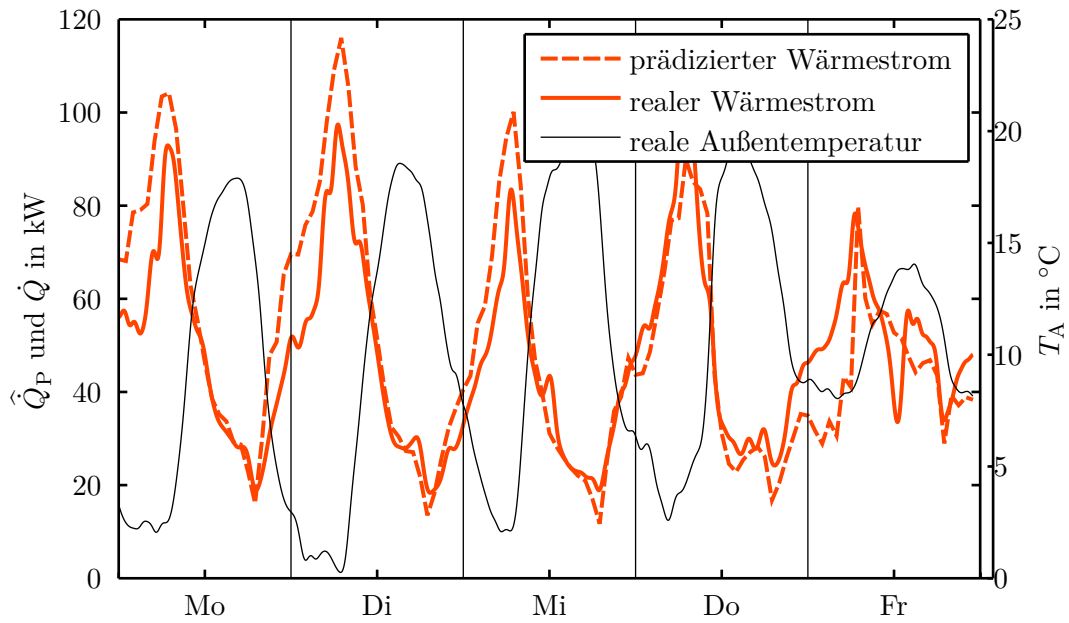


Abbildung A.3.: VB2 – KW13, MAE = 8,96 kW, MPE = -3,4 %, MAPE = 17,8 %.

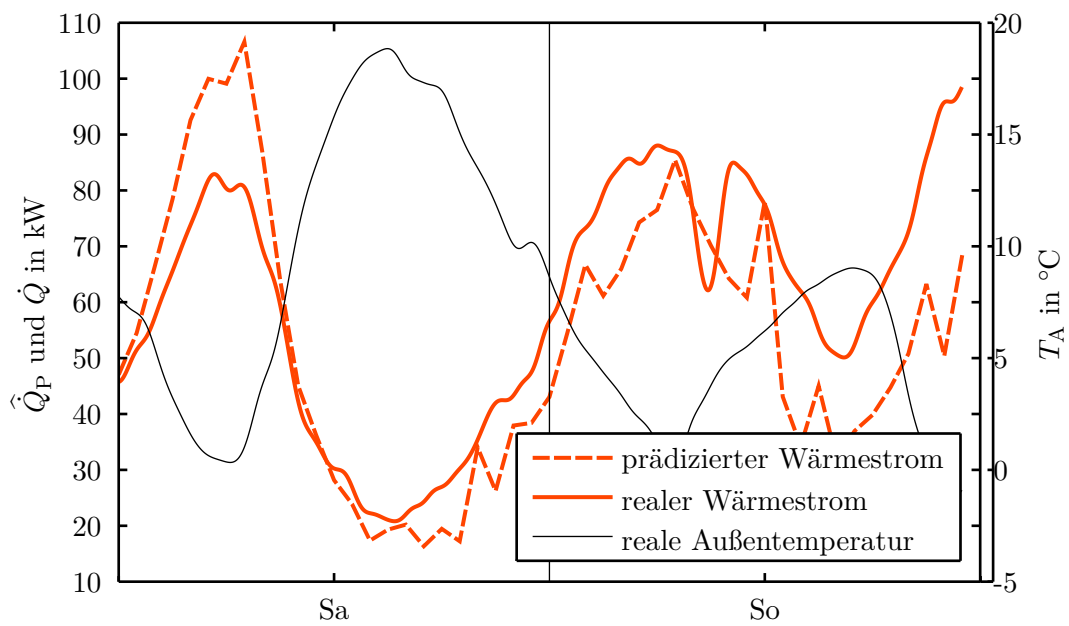


Abbildung A.4.: VB2 – KW13, MAE = 12,50 kW, MPE = 11,7 %, MAPE = 20,3 %.

