

Deliverable D7

Beschreibung der Forecasts und Flexibilitätspotentiale

Beschreibung der Forecasts zur Berechnung der Flexibilitätspotentiale: Analyse für jede Technologie

Version 1.4

Deliverable

Georgios Chasparis, Mario Pichler, Software Competence Center Hagenberg GmbH

Patrick Summerer, World Direct

Carlo Corinaldesi, Energy Economics Group – TU Wien

Diewald Nicole, Fronius International GmbH

Mark Stefan, Thomas Brunnhofer, AIT Austrian Institute of Technology



Energieforschungsprogramm - 4. Ausschreibung

Klima- und Energiefonds des Bundes – Abwicklung durch die Österreichische Forschungsförderungsgesellschaft FFG

INTERNE REFERENZIERUNG

- **Deliverable Nr.:** D7
- **Deliverable Name:** Beschreibung der Forecasts und Flexibilitätspotentiale für jeden Use Case und für jede Technologie sowie unter Berücksichtigung der Eigeninteressen
- **Lead Partner:** Software Competence Center Hagenberg GmbH
- **Work Package Nr.:** WP3
- **Task Nr. & Name:** Task 3.2 / Flexibilitäts-Forecast
- **Dokument (File):** FLEX+_DeliverableD7-v0.8.docx
- **Speicher Datum:** ~~2020-06-25~~2020-02-04

DOKUMENT SENSIBILITÄT

- Öffentlich**
- Konsortium und ausgewählte Review Partner**
- Ausschließlich Konsortialpartner**
- Vertraulich zwischen ausgewählten Projektpartnern**

ÜBERARBEITUNGSVERLAUF

Version	Datum	Autor	Änderung
0.1	15.9.2019	SCCH	Struktur
0.2	02.10.2019	EEG	E-car Nutzung
0.3	21.10.2019	SCCH	Unflexible Stromlast
0.4	24.10.2019	WD	Warmwassernutzung
0.5	29.10.2019	Fronius	PV-Erzeugung
0.6	04.11.2019	SCCH	Unflexible Stromlast
0.7	08.11.2019	SCCH	Unflexible Stromlast
0.8	12.11.2019	SCCH	Unflexible Stromlast
0.9	16.12.2019	AIT	Strompreis- & Warmwasserprognose
1.0	18.12.2019	SCCH	kleine Änderungen
1.1	09.01.2020	SCCH	Kleine Änderungen
1.2	10.01.2020	EEG	E-car Nutzung
1.3	10.01.2020	SCCH	Einführung
1.4	08.04.2020	SCCH	Titel ändern

DOKUMENTEN STATUS

	Datum	Person(en)	Organisation
Autor(en)			
Verifizierung			
Freigabe			

ZUSAMMENFASSUNG

In den letzten Jahren wird die Relevanz von flexiblen Prosumern in Energiemärkten auf europäischer Ebene untersucht. Die aktive Beteiligung mehrerer fernsteuerbarer Prosumer-Komponenten wie Wärmepumpen, Speichersysteme, Boiler, Photovoltaik und E-Mobilität ermöglicht den einzelnen Prosumern unterschiedliche systemdienliche Dienstleistungen zu erbringen. Im Flex+ Projekt werden dazu skalierbare Optimierungsalgorithmen entwickelt, die unter Berücksichtigung der wirtschaftlichen und nicht-wirtschaftlichen Interessen aller Teilnehmer eine optimale Nutzung und Vermarktung der vorhandenen Flexibilität ermöglichen.

Damit die vorhandene Flexibilität jedoch optimal genutzt werden kann, ist es notwendig, dass wir die Merkmale und Faktoren, die sie beeinflussen, über einen kurzen Zeithorizont (ein oder zwei Tage im Voraus) genau vorhersagen. Zu diesem Zweck werden in diesem Dokument Prognosemodelle erarbeitet, mit denen die Entwicklung mehrerer wichtiger Merkmale, die sich direkt auf die verfügbare Flexibilität auswirken, vorhergesagt werden kann. Insbesondere beschreiben wir ausführlich die Parameter, die prognostiziert werden müssen, und die für die Gesamtoptimierungsleistung entscheidend sind. Darüber hinaus stellen wir eine detaillierte Beschreibung der Methoden zur Ableitung dieser Prognosen sowie deren Leistungsvermögen zur Verfügung.

INHALTSVERZEICHNIS

1	EINFÜHRUNG	7
2	EXTERNE PROGNOSEN	8
2.1	Solareinstrahlung und PV-Erzeugung	8
2.2	Strom- und Regelenergiepreise	8
3	E-CAR-NUTZUNG	10
4	UNFLEXIBLE STROMLAST	12
4.1	Datenbeschreibung und –aufbereitung	12
4.2	Methodik	13
4.2.1	Persistenzmodelle (“copy-last-days”)	14
4.2.2	Auto-Regressive (AR)-basierte Modelle	14
4.2.3	Kombinierte Auto-Regressive (AR)- und Persistenz-Modelle	15
4.2.4	Training und Implementierung von Auto-Regressive (AR)-basierten Modellen	16
4.2.5	Feature-Extraktionsmodell	16
4.2.6	Holt-Winters Modell	17
4.2.7	SARIMA Modell	18
4.3	Ergebnisse	18
4.3.1	Persistenzmodelle (“copy-last-days”)	18
4.3.2	Kombinationen von Persistenz (“copy-last-days”) und Auto-Regressive (AR)-basierten Modellen	20
4.3.3	Holt-Winters Modell	21
4.3.4	SARIMA Modell	23
4.3.5	Feature-Extraktionsmodell	24
4.3.6	Gesamtleistungsvergleiche.....	25
5	WARMWASSERNUTZUNG	26
5.1	Datenbeschreibung und –aufbereitung	26
5.2	Methodik	27
5.2.1	Fehlermetriken	28
5.2.2	Zeitreihenanalyse	29
5.2.3	Exponential Smoothing.....	30
5.2.4	SARIMA	31
5.2.5	LSTM	31
5.3	Evaluierung	32
5.3.1	Evaluierungsstrategie.....	32

5.3.2 Ergebnisse	34
5.4 Erweiterung der Methodik: LSTM und FFNN	36
5.4.1 Prognose – 24-Stunden Warmwasserverbrauch	36
5.4.2 Error- und Performance-Metric	37
5.4.3 Trainings- und Testdaten	38
5.4.4 Ergebnisse	39
6 REFERENZEN.....	43
7 ANHANG	44
7.1 Abbildungsverzeichnis.....	44
7.2 Tabellenverzeichnis	45

1 Einführung

Eines der Hauptziele des Flex+-Projekts ist die Entwicklung genauer Prognosemodelle zur Erzeugung von Vorhersagen über jene Parameter, die für den Betrieb der Komponenten (Batteriespeicher, Wärmepumpen, Boiler und Elektroautos) entscheidend sind. Solche Vorhersagemodelle werden direkt bei der Ableitung der Optimierungsstrategien für das Flexibilitätsangebot in den Strommärkten verwendet. Daher kann die Vorhersagegenauigkeit der abgeleiteten Modelle die Leistung der Optimierung erheblich beeinflussen.

In diesem Ergebnisdokument beschreiben wir detailliert die Parameter, die prognostiziert werden müssen, um die Flexibilitätspotenziale für jeden Use-Case und für jede Technologie (Batteriespeicher, Wärmepumpen, Boiler und Elektroautos). Diese Prognosen werden dann direkt in die Optimierung zur Berechnung der optimalen Flexibilitätsangebot in den Strommärkten einbezogen. Da die Berechnung des Flexibilitätspotenzials direkt von der Modellierung jeder in der Optimierung verwendeten Technologie abhängt (die detailliert in Deliverable D9 dargestellt wird), stellen wir hier nur die Prognosen der wichtigen Parameter für jede Technologie dar und nicht direkt die Prognose des gesamten Flexibilitätspotenzials.

Darüber hinaus geben wir eine detaillierte Beschreibung der Methoden zur Ableitung dieser Prognosen sowie deren Leistungsvermögen. Im weiteren Verlauf des Dokuments werden in Kapitel 2 die externen Prognosen zusammengefasst, die außerhalb des Betriebs der Komponenten liegen aber indirekt die Optimierung der Flexibilität beeinflussen. In den Kapiteln 3, 4 und 0 beschreiben wir die Prognosemethodik für die Nutzung des Elektroautos, den unflexiblen Teil der Lasten und die Warmwassernutzung.

2 Externe Prognosen

2.1 Solareinstrahlung und PV-Erzeugung

Die Solaren Einstrahlungsprognosen und die PV-Erzeugungsprognosen werden von den Projektpartnern (Fronius, IDM, etc.) durch Drittanbieter bezogen. Die Prognosen basieren auf PV Modellen, welche die Nennleistung der Anlage, sowie Modulausrichtung und Aufstellwinkel berücksichtigen. Für ein PV-System können maximal zwei unterschiedliche Ausrichtungen angegeben werden. Der PV-Forecast kombiniert die beiden Ausrichtungen dann zu einer Gesamtleistung für ein PV-System.

Für die prognostizierten Leistungswerte sind je nach Prognosehorizont unterschiedliche Granularitäten in Watt verfügbar:

- 0-24 Stunden: 15 Minuten Werte
- 24-48 Stunden: 1 Stunden Werte
- 48-216 Stunden: 6 Stunden Werte

2.2 Strom- und Regelenergiepreise

Im Rahmen von dem Projekt wurden die langfristigen Preistrends analysiert, wie in Deliverable 5 (Corinaldesi, et al., 2019) beschrieben. Die kurzfristigen Prognosen werden durch die Lieferanten und den Regelenergieanbieter bereitgestellt. Es werden die folgenden Preisprognosen für die Simulationen bzw. in weiterer Folge für die Demonstrationen verwendet:

- Der stündliche Day-ahead Spotpreis der EPEX Spot wird von allen Stromlieferanten für die nächsten 72h vorhergesagt. Die Prognose wird von einem externen Dienstleister bezogen und an die flex+ Plattform weitergeleitet. Wichtige Eingangsparameter für die Prognose sind die Preise für Kohle, Gas und Öl sowie die Kosten für CO2 Zertifikate. Eine weitere wichtige Rolle spielt die zu erwartende Einspeisung aus erneuerbaren Energien wie Wind und PV. Für die Ermittlung der Stromnachfrage werden neben der Temperatur auch die Typtagprofile und spezifische Kalenderinformationen genutzt.
- Regelleistung und Regelenergie, Abrufwahrscheinlichkeiten: Die Regelleistungs- und Regelenergiepreise sowie die Abrufwahrscheinlichkeiten werden basierend auf den historischen Daten der letzten Wochen sowohl für eine untere als auch die obere Position auf der Merit Order abgeschätzt. Dabei ist bei der unteren Position auf der Merit Order der Preis für die Regelenergie geringer, dafür ist die Abrufwahrscheinlichkeit höher, während sich an den oberen Positionen die niedrigen Abrufwahrscheinlichkeiten mit hohen Preisen befinden. Der optimierte Fahrplan kann Angebote für diese zwei Positionen miteinbeziehen.
- Durchschnittlicher Intraday-Preis „ID3“: Intraday wird, im Gegensatz zu den Day-ahead Spot Preisen, pay-as-bid gehandelt, das heißt, dass jeder den Preis bekommt, den er angeboten hat. Dementsprechend können abhängig vom Zeitpunkt des Trades die Preise variieren. Ein typischer Preis für die Bewertung von Intraday sind die ID3-Preise, also der Durchschnitt aller Angebote der letzten 3 Stunden. Diese Preise werden von den Energielieferanten für die nächsten 3 Stunden abgeschätzt.
- Rollierende Bid/Ask-Intraday-Preise: Außerdem werden die realen rollierenden, sich stündlich verändernden Ask- und Bid-Preise für den Intraday-Preis für jede Viertelstunde bzw. Stunde für die nächsten

drei Stunden bereitgestellt. Ab dann sind nur noch Prognosen verfügbar, wobei die Intraday-Prognose dem Day-Ahead-Forecast entspricht.

3 E-Car-Nutzung

Mit der stetig wachsenden Marktdurchdringung von Elektrofahrzeugen und der damit einhergehenden steigenden Nachfrage elektrischer Energie, gewinnen die Verbrauchsprognosen von E-Autos-Ladevorgängen an Bedeutung. Die Auswirkung des Ladens von Elektrofahrzeugen auf das elektrische Netz wird nicht nur von der gesamten benötigten Energie und der maximalen erlaubten Leistung bestimmt, sondern auch von der Zeitspanne und vom geographischen Ort, wo die Ladevorgänge stattfinden.