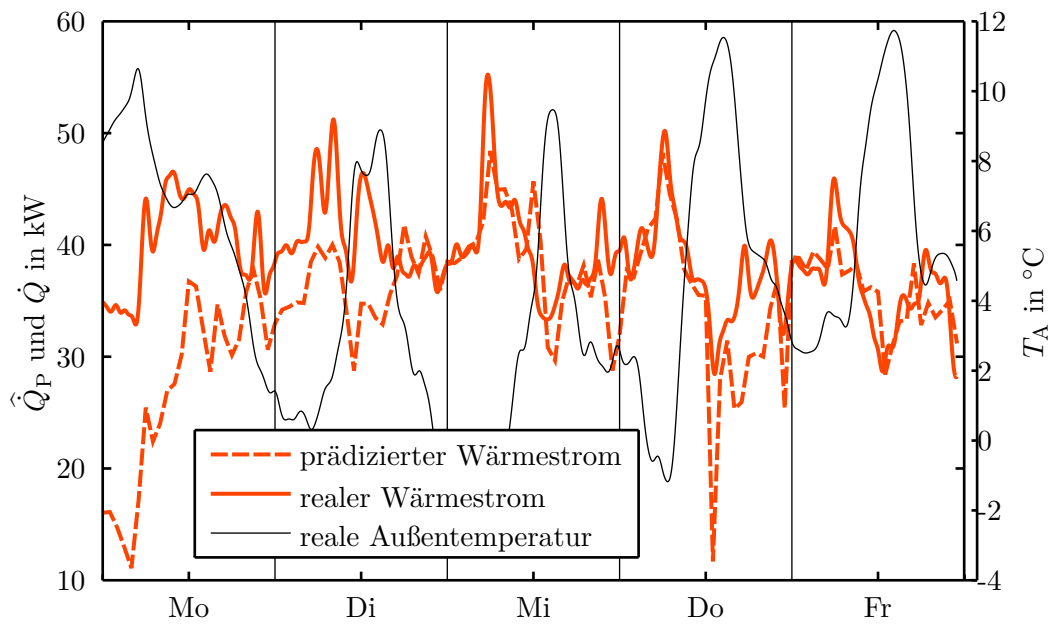


Abbildung A.5.: VB3 – KW45, MAE = 5,39 kW, MPE = 11,5 %, MAPE = 13,8 %.

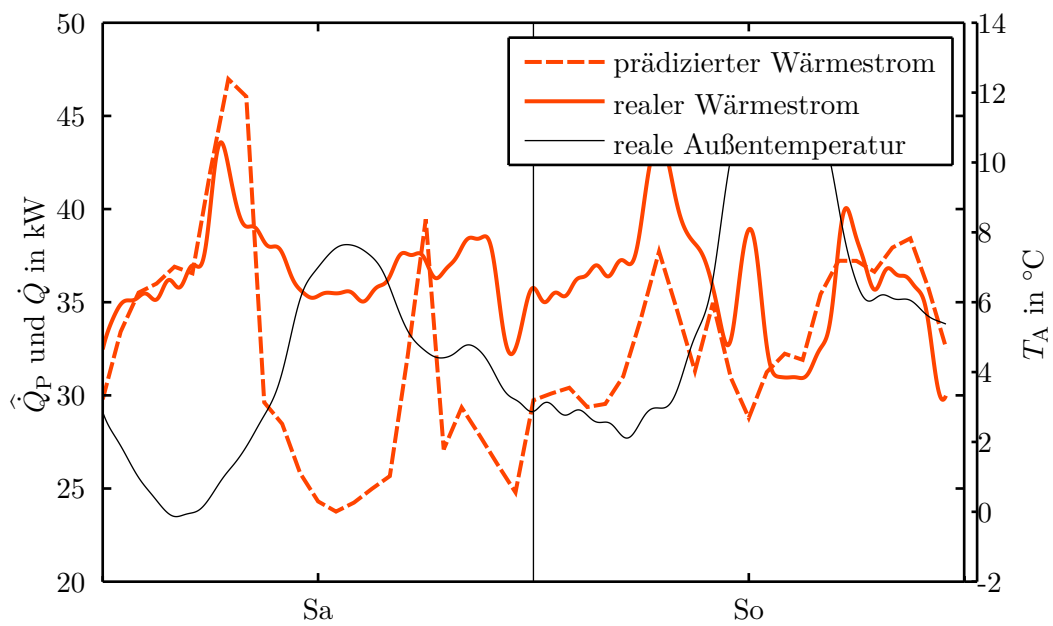


Abbildung A.6.: VB3 – KW45, MAE = 5,12 kW, MPE = 10,7 %, MAPE = 14,0 %.