Elektrofahrzeuge erzeugen zu bestimmten Zeitpunkten an bestimmten geographischen Punkten des Netzes einen konzentrierten Leistungsbedarf. Um jederzeit eine hohe Anzahl von Ladevorgängen gewährleisten zu können, wäre heute ein bedeutender Ausbau des elektrischen Netzes notwendig (Pollok, Hille, & Schnettler, 2009). Zeitliche Prognosen von E-Autos-Ladevorgängen haben aus diesem Grund ein großes Potential den Ausbau des Stromnetzes und den damit verbundenen Investitionen zu reduzieren, denn sie würden ein effizienteres Management der Leistungsflüsse auf Verteilungsebene ermöglichen.

Die traditionellen Verbrauchsprognosen von kleinen Verbrauchern basieren hauptsächlich auf dem Wetterfaktor (wie Temperatur und Luftfeuchtigkeit). Verbrauchsprognosen von E-Autos-Ladevorgängen sind komplizierter, denn sie sind auch von Benutzeraktionen stark abhängig.

Einige Studien analysieren das Ladeverhalten von Elektrofahrzeugen und die daraus resultierenden Lastprofile unter Berücksichtigung der Unsicherheiten und stochastischen Eigenschaften. (Sokorai & A. Fleischhacker, 2018) entwickelt ein auf einer Markov-Chain basierendes Tool, das die stochastische Natur der täglichen Nutzung einer Ladestation modellieren kann. Das Tool gibt als Ergebnis die Wahrscheinlichkeitsdichtefunktionen zum Laden von Elektrofahrzeugen an, die in Abbildung 1 dargestellt sind.

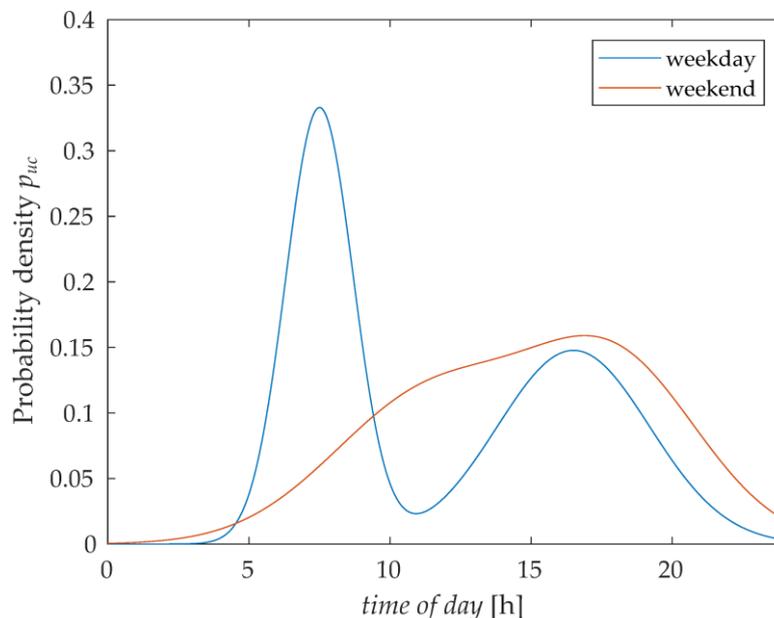


Abbildung 1 Wahrscheinlichkeitsdichtefunktionen von Plug-In-Zeiten.

Im Artikel (Zhuowei, 2011), wurde das Monte-Carlo-Simulationsverfahren angewendet, um die Startzeit des Ladevorganges basierend auf Wahrscheinlichkeitsverteilungen zu bestimmen. Diese Studie unterscheidet zwischen mehreren Verbrauchern: Elektrobussen, Taxis, Dienstwagen und Privatwagen. Wie in Abbildung 2 gezeigt, unterscheiden sich die Ladeverhalten von den Verbrauchern erheblich.

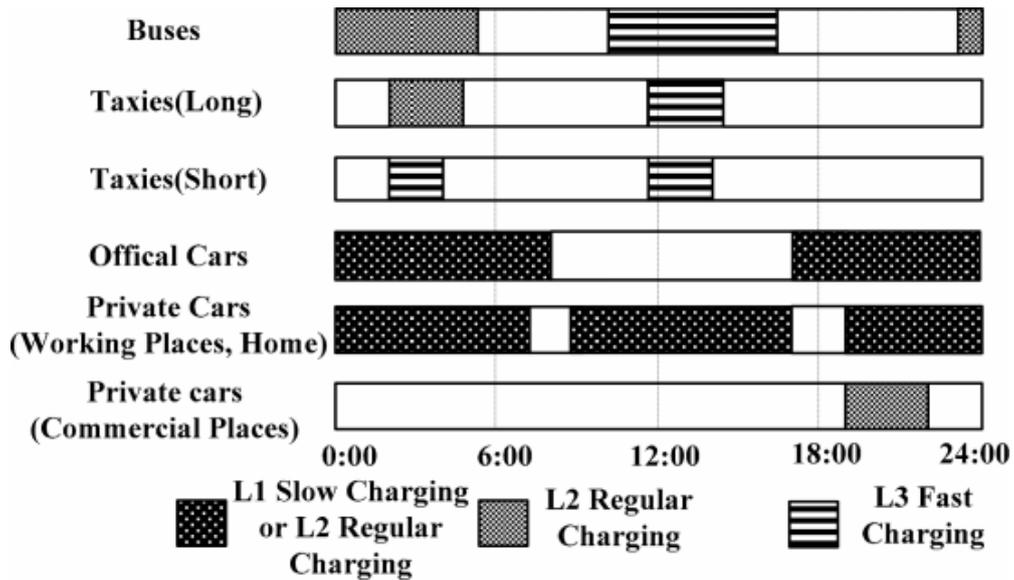


Abbildung 2 Ladezeiten der unterschiedlichen Verbraucher

Studien, wie (Feixiang, 2011) untersuchen die Eigenschaften des Verbraucherverhalten an den Ladestationen durch stochastischen Modellen wie neuronalen Backpropagation-Netze. Durch diese Methode wurde eine Eigenschaft der Ladevorgänge festgestellt: Der Verbrauch am Prognosestag ist der Verbrauch am selben Tag der letzten Woche sehr ähnlich. Es wird hier gezeigt, dass unter Berücksichtigung weiterer Einflussfaktoren (wie Wetter und Temperatur) der Prognosen Fehler dann erheblich verringert werden kann.

Genau und angemessene Lastprognosemethoden zur Beschreibung der Ladeverhalten von Elektrofahrzeugen stellen heutzutage eine wichtige Garantie für die Planung und den Betrieb des zukünftigen Versorgungssystems dar und werden unvermeidlich eine wichtige Rolle bei der Automatisierung des elektrischen Systems spielen.

4 Unflexible Stromlast

4.1 Datenbeschreibung und -aufbereitung

Messungen des unflexiblen Lastverbrauchs jedes Haushalts wurden mit einer Granularität von 15 Minuten gesammelt. Damit stehen uns an jedem Tag 96 sequentielle Messungen der unflexiblen Last aus jedem Haushalt zur Verfügung. Jede dieser Messungen zum Zeitpunkt t stellt den Lastverbrauch während des letzten 15min-Intervalls dar, d.h. den Energieverbrauch innerhalb des letzten 15min-Intervalls. In der folgenden [Abbildung 3](#) und [Abbildung 4](#) stellen wir eine Stichprobe des gemessenen Lastverbrauchs von einem Haushalt innerhalb eines Monats bzw. eines Tages dar.

Formatie

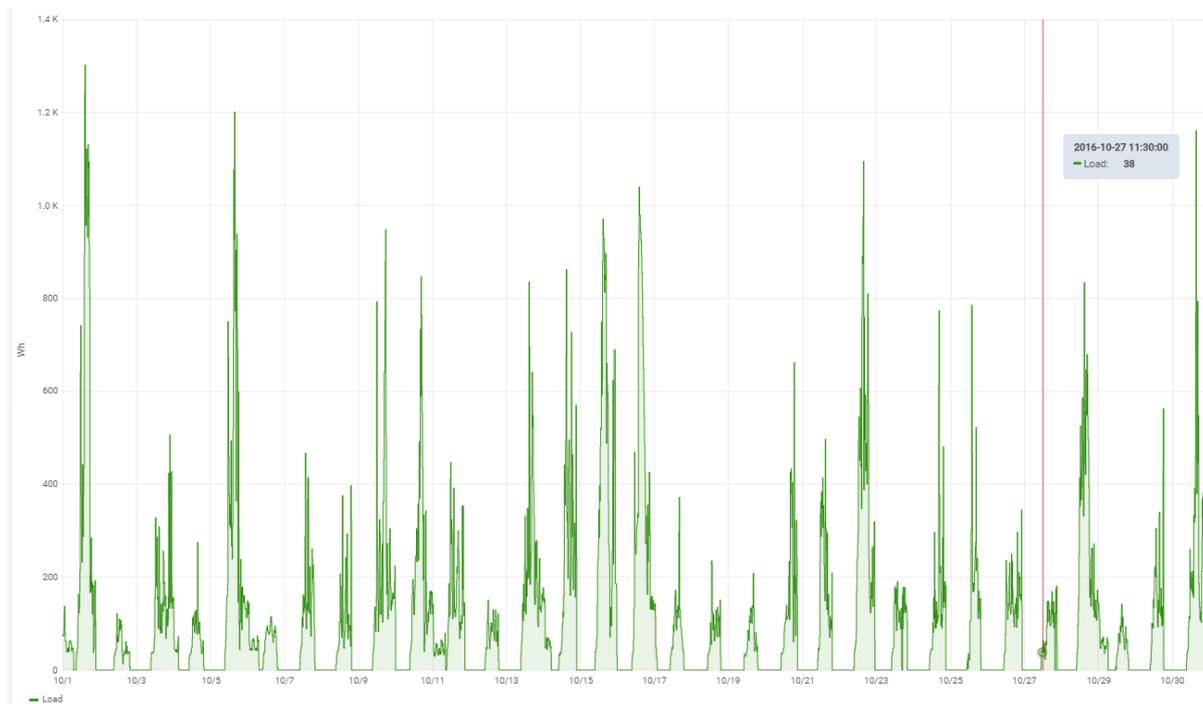


Abbildung 3 Unflexibler Lastverbrauch eines Haushalts in einem Monat (von 01.10.2016 bis 31.10.2016).

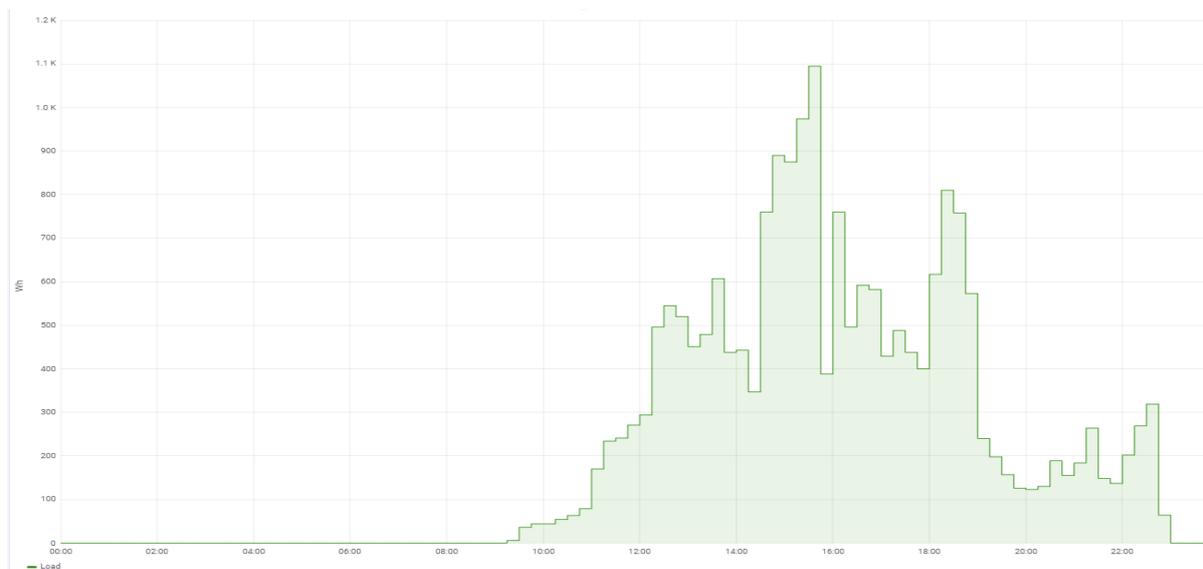


Abbildung 4 Unflexibler Lastverbrauch eines Haushalts an einem Tag (22.10.2016).

Solche Messungen, die über den Zeitraum von einem Jahr oder mehr gesammelt werden, reichen in der Regel aus, um ein Prognosemodell zu erstellen. Ein solches Prognosemodell wird verwendet, um Prognosen der Last für den nächsten Kalendertag zu erstellen, die es uns dann ermöglichen, die Flexibilitätsnutzung zu optimieren.

4.2 Methodik

In diesem Unterabschnitt beschreiben wir im Detail die wichtigsten Methoden, die zur Vorhersage der unflexiblen Lasten verwendet werden. Unter unflexiblen Lasten verstehen wir alle elektrischen Lasten außer der Last der steuerbaren Komponenten (Wärmepumpe, Boiler, Elektroauto). Unser Hauptaugenmerk liegt auf der Entwicklung von Prognosemethoden, die a) rechentechnisch effizient (ermöglicht eine Ad-hoc-Berechnung der Prognosen auch in Echtzeit) und b) zuverlässig in Bezug auf die abgeleiteten Prognosen sind. Zu diesem Zweck haben wir die folgenden wichtigsten Methoden untersucht:

- a) Persistenzmodelle
- b) Auto-Regressive (AR)-basierte Modelle
- c) Kombinationen von Persistenz und Auto-Regressive (AR)-basierten Modellen
- d) Feature-Extraktionsmodell
- e) Holt-Winters Modell
- f) SARIMA Modell

Die erste Klasse von Modellen (Persistenzmodelle) verwendet das zuletzt beobachtete Verhalten über einen Zeitraum der letzten Tage. Einfach gesagt, Persistenzmodelle versuchen, das beobachtete Verhalten über mehrere (letzte) Tage zu erfassen. Die zweite Klasse von Modellen (Auto-Regressive-basierte Modelle) verwendet das zuletzt beobachtete Verhalten während des aktuellen Tages (der letzten Stunden), um eine Vorhersage für die nächsten Stunden zu erstellen. Kurz gesagt, Auto-Regressive-basierte Modelle erfassen zeitliche Abhängigkeiten von zuvor beobachtetem Verhalten innerhalb eines Tages. Informell können wir sagen, dass Persistenzmodelle versuchen, niedrige Frequenzen in den Profilen der nicht-flexiblen Last zu erfassen, während Auto-Regressive-basierte Modelle versuchen, höhere Frequenzen in den Profilen der nicht-flexiblen Last zu erfassen.

Eine wichtige Beobachtung bei der Untersuchung dieser Methoden ist, dass eine Kombination der beiden Methoden (a) und (b) genauere Vorhersagen über die unflexiblen Lasten liefern sollte. Wir möchten sowohl die nieder- als auch die hochfrequenten Phänomene erfassen, die auftreten können. Diese Notwendigkeit wird auch dadurch begründet, dass der große Zeithorizont der angeforderten Prognosen (in der Regel 1 Tag im Voraus) zusätzliche Herausforderungen (z.B. seltene Ereignisse) mit sich bringt, die mit keiner der beiden Methoden allein erfasst werden können. Wir untersuchen die Kombination der beiden Modelle in der Modellkategorie (c).

Diese Idee (der Kombination alternativer Modelle und Faktoren innerhalb eines einzigen Modells) wurde in der Modellkategorie (d) weiter genutzt; diese versucht, eine größere Anzahl von Merkmalen (d.h. eine größere Variation von nieder- und hochfrequenten Persistenz Faktoren) zu integrieren und optimal unter ihnen auszuwählen. Schließlich stellen die letzten beiden Modellkategorien (e) und (f) zwei eher standardisierte Ansätze dar, nämlich die Holt-Winters und die SARIMA-Prognosemodelle.

Unser Ziel ist es, einen Vergleich zwischen dieser großen Sammlung von Modellen durchzuführen und den Unterschied zwischen den Modellen (c) und (d) zu identifizieren, die speziell auf die Lastprognose zugeschnitten sind, im Vergleich zu generischen Black-Box-Modellen, wie den Modellen (e) und (f). In den folgenden Unterabschnitten stellen wir die Details der oben genannten Methoden vor.

4.2.1 Persistenzmodelle ("copy-last-days")

Persistenz-Prognosemodelle werden in der Regel verwendet, um Referenzmodelle (Baseline) zu erstellen, die dann für Vergleichstests verwendet werden können. In vielen Fällen ist es von Vorteil zu wissen, ob ein entwickeltes Prognosemodell bessere Vorhersagen liefern kann als ein solches Referenz- oder Basismodell. Persistenzmodelle gehören zu den "trivialsten" und basieren auf dem Prinzip, dass "die Dinge gleich bleiben", d.h. die Prognose ist immer gleich dem letzten bekannten Datenpunkt.

Gemäß Referenz (Notton & Voyant, 2018) würde ein Persistenzmodell davon ausgehen, dass die (unflexible) Last zum Zeitpunkt $t + 1$ gleich der Last zum Zeitpunkt t ist. Wie genau sollten jedoch die Zeitinstanzen $t + 1$ und t definiert sein? Bei der (unflexiblen) Lastprognose müssen wir eine Lastprognose über 24 Stunden im Voraus (d.h. über den nächsten Tag) erstellen. Natürlich würde in diesem Zusammenhang ein Persistenzmodell, das davon ausgeht, dass die Last am nächsten Tag konstant bleibt und dem aktuellen entspricht, höchstwahrscheinlich die zeitlichen Schwankungen im Lastprofil nicht gut vorhersagen können. Stattdessen wäre ein Persistenzmodell genauer, wenn es davon ausgeht, dass die (unflexible) Last zur Zeit t des Tages $d + 1$ (kurz $(t, d + 1)$) gleich wie die Last zur gleichen Zeit t des Vortages oder des Vortages d wäre. Eine zusätzliche Variation dieses Modells würde auch mehr als einen Vortag berücksichtigen (z.B. den durchschnittlichen Verbrauch zur gleichen Zeit an N Vortagen).

Zu diesem Zweck werden wir in der kommenden Evaluierung ein Persistenzmodell betrachten, das N letzte gleiche Wochentage verwendet. Wenn wir beispielsweise Prognosen für den nächsten Tag erstellen müssen, und dieser Tag ist ein Montag, dann entspricht diese Prognose dem durchschnittlichen Lastverbrauch der letzten N Montage zur gleichen Zeit. Wir werden die Abkürzung CLD verwenden, die für "copy-last-days" steht, um auf diese Modelle zu verweisen.

Formal gesehen bezeichnet $l_d(t)$ die (unflexible) Last eines Haushalts zum Zeitpunkt t am Tag d . Dann geht das 1-Tages-Persistenzmodell davon aus, dass

$$\hat{l}_d^{PM}(t) = l_{d-1}(t)$$

Analog dazu können wir das N -Tages-Persistenzmodell wie folgt definieren:

$$\hat{l}_d^{PM}(t) = \frac{1}{N} \sum_{i=d-N}^{d-1} l_i(t)$$

Mit anderen Worten, das N -Tages-Persistenzmodell nimmt die durchschnittliche Last von N Vortagen zur gleichen Zeit. Da die Stromlast jedoch stark mit der Anwesenheit der Menschen in einem Haushalt (d.h. mit dem Tagesablauf der Menschen) korreliert, können wir das oben genannte N -Tages-Persistenzmodell weiter verbessern, indem wir nur die N vorherigen gleichen Tage berücksichtigen. Wenn $d + 1$ also einem "Montag" entspricht, dann müssen wir, um unsere Prognose für den Zeitpunkt t zu erstellen, die durchschnittliche Last zur gleichen Zeit an den letzten N vorhergehenden Montagen errechnen. Wir bezeichnen dieses Modell als das N -Gleiche-Tage-Persistenzmodell, nach dem die Prognosen wie folgt berechnet werden

$$\hat{l}_d^{PM}(t) = \frac{1}{N} \sum_{i=d-7N}^{d-7} l_i(t)$$

4.2.2 Auto-Regressive (AR)-basierte Modelle

Wie wir im vorherigen Unterabschnitt gesehen haben, versuchen die diskutierten Persistenzmodelle, zeitliche Abhängigkeiten zu erfassen, deren Häufigkeit sich über mehrere Tage oder Wochen erstreckt (z.B. aufgrund von Ähnlichkeiten im Tagesablauf der Bewohner an denselben Tagen). Wir werden solche Abhängigkeiten als niederfrequente zeitliche Abhängigkeiten bezeichnen.

Es kann aber auch zeitliche Abhängigkeiten der Stromlast innerhalb desselben Tages geben. Es ist demnach sehr wahrscheinlich, dass die zur Tageszeit t gemessene Last von der zur vorherigen Zeitinstanz $t - 1$ desselben Tages gemessenen Last abhängt, vor allem wenn das Zeitintervall zwischen diesen beiden Zeitinstanzen eher klein ist (hierbei gehen wir von Zeitintervallen von 15min aus). Wir werden solche zeitlichen Abhängigkeiten der unflexiblen Last als hochfrequente zeitliche Abhängigkeiten bezeichnen.

Mit Hilfe von Auto-Regressive (AR) Prognosemodellen können solche (hochfrequenten) zeitlichen Abhängigkeiten der Last innerhalb eines Kalendertages erfasst werden. Das wahrscheinlich einfachste derartige Modell ist das Auto-Regressive Modell (kurz AR-Modell), nach dem die Vorhersage der Last zum Zeitpunkt $t + 1$ durch eine lineare Kombination der Last zu früheren Zeitpunkten gegeben ist. Es gilt also:

$$\hat{l}_d^{AR}(t) = a_1 l_d(t - 1) + a_2 l_d(t - 2) + \dots + a_n l_d(t - n)$$

(das sich aus einem Maximum-a-posteriori Prädiktor eines ursprünglichen Weißrausch-Störprozesses ergibt, vgl. (Ljung, 1999), Kapitel 4).

Falls wir Vorhersagen über mehrere Zeitinstanzen im Voraus (in der Regel einen Tag im Voraus) erstellen möchten, dann können wir eine Variation des obigen Modells implementieren, die in der Regel als Pseudo-Regressionsmodell bezeichnet wird und die folgende Form annimmt:

$$\hat{l}_d^{AR}(t|a_1, \dots, a_n) = a_1 \hat{l}_d^{AR}(t - 1) + a_2 \hat{l}_d^{AR}(t - 2) + \dots + a_n \hat{l}_d^{AR}(t - n)$$

Mit anderen Worten, wenn die unflexible Last zum Zeitpunkt $t - j$, $j = 1, \dots, n$, nicht bekannt ist, wird sie durch die für diesen Zeitpunkt verfügbare Vorhersage ersetzt.

Beachten Sie, dass alternative Auto-Regressive Modelle definiert werden können. Beispielsweise werden häufig Modelle verwendet, die auch Moving-Average (MA) Rausch-Terme enthalten. Solche Moving-Average Terme können verwendet werden, um die Auswirkungen niederfrequenten Rauschens im Profil zu erfassen. Unsere Untersuchung zeigt jedoch, dass der Gewinn an Vorhersagegenauigkeit durch das Hinzufügen solcher Terme vernachlässigbar ist.

Einer der Hauptnachteile solcher Methoden ist die Tatsache, dass sie eher für kurzfristige Zukunftsprognosen (im Bereich von wenigen Stunden) geeignet sind. Tatsächlich ist es einfach zu erkennen, dass sich selbst kleine Prognosefehler in den kurzfristigen Zukunftsprognosen unvorhersehbar ausbreiten können, wenn man langfristige Zukunftsprognosen formuliert, die sich über einen Tag im Voraus erstrecken. Aus diesem Grund stellen wir im nächsten Unterabschnitt eine Kombination aus dem im vorherigen Abschnitt beschriebenen N-Gleiche-Tage-Persistenzmodell und dem oben beschriebenen AR-basierten Modell vor.

4.2.3 Kombinierte Auto-Regressive (AR)- und Persistenz-Modelle

Wie bereits erwähnt können die Persistenzmodelle niederfrequente zeitliche Abhängigkeiten im Lastprofil (über mehrere Tage oder Wochen) erfassen, während Auto-Regressive-Modelle dies für hochfrequente zeitliche Abhängigkeiten (innerhalb desselben Kalendertages) können. Darüber hinaus funktionieren Auto-Regressive-Modelle nur innerhalb eines kurzfristigen Zeithorizonts gut. Zu diesem Zweck möchten wir in diesem Unterabschnitt auch die Möglichkeit einer optimalen Kombination der beiden Modelltypen berücksichtigen. Im Prinzip passt diese Idee gut zu den in (Cesa-Bianchi & Lugosi, 2006) diskutierten expertenbasierten Prognosemethoden und Transfer-Lernmethoden wie (Grubinger, Chasparis, & Natschlaeger, 2017).

Kurz gesagt setzt solch eine optimale Kombination der beiden Prognoseverfahren eine kombinierte Vorhersage der folgenden Form voraus:

$$\hat{l}_d^{AR}(t|a_1, \dots, a_n, b_0) = a_1 \hat{l}_d^{AR}(t - 1) + a_2 \hat{l}_d^{AR}(t - 2) + \dots + a_n \hat{l}_d^{AR}(t - n) + b_0 \hat{l}_d^{PM}(t)$$

In diesem Fall möchten wir eine neue Menge von Gewichten $a_1, a_2, \dots, a_n, b_0$ berechnen, die die optimale Kombination aus den hochfrequenten zeitlichen Abhängigkeiten (erfasst durch

die auto-regressiven Terme) und den niederfrequenten zeitlichen oder saisonalen Abhängigkeiten (erfasst durch den abschließenden Persistenzterm) finden.