A. Zusatzmaterial

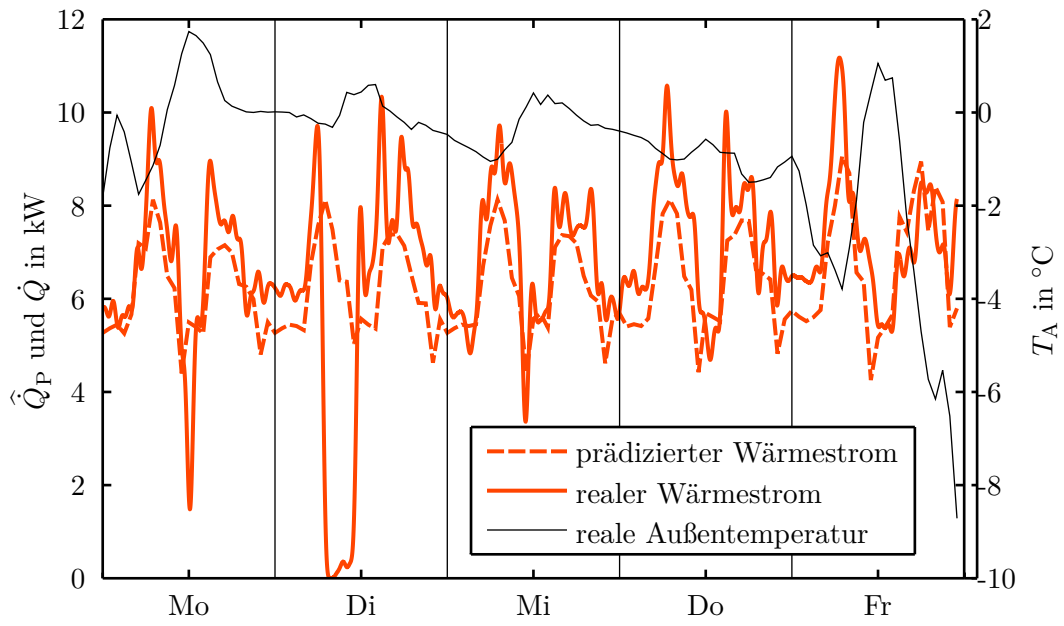


Abbildung A.7.: VB4 – KW4, MAE = 1,15 kW, MPE = 5,8 %, MAPE = 16,3 %.

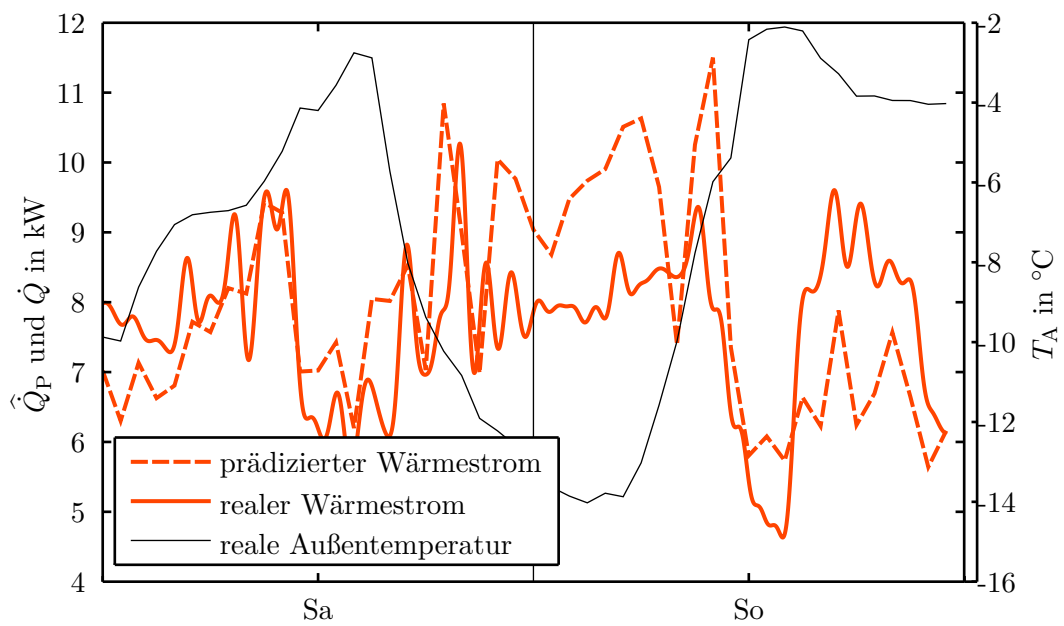


Abbildung A.8.: VB4 – KW4, MAE = 1,14 kW, MPE = -4,9 %, MAPE = 14,9 %.