4.2.4 Training und Implementierung von Auto-Regressive (AR)-basierten Modellen

In diesem Unterabschnitt möchten wir auf weitere Details zur Formulierung der erforderlichen unflexiblen Lastprognosen (für den nächsten Tag) sowie zum Trainieren dieser Modelle eingehen.

Nehmen wir an, wir haben den unflexiblen Lastverbrauch eines Haushalts über die Dauer von $d - 1 > 0$ Tagen gemessen und Messungen in Intervallen von jeweils 15min gesammelt. Infolgedessen verfügen wir an jedem dieser Tage über 96 sequentielle Lastmessungen. Jede dieser Messungen zum Zeitpunkt t stellt den Lastverbrauch während des letzten 15min Intervalls dar. Anhand der verfügbaren Messungen über alle vorangegangenen $d - 1$ Tage möchten wir den Lastverbrauch über den nächsten Tag d vorhersagen. Schematisch ist dies in der folgenden Abbildung 5 dargestellt.

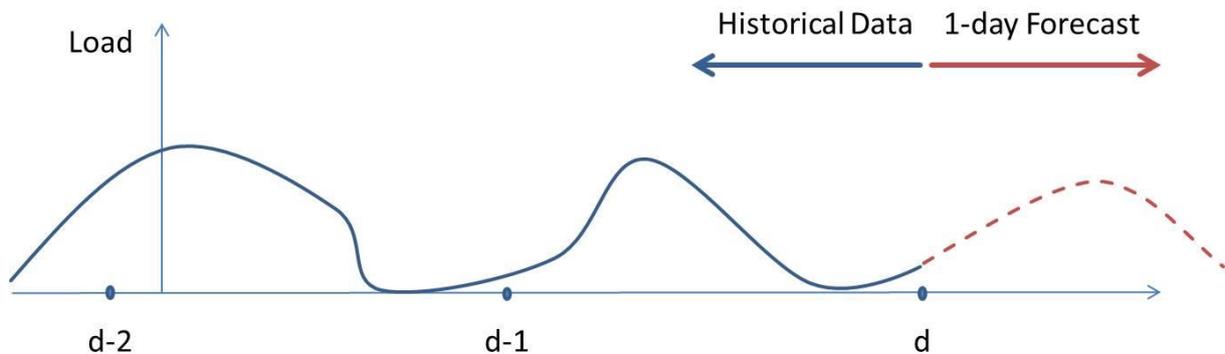


Abbildung 5 Schematische Lastprognose unflexibler Verbraucher für den nächsten Tag.

Im Falle des Persistenzmodells von Abschnitt 4.2.1 berechnen wir für jedes der 15min Intervalle des nächsten Tages einfach den Durchschnitt des Lastverbrauchs des gleichen Intervalls an den N gleichen Vortagen.

Im Falle eines AR-basierten Modells müssen wir zunächst die optimalen Parameter (oder Gewichte) des Modells berechnen. Die optimale Berechnung dieser Gewichte wird als Regressionsproblem der folgenden Form formuliert:

$$\min_{a_1, a_2, \dots, a_n} \sum_{t=1}^{T=96(d-1)} \lambda^{T-t} |\hat{l}_d^{AR}(t|a_1, \dots, a_n) - l_d(t)|^2$$

Es minimiert den Vorhersagefehler über alle vorangegangenen Zeitintervalle. Der Parameter λ entspricht einem Gewichtungsfaktor, der den aktuellen Messungen mehr Gewicht verleiht (wodurch wir auch saisonale Abhängigkeiten des Lastverbrauchs erfassen können). In dieser Implementierung setzen wir ihn auf $\lambda = 0.999$. Aus berechnungstechnischen Gründen lösen wir das Problem der linearen Regression mit einer Recursive-Least-Squares (RLS) Implementierung (vgl. Abschnitt 12.3 von (Sayed, 2003)).

4.2.5 Feature-Extraktionsmodell

Eine der Schlussfolgerungen aus der Kombination der von uns in Abschnitt 4.2.3 beschriebenen Auto-Regressive und Persistenzmodelle ist die Tatsache, dass es in historischen Daten (Vortage) Informationen gibt, welche die Prognosegenauigkeit signifikant erhöhen können. Zu diesem Zweck beschreiben wir in diesem Abschnitt eine weitere Variante von Auto-Regressive und Persistenzmodellen, die versuchen, diese historischen Merkmale besser zu verfolgen. Wir nennen dieses Modell Feature-Extraktionsmodell, da es mehrere historische Persistenzmerkmale kombiniert und versucht, ihre optimale relative Bedeutung zu finden. Wir formulieren dieses Prognosemodell als Regressionsproblem und versuchen, die relative

Bedeutung für die Lastprognose zu verstehen. Die Regressionsparameter werden nach dem in Abschnitt 4.2.4 vorgestellten Algorithmus der Recursive Least Squares trainiert. Die betrachteten Merkmalskategorien sind in der folgenden [Tabelle 1](#) dargestellt.

Tabelle 1 Feature-Kategorien im Feature-Extraktionsmodell

<i>RA</i>	Rollender Mittelwert der unflexiblen Last während des vorherigen Stundenfensters (vier 15-Minuten-Intervalle).
<i>d</i>	Werktag oder Wochenende, $d \in [0, 1]$
<i>Lh</i>	Gesamtlast innerhalb des aktuellen Stundenintervalls. Das Stundenintervall ist definiert als der Stundenzeitstempel des aktuellen 15min Intervalls.
<i>Ld</i>	Anteil der Last des aktuellen 15min Intervalls, im Vergleich zur mittleren Last des Tages.
<i>DLh</i>	Differenz in der Last innerhalb der vorherigen Stunde.
<i>LC</i>	Erkennung von niedrigem Verbrauch, wenn $Ld < 0.2$.
<i>PC</i>	Erkennung von Spitzenverbrauch, wenn $Ld > 1.5$.

Die genauen Merkmale, die als Regressoren verwendet werden, sind in der folgenden [Tabelle 2](#) dargestellt. Beachten Sie, dass wir für alle Merkmalskategorien mit Ausnahme des Tagesindex das entsprechende Merkmal vom Vortag und vom gleichen Tag in der Vorwoche verwenden.

Tabelle 2 Merkmale, die im Feature-Extraktionsmodell als Regressoren betrachtet werden und zugeordnete Gewichte.

Day	LOAD	RA	<i>d</i>	<i>Lh</i>	<i>Ld</i>	<i>DLh</i>	LC	PC
Target X	-	-	w_4	-	-	-	-	-
X-1	w_0	w_2	-	w_5	w_7	w_9	w_{11}	w_{13}
X-7	w_1	w_3	-	w_6	w_8	w_{10}	w_{12}	w_{14}

Wir erwarten, dass dieses Modell Abhängigkeiten erfasst, die auch im Persistenzmodell und in den Auto-Regressive Modellen vorhanden sind. Es versucht jedoch auch, den Gesamtverbrauch, der am Vortag/Woche aufgetreten ist, sowie das durchschnittliche Verhalten in der letzten Stunde zu verfolgen. In gewisser Weise können wir argumentieren, dass ein solches Modell reichhaltiger ist als die Kombination aus Auto-Regressive- und Persistenzmodell, die in Abschnitt 4.2.3 vorgestellt wird.

4.2.6 Holt-Winters Modell

Das Holt-Winters Prognosemodell ist in der Prognoseliteratur gut bekannt, vgl. (Szmit, Szmit, Slawomir, & Bugala, 2012). Es zerlegt den vorhergesagten Parameter in drei Komponenten, nämlich eine Niveauelemente, eine (hochfrequente) Trendkomponente und eine (niederfrequente) Saisonkomponente. Wir können argumentieren, dass es Elemente der in den vorangegangenen Abschnitten vorgestellten Modelle umfasst. So erfasst beispielsweise die Trendkomponente zeitliche Abhängigkeiten, die auch in den Auto-Regressive (AR)-basierten Modellen vorhanden sind, während die saisonale Komponente niederfrequente Abhängigkeiten erfasst, die auch in den Persistenzmodellen vorhanden sind. Die

Details zu diesem Modell finden Sie in der obigen Referenz sowie weiter hinten in Abschnitt 5.2.3. Wir verwenden dieses Modell, um die Leistung der abgeleiteten Auto-Regressive und persistenzbasierten Prognosemodelle, die in den Abschnitten 4.2.3 und 4.2.5 entwickelt wurden, besser bewerten zu können.

4.2.7 SARIMA Modell

Das ARIMA-Modell (Auto Regressive Integrated Moving Average) ist eine der gängigsten Methoden zur Analyse und Prognose von Zeitreihen. Es handelt sich um eine Erweiterung der ARMA-Modelle für nicht-stationäre Zeitreihen, die stationär gemacht werden können, indem Unterschiede einer bestimmten Ordnung von der ursprünglichen Zeitreihe - integrierte oder differenzstationäre Zeitreihen - übernommen werden.

ARIMA verwendet Hauptparameter (p, d, q) , die als ganze Zahlen ausgedrückt werden. Diese drei Parameter berücksichtigen zusammen Saisonalität (Periodizität), Tendenz und Rauschen in den Datensätzen:

- p ist die auto-regressive Ordnung, die es ermöglicht, frühere Werte der Zeitreihe zu berücksichtigen.
- d ist die Reihenfolge der Integration, die es ermöglicht, frühere Unterschiede der Zeitreihe zu berücksichtigen,
- q ist die Reihenfolge des gleitenden Mittelwerts, die es ermöglicht, den Modellfehler als lineare Kombination von zuvor beobachteten Fehlerwerten einzustellen.

Der Hauptnachteil dieses Modells ist, dass es keine saisonalen Zeitreihen unterstützt, was es unmöglich macht, damit Zeitreihen des Energieverbrauchs vorherzusagen, die durch starke Periodizität gekennzeichnet sind, wie z.B. Tages- oder Wochenzeiten.

Eine Variation des ARIMA-Modells, nämlich SARIMA, kann stattdessen verwendet werden, um auch die Periodizität zu verfolgen. In diesem Modell sind die Parameter (p, d, q) die nicht-saisonalen Parameter, die wie oben beschrieben gleich bleiben. Zusätzlich zu diesen Parametern stellen wir auch Parameter (P, D, Q) vor, die ähnlich wie (p, d, q) definiert sind, aber stattdessen für die saisonale Komponente der Zeitreihe gelten. Schließlich beschreibt der Parameter S den Zeitraum der Saison in der Zeitreihe (96, wenn die Periode einem Tag entspricht, 7×96 , wenn die Periode einer Woche entspricht, etc., wobei 96 sich auf die Granularität der Sensordaten innerhalb eines Tages bezieht). Ähnlich wie das Holt-Winters-Modell ist dies auch ein Black-Box-Modell, das auch die saisonalen Effekte erfasst und verwendet wird, um die Leistung der abgeleiteten, Auto-Regressive Prognosemodelle besser zu bewerten. Weitere Informationen zu diesem Modell finden Sie auch im Abschnitt 5.2.4 weiter unten.

4.3 Ergebnisse

In diesem Unterabschnitt stellen wir die Ergebnisse der betrachteten Prognosemethoden für die unflexiblen Lasten in Wohngebäuden vor. Insbesondere stellen wir die Leistungsfähigkeit des Persistenzmodells (Abschnitt 4.2.1), der Kombination aus Persistenzmodell und Auto-Regressive Modell (Abschnitt 4.2.3), des Feature-Extraktionsmodells (Abschnitt 4.2.5), des Holt-Winters-Modells (Abschnitt 4.2.6) und des SARIMA-Modells (Abschnitt 4.2.7) vor. Die Auswertung aller Modelle erfolgt in Bezug auf den Root-Mean-Square-Error (RMSE).

4.3.1 Persistenzmodelle ("copy-last-days")

In diesem Abschnitt stellen wir die Prognoseperformance des Persistenz-Prognosemodells vor (kurz "Copy-last-days"), wobei wir die letzten $N=3$ gleichen Tage verwenden. In [Abbildung 6](#) und [Abbildung 7](#) demonstrieren wir die One-Day-Ahead-Prognose des Persistenzmodells und der tatsächlichen unflexiblen Last über die Dauer von einem Monat bzw. einem Tag.

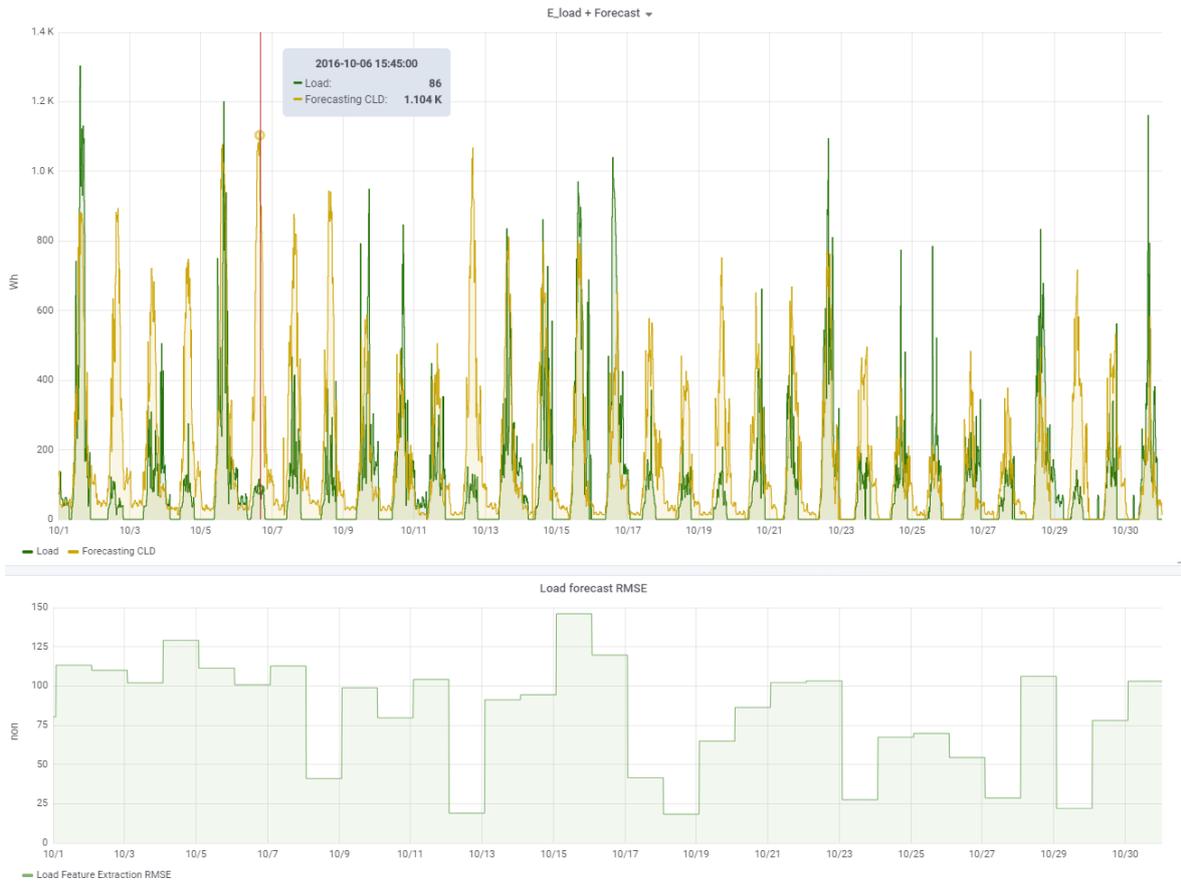


Abbildung 6 Persistenzmodell ("Copy-last-days") mit $N=3$ letzten (gleichen) Tagen, die über einen Monat ausgewertet wurden (vom 01.10.2016 bis 31.10.2016). Der resultierende $RMSE=189$. Wir verwenden die Abkürzung CLD zur Bezeichnung des "copy-last-days"-Persistenzmodells.

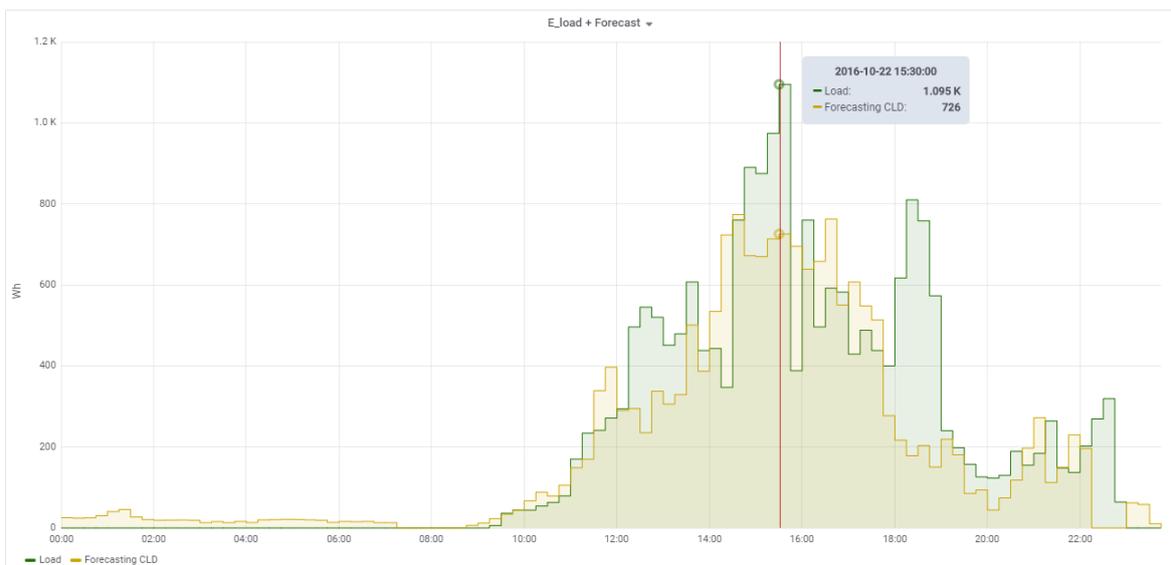


Abbildung 7 Persistenzmodell ("Copy-last-days") mit $N=3$, das über einen Tag (22.10.2016) ausgewertet wurde. Der resultierende $RMSE=153$.

4.3.2 Kombinationen von Persistenz ("copy-last-days") und Auto-Regressive (AR)-basierten Modellen

In diesem Abschnitt stellen wir die Vorhersageleistung des kombinierten auto-regressiven (AR) und Persistenzmodells ("copy-last-days") vor. Wie bereits beschrieben, versucht ein solches kombiniertes Modell die Persistenz im Lastprofil durch ähnliche Tage (Persistenzkomponente) und zeitliche Effekte innerhalb desselben Tages (AR-Komponente) zu erfassen.

Die folgende [Abbildung 8](#) zeigt die Reaktion dieses kombinierten Modells über den Zeitraum von einem Monat, während die [Abbildung 9](#) die entsprechende Reaktion über den Zeitraum von einem Tag darstellt.

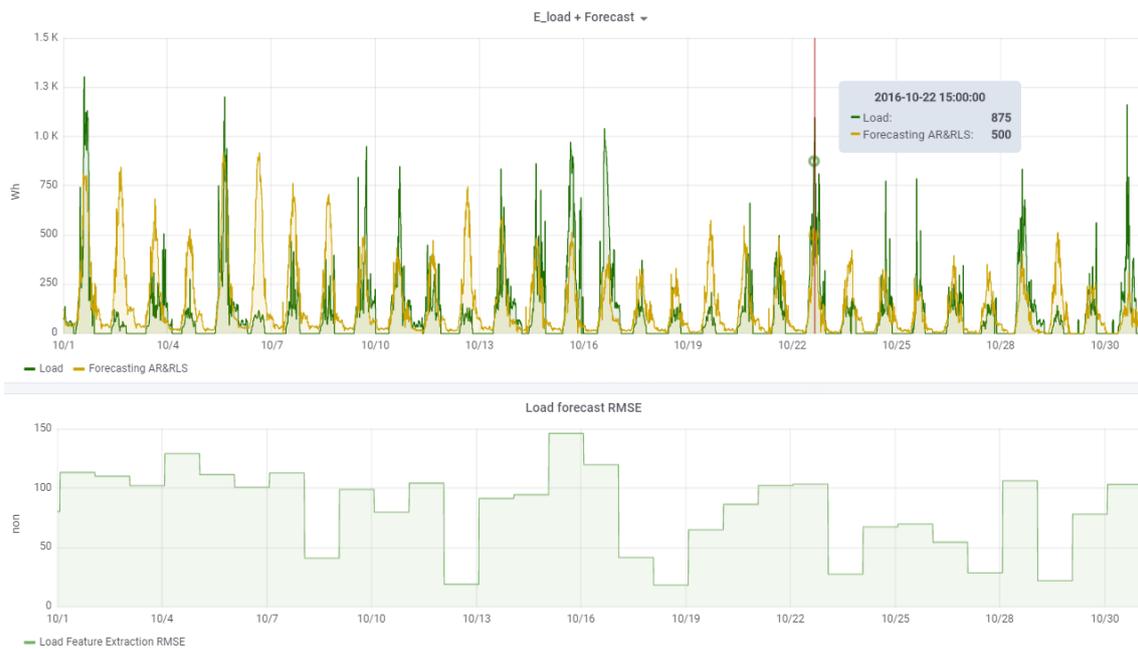


Abbildung 8 Kombiniertes AR mit Persistenzmodell ("Copy-last-days") mit N=3 letzten (gleichen) Tagen, die über einen Monat ausgewertet wurden (vom 01.10.2016 bis 31.10.2016). Der resultierende RMSE=159.

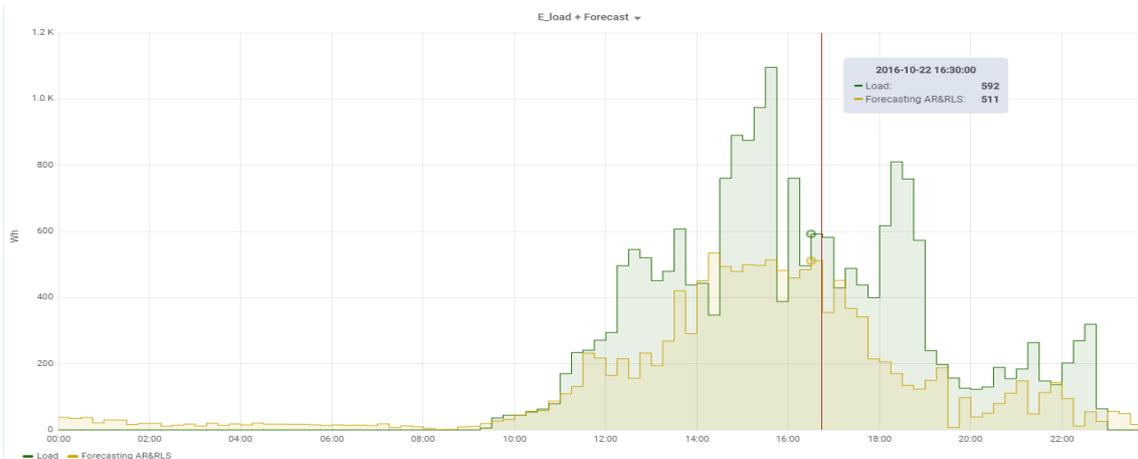


Abbildung 9 Kombiniertes AR mit Persistenzmodell ("Copy-last-days") mit N=3 letzten (gleichen) Tagen, die über einen Tag (22.10.2016) ausgewertet wurden. Der resultierende RMSE=174.

Es ist offensichtlich (durch direkten Vergleich mit dem Persistenzmodell), dass die Leistungsfähigkeit des kombinierten Modells über einen Monat jene des Persistenzmodells um etwa 15% übertroffen hat, was zeigt, dass zeitliche Abhängigkeiten innerhalb desselben Tages die Lastprognose signifikant beeinflussen.

4.3.3 Holt-Winters Modell

In diesem Abschnitt beschreiben wir die Leistung des in Abschnitt 4.2.6 kurz vorgestellten Holt-Winters-Prognosemodells. In der Auswertung haben wir die Periode entweder als 1 Woche oder als 1 Tag definiert. Da diese saisonale Periodizität vorgegeben und benutzerdefiniert ist, ist diese Methode sehr empfindlich gegenüber dem für die Auswertung verwendeten Datensatz (da die Periodizität im Verhalten der Bewohner zwischen den Haushalten stark variieren kann).

Insbesondere in ~~Abbildung 10~~ ~~Abbildung 10~~ und ~~Abbildung 11~~ ~~Abbildung 11~~ stellen wir die Leistung dieses Modells über den Zeitraum von einem Monat bzw. einem Tag dar, bei einer saisonalen Komponente von 1 Woche.