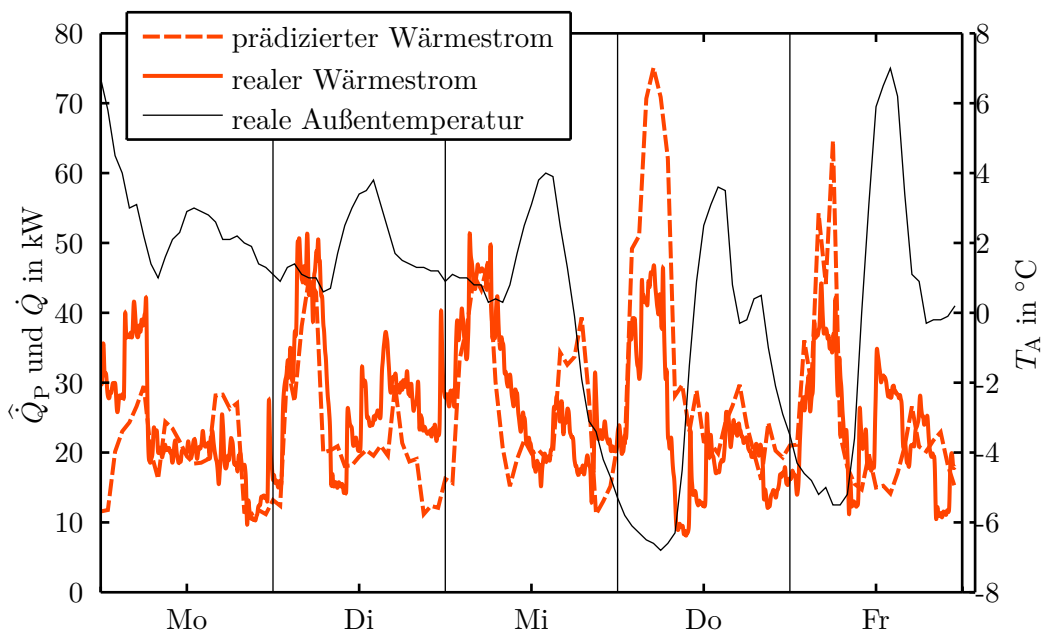


Abbildung A.9.: VB5 – KW48, MAE = 7,74 kW, MPE = -6,7%, MAPE = 33,1%.