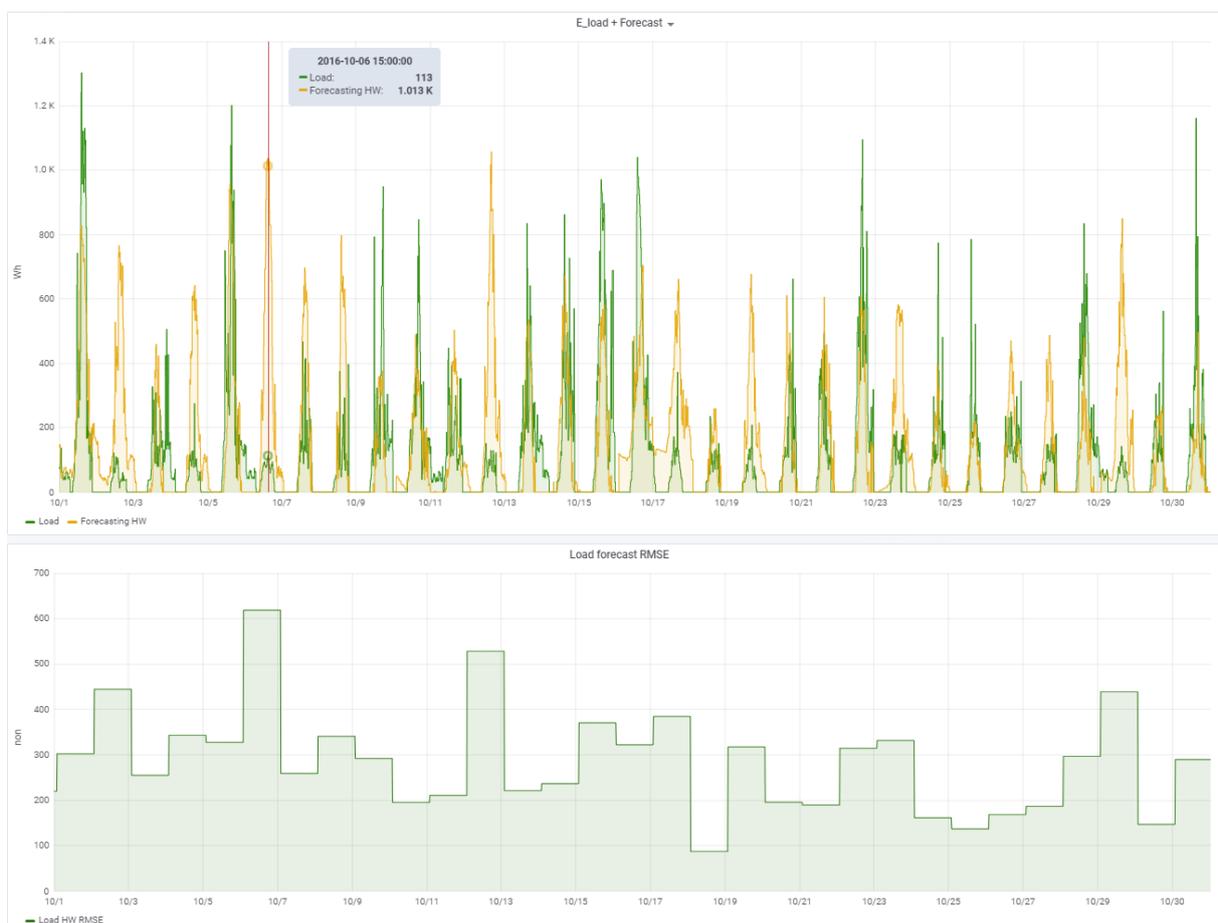


Abbildung 10 Holt-Winters Auswertung über einen Monat (vom 01.10.2016 bis 31.10.2016) für eine saisonale Komponente von 1 Woche. Der resultierende RMSE=288.

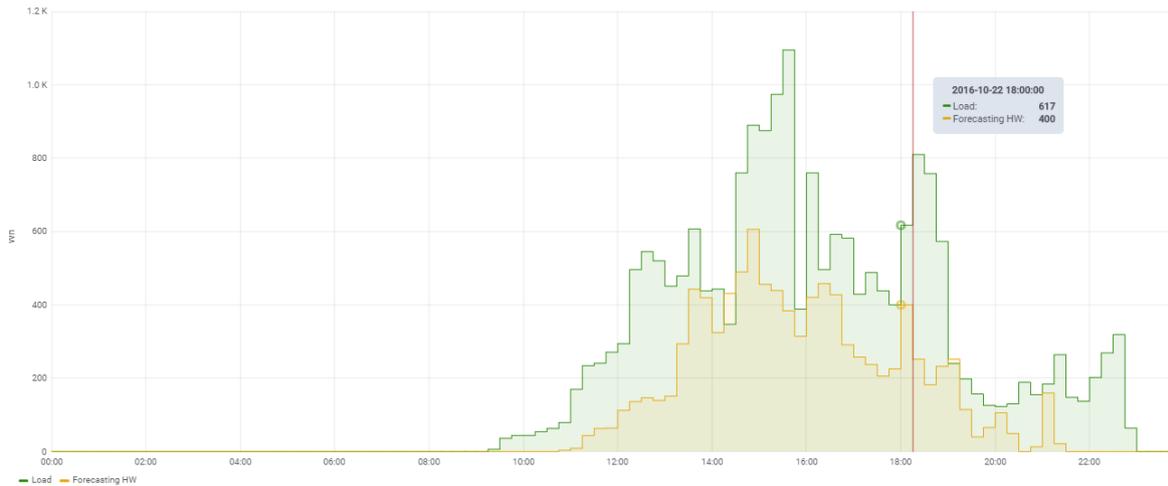


Abbildung 11 Holt-Winters Auswertung über einen Tag (22.10.2016) für eine saisonale Komponente von 1 Woche. Der resultierende RMSE=305.

Außerdem stellen wir in [Abbildung 12](#) und [Abbildung 13](#) die Leistungsfähigkeit dieses Modells über den Zeitraum von einem Monat bzw. einem Tag bei einer Periodizität von 1 Tag dar. Insgesamt konnten wir eine Verbesserung der Leistung feststellen, wenn die saisonale Komponente auf der Grundlage eines Tages definiert wurde, was jedoch sehr spezifisch für den für diese Auswertung verwendeten Datensatz sein könnte.

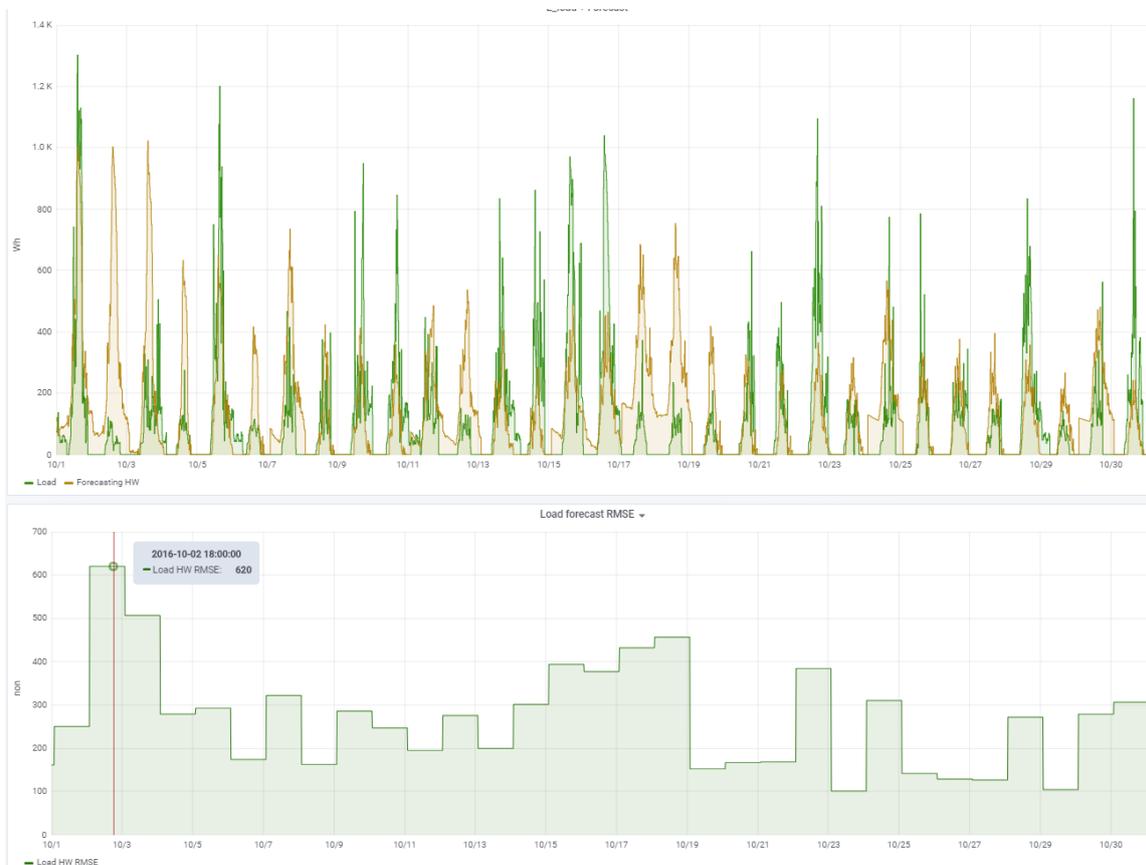


Abbildung 12 Holt-Winters Auswertung über einen Monat (vom 01.10.2016 bis 31.10.2016) und für eine saisonale Komponente von 1 Tag. Der resultierende RMSE=271.

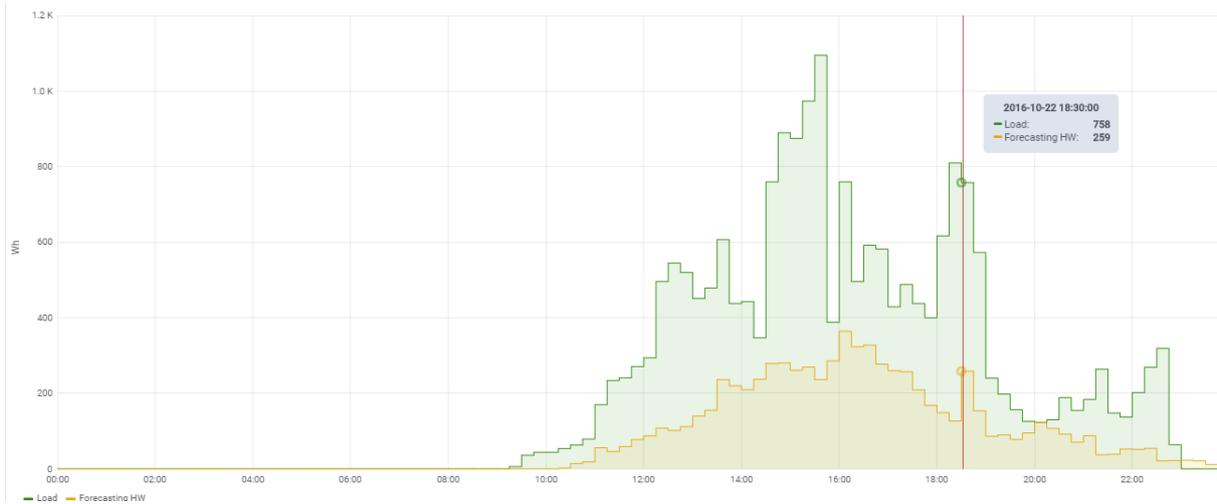


Abbildung 13 Holt-Winters Auswertung über einen Tag (22.10.2016) und für eine saisonale Komponente von 1 Tag. Der resultierende RMSE=366.

4.3.4 SARIMA Modell

Die Leistung des SARIMA-Modells über einen Zeitraum von einem Monat ist in [Abbildung 14](#) und für einen Tag in [Abbildung 15](#) zu sehen. Beachten Sie, dass die Performance des SARIMA-Modells besser ist als die entsprechende Performance der Holt-Winters-Modelle und fast identisch mit dem Persistenzmodell.