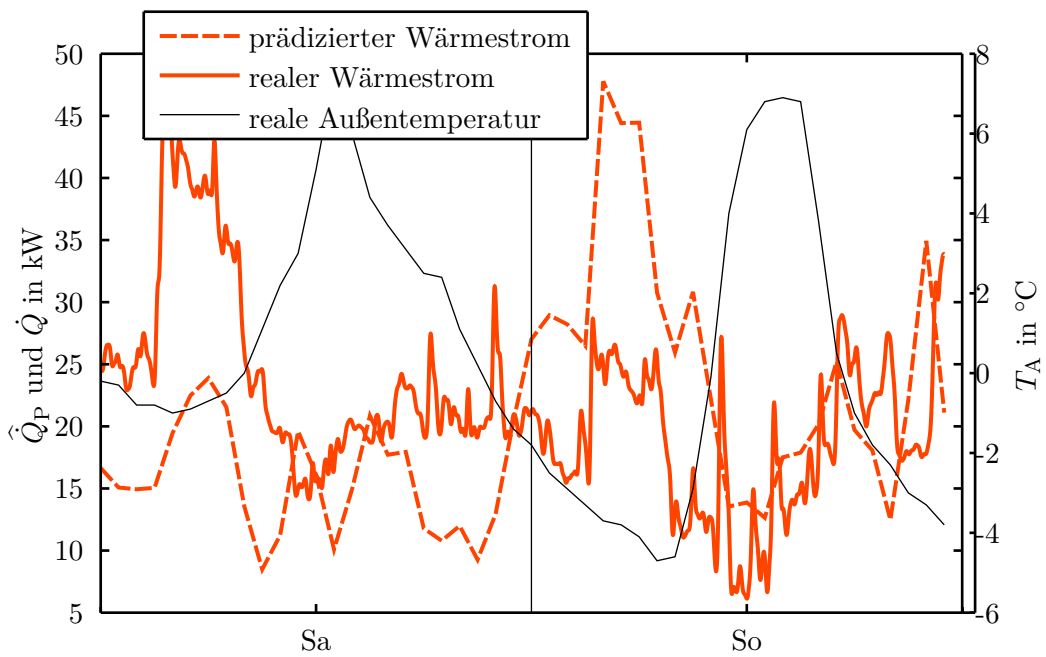


Abbildung A.10.: VB5 – KW48, MAE = 9,26 kW, MPE = -5,3%, MAPE = 44,4%.

A. Zusatzmaterial

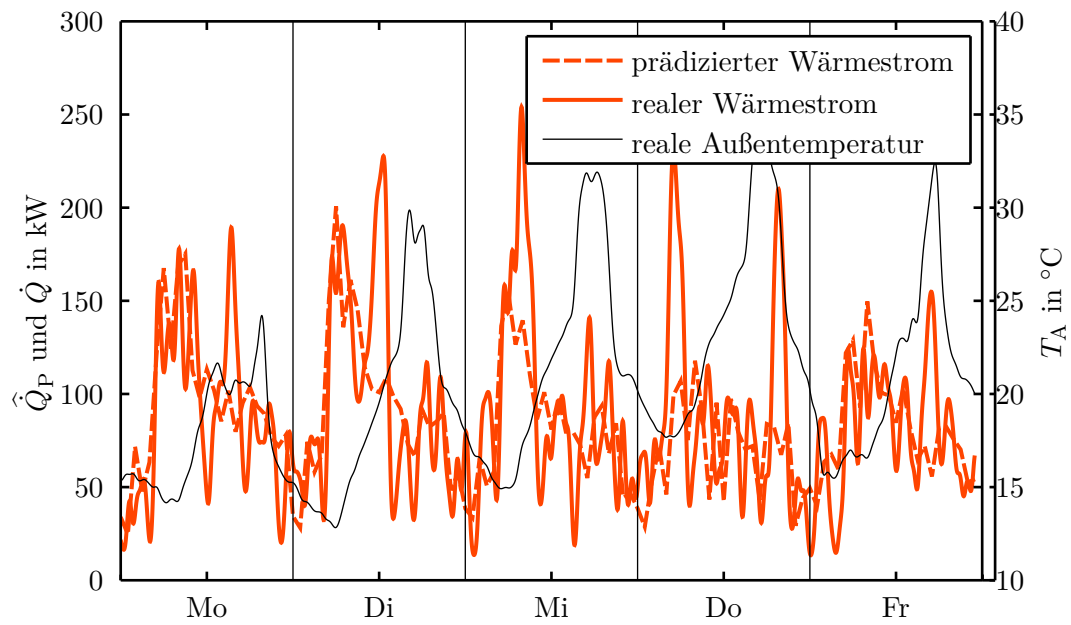


Abbildung A.11.: VB6 – KW31, MAE = 28,44 kW, MPE = 14,5 %, MAPE = 40,0 %.

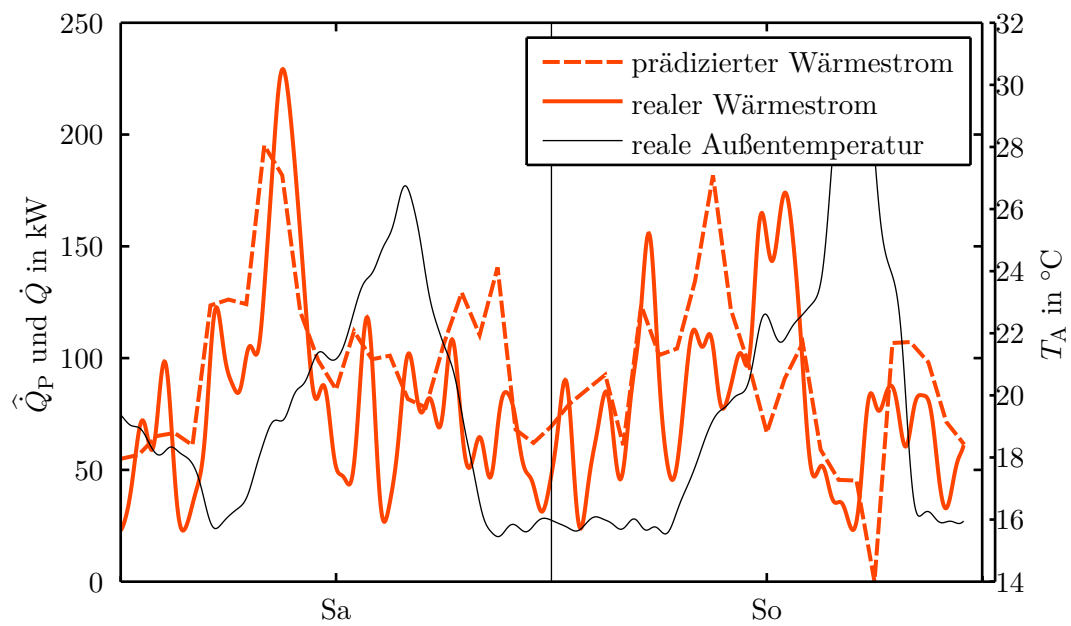


Abbildung A.12.: VB6 – KW31, MAE = 31,36 kW, MPE = -35,8 %, MAPE = 49,4 %.

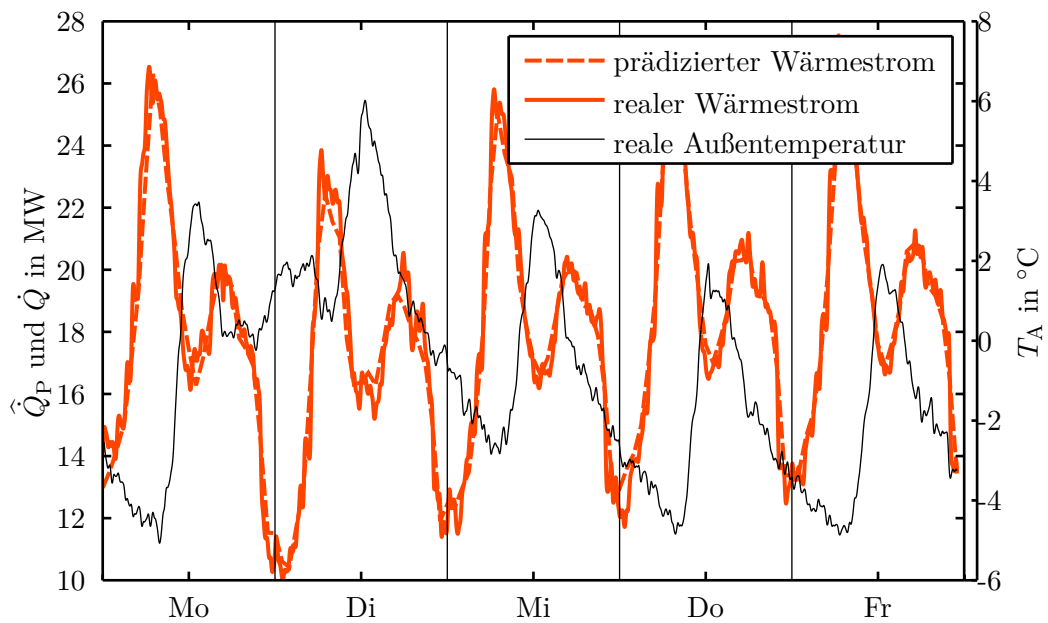


Abbildung A.13.: VB7 – KW50, MAE = 565,24 kW, MPE = -0,1 %, MAPE = 3,2 %.