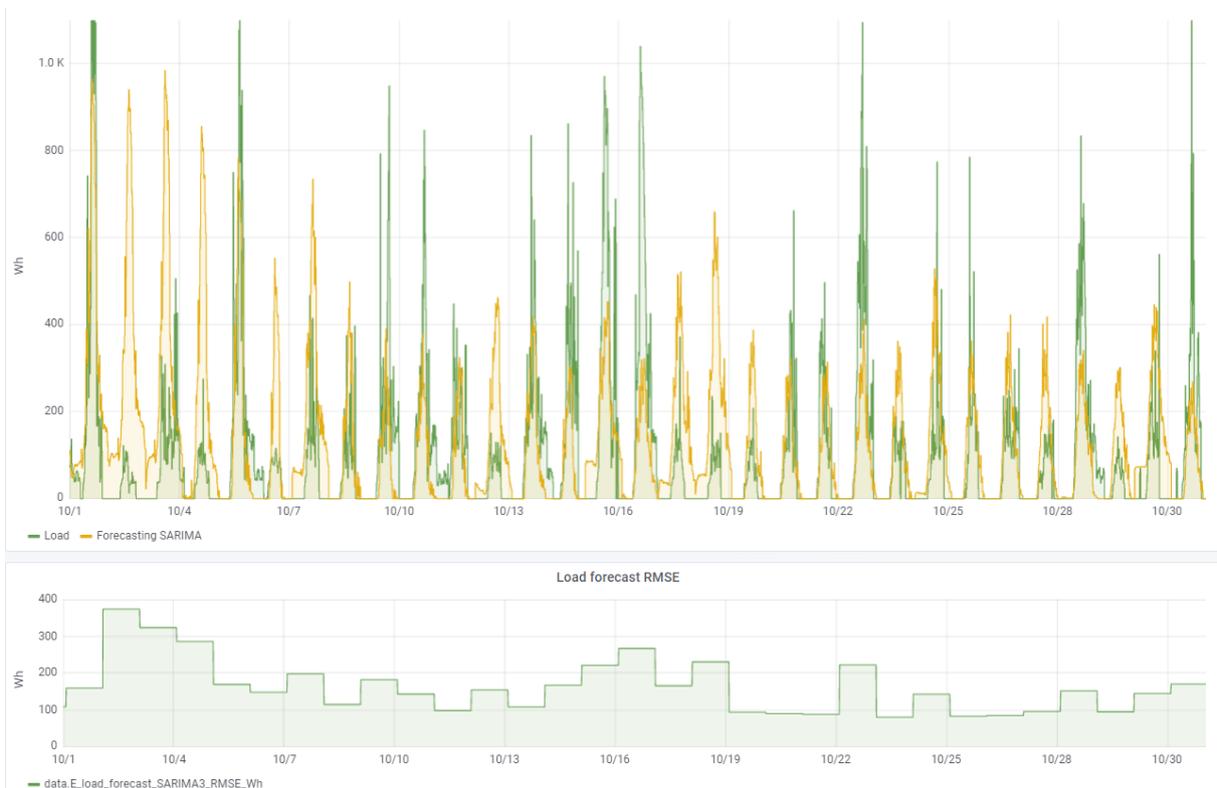


Abbildung 14 SARIMA-Modell vom Typ $(1,1,1)(1,1,1)_{96}$ mit einer Auswertung über einen Monat (vom 01.10.2016 bis 31.10.2016). Der resultierende RMSE=163.

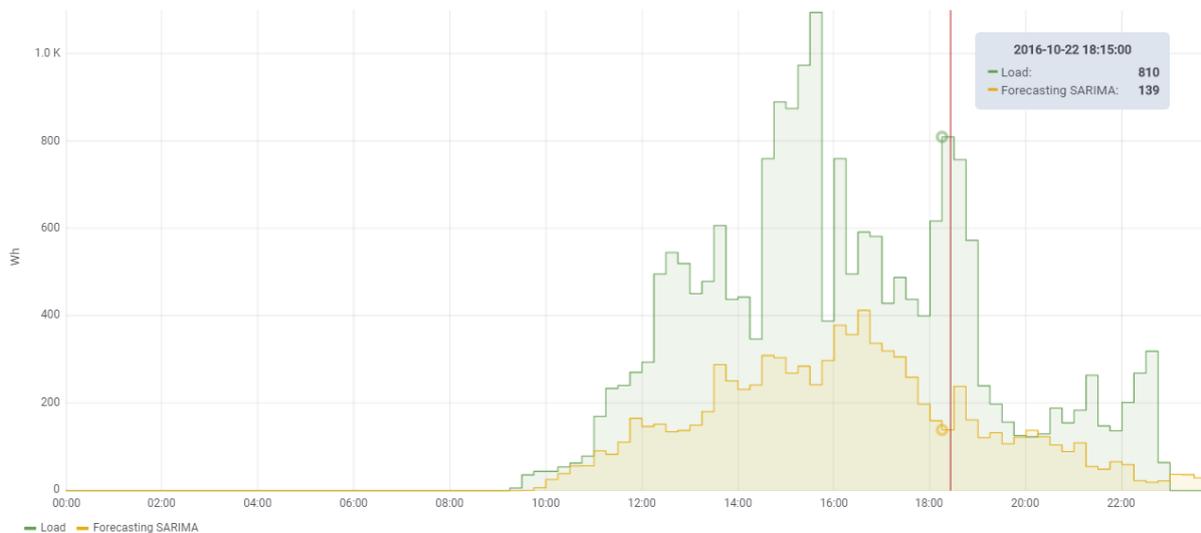


Abbildung 15 SARIMA-Modell vom Typ $(1,1,1)(1,1,1)_{96}$ mit einer Auswertung über einen Tag (22.10.2016). Der resultierende RMSE=211.

4.3.5 Feature-Extraktionsmodell

Die Leistung der Feature-Extraktionsmethode über einen Zeitraum von einem Monat und einem Tag ist in [Abbildung 16](#) und [Abbildung 17](#) dargestellt. Wir haben festgestellt, dass die Feature-Extraktionsmethode Ähnlichkeiten mit früheren Perioden (d.h. saisonale Effekte) besser erfassen kann als das kombinierte Auto-Regressive- und Persistenzmodell. Der daraus resultierende Prognosefehler (RMSE) wurde um 50% reduziert.

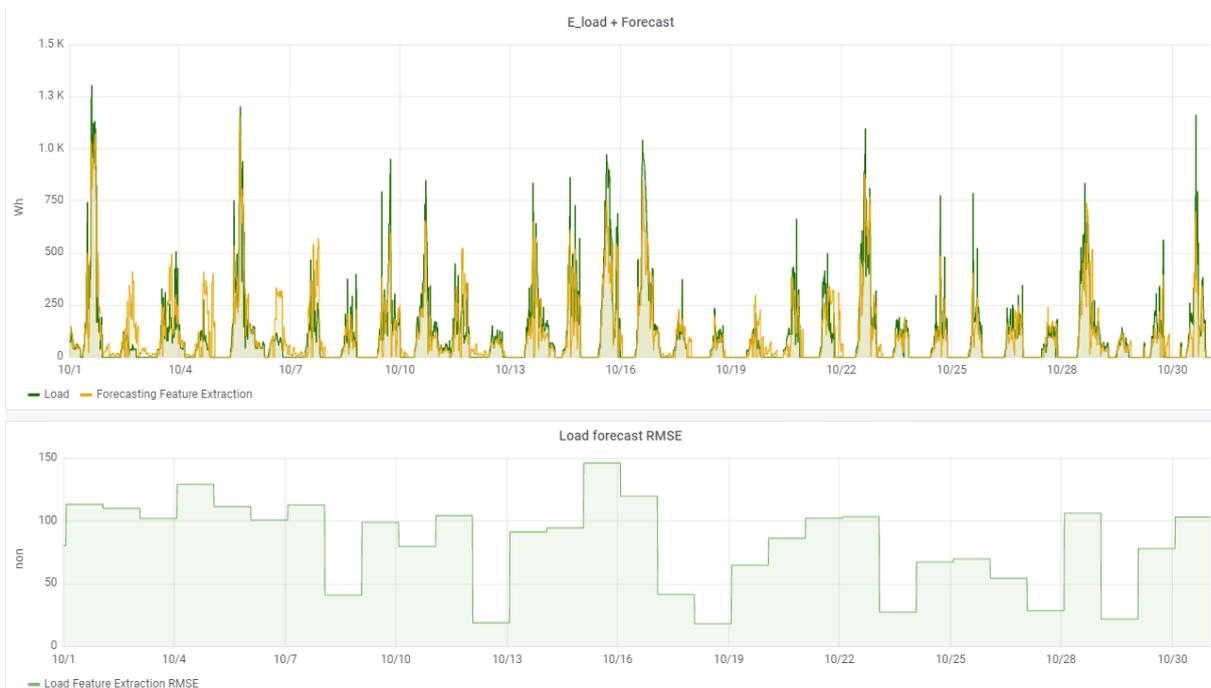


Abbildung 16 Feature-Extraktionsmethode über einen Monat (vom 01.10.2016 bis 31.10.2016). Der resultierende RMSE=82.

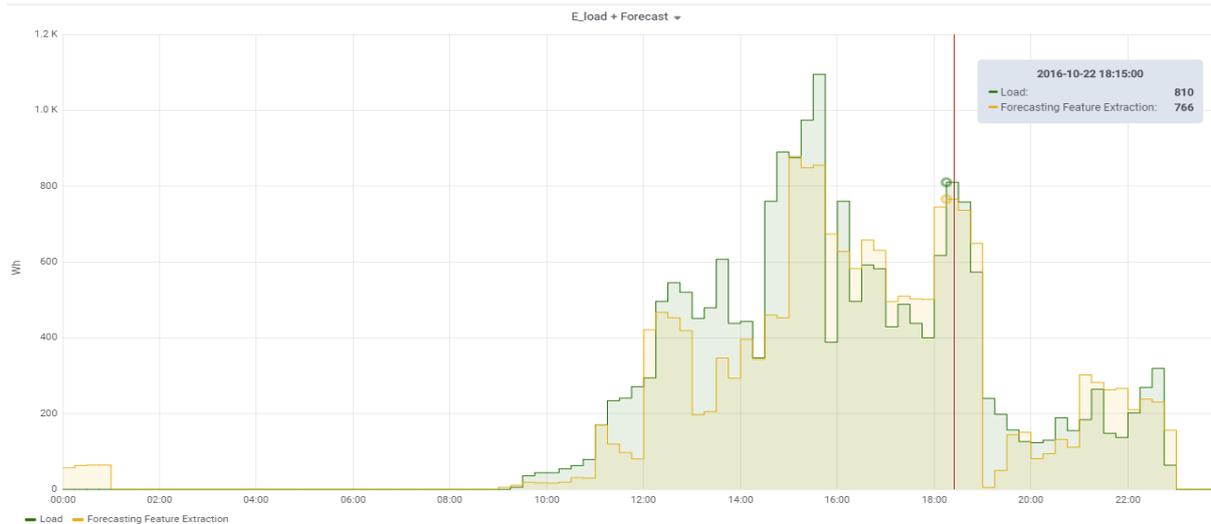


Abbildung 17 Feature-Extraktionsmethode mit einer Auswertung über einen Tag (22.10.2016). Der resultierende RMSE=103.

4.3.6 Gesamtleistungsvergleiche

In der folgenden ~~Tabelle 3~~ [Tabelle 3](#) haben wir die resultierenden RMSE-Ergebnisse aller Methoden über die Dauer von 1 Tag und 1 Monat zusammengefasst. Im Allgemeinen beobachten wir, dass die Feature-Extraktionsmethode, die mehrere Persistenz-Features (die für die Lastprognose sehr spezifisch sind) kombiniert, alle anderen Methoden übertrifft.

Tabelle 3 RMSE-Vergleich der untersuchten Lastprognosemethoden.

Duration	RMSE (1 day)	RMSE (1 month)
Persistenz	153	189
Holt-Winters	305	288
AR+Persistenz	174	159
SARIMA	211	163
Feature Extraktion	103	82

5 Warmwassernutzung

5.1 Datenbeschreibung und –aufbereitung

Die Datenbasis für die Erstellung der Warmwassernutzungsprognosen sind 20 Smart-Boiler, die bei Kunden der TIWAG verbaut sind und ca. seit März 2019 Daten liefern. Der integrierte Hardware-Controller in den elektrischen Warmwasserspeichern liefert mehrmals pro Minute Werte von Temperatursensoren, die an unterschiedlichen Stellen im Speicher angebracht sind. Daraus wird mit einer speziellen Berechnungsmethode der aktuelle Energiegehalt (=Kapazität) des Wassers im Boiler berechnet (in kWh). Weiters liefert der Controller die momentane elektrische Leistungsaufnahme. Leider ist kein Durchflusssensor verbaut, um die Menge von Warmwasserzapfungen zu ermitteln, weshalb eine alternative Approximation mit Hilfe der Änderung der Kapazität angestrebt wurde. Die verfügbaren Daten wurden exportiert und in folgender Aufbereitungs-Pipeline durchlaufen, um entsprechend bereinigte und vorberechnete Daten für die Prognosemodelle zu erhalten.

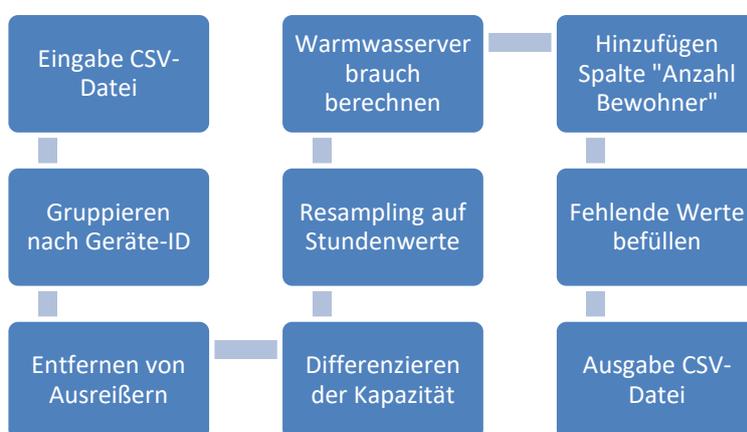


Abbildung 18 Die Schritte der Datenaufbereitung dienen zum Erstellen von passenden Eingabedaten für die Prognosemodelle.