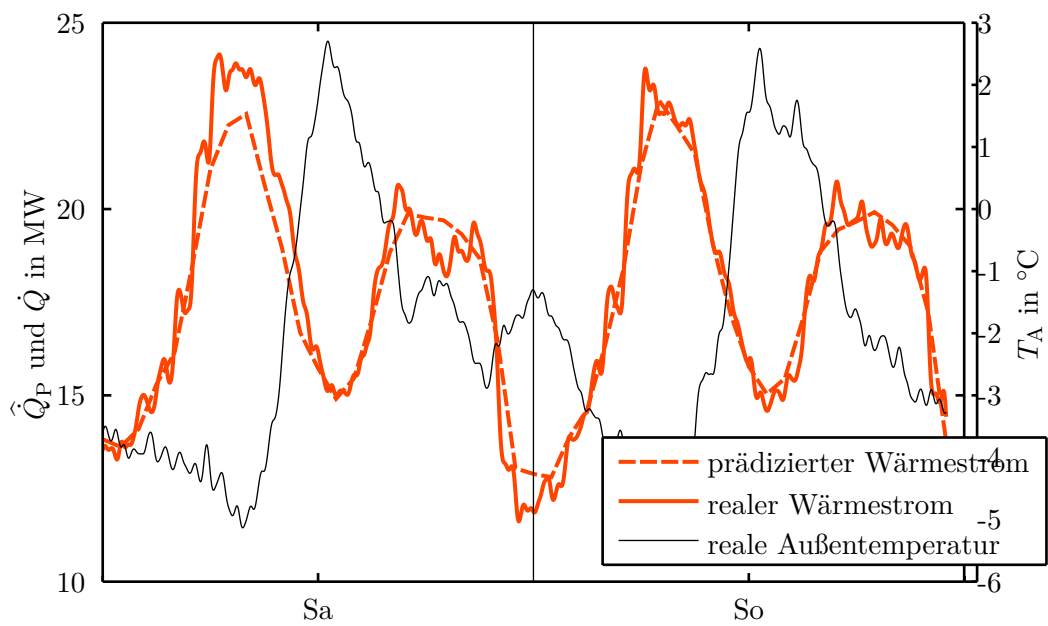


Abbildung A.14.: VB7 – KW50, MAE = 626,81 kW, MPE = 0,6 %, MAPE = 3,4 %.

A.2. Lineare Programme in Matlab

In diesem Abschnitt werden die drei betrachteten Methoden zur Lösung des Optimierungsproblems für die Parameterbestimmung des Prädiktors mittels Matlab beschrieben. Für dieses Optimierungsproblem wurde die LAV-Kostenfunktion verwendet, die auch mittels Linearem Programm (LP) formuliert werden kann. In diesem Anhang liegt der Fokus auf der Umsetzung in Matlab und für die Theorie zu Linearen Programmen wird auf einschlägige Literatur wie zum Beispiel [4, 50, 51] verwiesen.

Ziel ist die Lösung des überbestimmten Gleichungssystems

$$\mathbf{Ax} = \mathbf{y} \quad (\text{A.1})$$

mittels Minimierung der Summe der absoluten Fehler. Der Fehler (Residuum) wird definiert als

$$\mathbf{e} = \mathbf{y} - \mathbf{Ax}. \quad (\text{A.2})$$

Somit lautet das Optimierungsproblem für die LAV-Kostenfunktion

$$\min_{\mathbf{x}} \sum_i |e_i| = \min_{\mathbf{x}} \|\mathbf{e}\|_1 = \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{Ax}\|_1, \quad (\text{A.3})$$

wobei $\|\mathbf{e}\|_1$ die Einsernorm des Vektors beschreibt. Für die numerische Lösung von (A.3) sollen drei Matlab-Methoden betrachtet werden: `fminsearch`, `yalmip`¹ und `linprog`. Das Problem (A.3) kann direkt mit `fminsearch` gelöst werden.

```

%% FMINSEARCH
x0 = ones(n,1);
x_FMINSEARCH = fminsearch(@(x) norm(y-A*x,1), x0 );

```

Für die Lösung des LP mit `yalmip` muss das Problem um eine Schlupfvariable \mathbf{r} erweitert werden (siehe [51, S.294])

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{r}} \quad & \mathbf{1}^\top \mathbf{r} \\ \text{unter der Bedingung} \quad & -\mathbf{r} \leq \mathbf{y} - \mathbf{Ax} \leq \mathbf{r} \\ & \mathbf{r} \geq \mathbf{0} \end{aligned} \quad (\text{A.4})$$

wobei die zweite Bedingung $\mathbf{r} \geq \mathbf{0}$ laut [51, S.294] auch weggelassen werden kann. Sie wird aber trotzdem verwendet, da dies einer Standardform für LP entspricht. Das LP (A.4) kann in Matlab wie folgt geschrieben werden

```

%% YALMIP
[m,n] = size(A); % Bestimme Dimensionen
yalmip('clear'); % YALMIP starten

```

¹YALMIP ist eine Beschreibungssprache für die Formulierung und Lösung konvexer und nicht-konvexer Optimierungsprobleme. Sie ist als kostenfreie Toolbox für Matlab implementiert. <http://users.isy.liu.se/johanl/yalmip/>

```

x = sdpvar(n,1);    % definiere YALMIP-Variable
r = sdpvar(m,1);

C = set(-r <= y - A*x <= r); % C wie Constraints (Nebenbedingungen)
C = C + set(r >= 0);

diagnostic = solvesdp(C, ones(1,m)*r); % Löse(Nebenbedingung, Zielfunktion)

xYALMIP = double(x);
r = double(r);

```

Für die Verwendung von `linprog` muss das Problem (A.3) auf die Form

$$\min_{\tilde{\mathbf{x}}} \tilde{\mathbf{f}}^\top \tilde{\mathbf{x}} \quad \text{unter der Bedingung} \quad \tilde{\mathbf{A}}\tilde{\mathbf{x}} \leq \tilde{\mathbf{b}} \quad (\text{A.5})$$

gebracht werden. Hierfür wird die erste Bedingung von (A.4) umgeformt auf

$$\mathbf{y} - \mathbf{A}\mathbf{x} \leq \mathbf{r} \quad \text{und} \quad -\mathbf{y} + \mathbf{A}\mathbf{x} \leq \mathbf{r}. \quad (\text{A.6})$$

Durch umstellen dieser beiden Ungleichungen erhält man

$$-\mathbf{A}\mathbf{x} - \mathbf{r} \leq -\mathbf{y} \quad \text{und} \quad \mathbf{A}\mathbf{x} - \mathbf{r} \leq \mathbf{y}. \quad (\text{A.7})$$

Schreibt man die zweite Bedingung von (A.4) als $-\mathbf{r} \leq \mathbf{0}$ so erhält man mit (A.7) in Matrixnotation

$$\underbrace{\begin{bmatrix} -\mathbf{A} & -\mathbf{E} \\ \mathbf{A} & -\mathbf{E} \\ \mathbf{0} & -\mathbf{E} \end{bmatrix}}_{=\tilde{\mathbf{A}}} \underbrace{\begin{bmatrix} \mathbf{x} \\ \mathbf{r} \end{bmatrix}}_{=\tilde{\mathbf{x}}} \leq \underbrace{\begin{bmatrix} -\mathbf{y} \\ \mathbf{y} \\ \mathbf{0} \end{bmatrix}}_{=\tilde{\mathbf{b}}}, \quad (\text{A.8})$$

womit die Nebenbedingungen auf die Form (A.5) gebracht wurden. Die Zielfunktion wird entsprechend (A.4) formuliert, wobei für \mathbf{x} um Nullen erweitert wird

$$\tilde{\mathbf{f}}^\top \tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{0}^\top & \mathbf{1}^\top \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{r} \end{bmatrix}. \quad (\text{A.9})$$

Mit (A.8) und (A.9) entspricht nun das LP der geforderten Form (A.5) und kann mit Matlab gelöst werden

```

%% LINPROG
[m,n] = size(A);    % Bestimme Dimensionen

f_LP = [zeros(n,1); ones(m,1)];

A_LP = [-A          -eye(m); ...
        A           -eye(m); ...
        zeros(m,n) -eye(m)];

b_LP = [-y; y; zeros(m,1)];

```

A. Zusatzmaterial

```
options = optimset('LargeScale', 'off', 'Simplex', 'on', 'Display', 'off');  
xr_LP = linprog(f_LP, A_LP, b_LP, [], [], [], [], [], options);  
  
xLINPROG = xr_LP(1:n);
```

Bei `linprog` kann der Simplex-Algorithmus explizit als Lösungsalgorithmus angegeben werden. Beim Simplex-Algorithmus wird kein Startwert angegeben, sondern mit einem Ersatz-LP eine zulässige Basislösung für den Startwert bestimmt (siehe [50, Kap. 3.3]).

Ein Vergleich der drei Methoden für die Small-Scale-Problems dieser Arbeit hat gezeigt, dass `fminsearch` die schnellste Methode ist, aber das Ergebnis vom Startwert abhängt. `linprog` benötigt doppelt so viel Zeit wie `fminsearch` und `yalmip` 20mal so viel Zeit wie `fminsearch`. `yalmip` ist nur eine Methode zur Formulierung von Optimierungsproblemen und für deren Lösung wird auf Matlab-Methoden zurückgegriffen. Daher ist zu vermuten, dass `yalmip` in weitere Folge `linprog` als Lösungsmethode verwendet und der Overhead für die lange Rechenzeit verantwortlich ist. Bei `yalmip` und `linprog` sind die Ergebnisse identisch.

Wird die Matrix **A** um eine Null-Spalte erweitert, so liefert `yalmip` für den dazugehörigen Parameter im Vektor **x** NaN und `linprog` liefert 0. Für diesen Fall liefert `fminsearch` ein anderes Ergebnis als beim ursprünglichen Problem.

Für diese Arbeit wird `linprog` verwendet, da diese Methode 10mal so schnell ist wie `yalmip`, robuster ist als `fminsearch`, und mit Sicherheit den Simplex-Algorithmus verwendet, der am ehesten in der Industrie zur Verfügung steht.