Folgende Spalten sind in der Input-Datei enthalten: Die Geräte-ID, der UTC-Zeitstempel der Messung, elektr. Leistung in kW und Energieinhalt in kWh. Es wurde entschieden die Warmwasserzapfungen in einer Auflösung von einer Stunde zu berechnen, da höhere Auflösungen (z.B. 15 Minuten) schlechte Ergebnisse und viel längere Laufzeiten in ersten Experimenten zeigten. Als Beispiel dafür kann genannt werden, dass es wahrscheinlicher ist, dass eine morgendliche Dusche in derselben Stunde passiert als im selben 15 Minuten-Block. Der wichtigste Schritt in der Pipeline ist natürlich den Warmwasserverbrauch zu berechnen. Dazu wurde der Warmwasserspeicher als offenes thermodynamisches System modelliert, dem Energie über die elektrische Arbeit des Heizstabes und das einfließende Wasser zugeführt, und Energie über Zapfungen und Verlustleistung genommen wird.

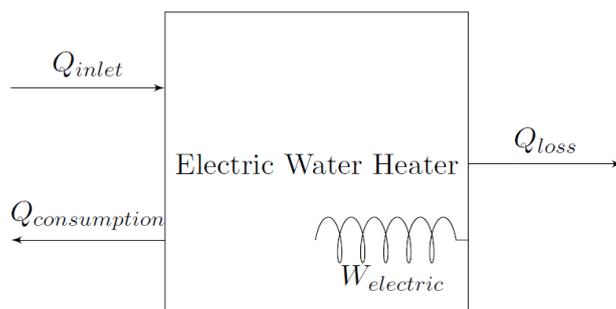


Abbildung 19 Der Boiler modelliert als offenes thermodynamisches System.

Der erste Hauptsatz der Thermodynamik besagt, dass keine Energie in einem thermodynamischen System verloren gehen kann. Wir erstellen eine vereinfachte Form der Energiebilanz mit folgender Gleichung:

$$\Delta Q_{heater} = W_{electric} + Q_{inlet} - Q_{consumption} - Q_{loss}$$

Die Änderung der internen Energiemenge ist bekannt, ebenso die elektrische Arbeit. Der Energie-Verlust, der trotz guter Isolierung entsteht, wurde empirisch ermittelt und mit dem Datenblatt des Boilers abgeglichen und beträgt ca. 0,05 kWh pro Stunde. Die Energie des zufließenden Wassers kann mit 0 angesetzt werden, da für die Berechnung der Energiemenge im Boiler die Leitungswassertemperatur der Bezugspunkt ist.

Daraus ergibt sich die folgende Gleichung zur Berechnung des approximierten Warmwasserverbrauchs:

$$\tilde{Q}_{consumption} = \max(0, W_{electric} - \Delta Q_{heater} - Q_{loss})$$

Um sicher zu gehen, dass keine negativen Werte für den Verbrauch entstehen im Falle eines Messfehlers bzw. durch die Annäherung, wird der Wertebereich auf nicht-negativ eingeschränkt. [Abbildung 20](#) zeigt die berechneten Warmwasserverbrauch (in Form von Energie) von zwei unterschiedlichen Haushalten.

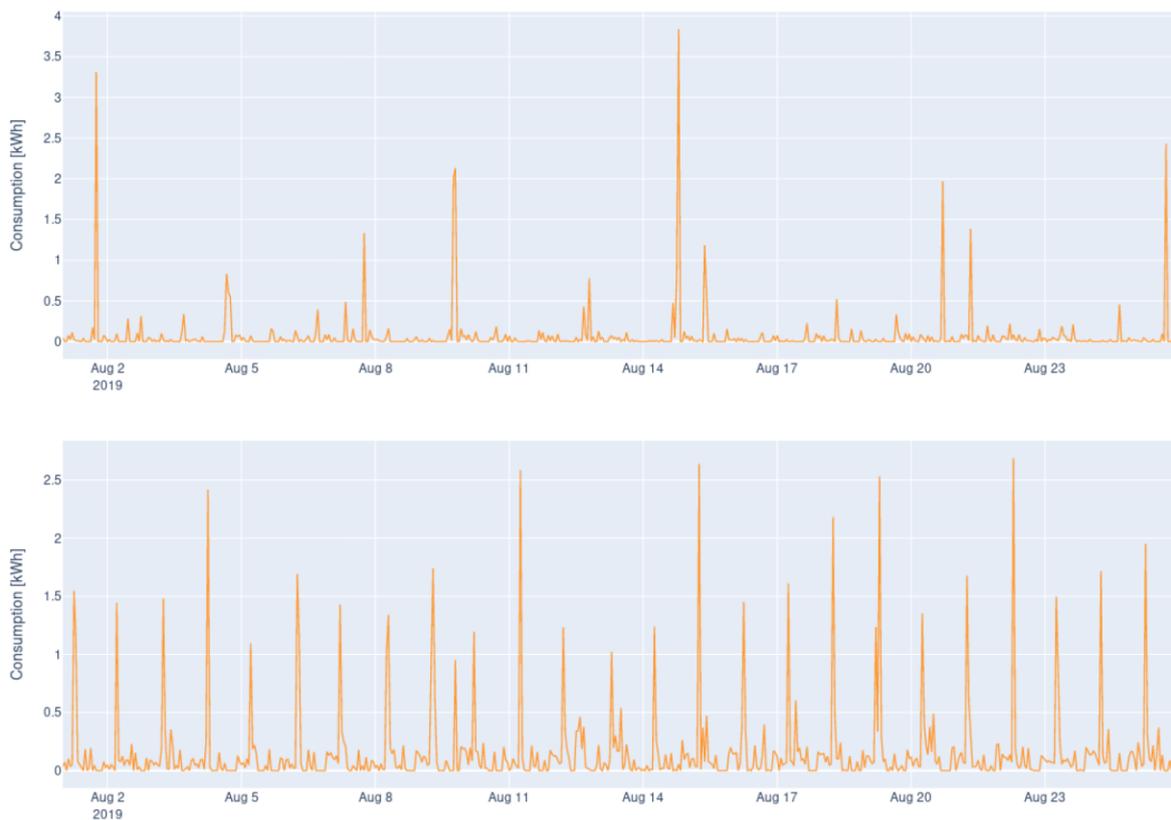


Abbildung 20 Beispiele von unregelmäßigem und regelmäßigem Warmwasserverbrauch in zwei verschiedenen Haushalten

5.2 Methodik

Es wurde angestrebt klassische statistische Prognosemethoden mit einem modernen künstlichen neuronalen Netz Ansatz zu vergleichen. Statistische Vorhersagemodelle werden schon sehr lange verwendet, erste ARMA Modelle wurden bereits in den späten 1930er Jahren entwickelt und Methoden der exponentiellen Glättung in den 1950ern (Majid & Mir, 2018). Diese beiden linearen Methoden sind laut Hyndman heute noch die am häufigsten

eingesetzten Modelle für die Prognose von Zeitreihen (Hyndman & Athanasopoulos, 2018). Das Problem von linearen Modellen ist die beschränkte Lernfähigkeit von nur linearen Zusammenhängen z.B. zwischen zeitlich versetzten Beobachtungen oder anderen Variablen. Reale Zeitreihen zeigen oft nichtlineares Verhalten, weshalb auch nichtlineare Modelle oft besser geeignet sind (Hagan & Behr, 1987). Künstliche neuronale Netze (ANNs) ermöglichen die Modellierung dieser nichtlinearen Zusammenhänge und wurden durch die steigende Rechnerleistung und Evolution von GPUs in den letzten 2 Jahrzehnten sehr populär.

Wir entschieden uns zwei statistische Modelle (SARIMA und Exponential Smoothing) mit einem speziellen Recurrent Neural Network, dem Long Short-Term Memory (LSTM), zu vergleichen. Der Prognose-Horizont, d.h. der Zeitraum für den eine Prognose erstellt wird, wurde mit 48 Stunden festgelegt.

5.2.1 Fehlermetriken

Speziell im Bereich von Prognosen wird die Qualität von Modellen meist in Form von Fehlermetriken gemessen, die den quantitativen Unterschied zwischen echten (y_t) und vorhergesagten (\hat{y}_t) Werten ausdrückt. Die Festlegung einer geeigneten Fehlermetrik ist eine sehr wichtige, aber keineswegs leichte Entscheidung. Eine gute Metrik sollte unverfälscht, symmetrisch (außer Asymmetrie ist erwünscht), universell anwendbar, Skalen-unabhängig (für Vergleiche) und einfach zu interpretieren sein. Unzählige Fehlermetriken existieren, welche grob in folgende fünf Kategorien eingeteilt werden können (Hyndman & Koehler, 2006):

- **Maßstabsabhängige Metriken:** Die Skala der Metrik hängt von der Skala der Daten ab, d.h. Daten mit unterschiedlicher Skalierung können nicht miteinander verglichen werden. Beispiele: Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE).
- **Prozentuale Metriken:** Der Fehler wird als prozentualer Wert angegeben, d.h. diese Kategorie der Metriken ist skalunabhängig. Die bekannteste Metrik aus dieser Klasse ist MAPE (Mean Absolute Percentage Error). MAPE hat zwei Nachteile: Erstens dürfen die Werte der Zeitreihe keine 0-Werte enthalten und zweitens werden negative Fehler ($y_t < \hat{y}_t$) stärker bewertet als positive (=Asymmetrie).
- **Metriken basierend auf relativen Fehlern:** Um Maßstabsunabhängigkeit zu erreichen werden Fehler von unterschiedlichen Methoden in Relation gesetzt. Meist wird für das zu vergleichende Modell ein einfaches (naives) Benchmark-Modell verwendet. Beispiel: Mean Relative Absolute Error (MRAE).
- **Relative Metriken:** Im Unterschied zur oberen Klasse werden hier nicht die einzelnen Fehler im Verhältnis betrachtet, sondern zwei fertig berechnete Metriken. Vorzugsweise ist der Nenner im Verhältnis dieselbe Metrik aber aus Vorhersagen mit dem Benchmark-Modell (*) berechnet. Bsp: Relative Mean Absolute Error (RelMAE = MAE/MAE*).
- **Skalierte Metriken:** In (Hyndman & Koehler, 2006) werden die Nachteile von relativen und auf relativen Fehlern basierenden Metriken erörtert. Deshalb wird die Metrik „Mean Absolute Scaled Error“ (MASE) vorgeschlagen, um alle Nachteile der anderen Kategorien zu umgehen. MASE relativiert den MAE der Prognose mit dem „in-sample“ MAE der Prognose der naiven Methode (Saisonal-Naiv/Random Walk). Sollte die Zeitreihe eine Saisonalität aufweisen, wird der Parameter m entsprechend der Periode gesetzt (Saisonal-Naiv), andernfalls wird $m = 1$ gesetzt, was einem Random-Walk-Modell entspricht.

$$MASE = \frac{MAE}{MAE_{in-sample}^*} = \frac{\frac{1}{N} \sum_{t=1}^N |e_t|}{\frac{1}{T-m} \sum_{t=m+1}^T |y_t - y_{t-m}|}$$

Obwohl die Fehlermetriken MSE/RMSE, MAE und MAPE am gebräuchlichsten sind, entscheiden wir uns für MASE als Metrik aufgrund seiner nützlichen Eigenschaften. MAPE konnte von Anfang an ausgeschlossen werden, da unsere Zeitreihen viele 0-Werte enthalten. Da wir unterschiedliche Modelle und auch Zapfprofile von 20 verschiedenen Haushalten mit unterschiedlichstem Zapfverhalten vergleichen, bietet sich MASE sehr gut an. Ein weiterer Vorteil ist, dass man auf einem Blick erkennt ob das Vorhersagemodell besser als das Benchmark-Modell ist, nämlich wenn $MASE < 1$ ist.

5.2.2 Zeitreihenanalyse

Zeitreihen können verschiedene Muster bzw. Komponenten aufweisen:

- **Trend:** Ist der Durchschnitt der Werte einer Zeitreihe nicht konstant über die Zeit, wird von einem Trend gesprochen
- **Saisonalität:** Ein saisonales Muster besteht aus Fluktuationen in der Zeitreihe, die durch bestimmte zeitliche Faktoren (z.B. spezifische Monate eines Jahres oder die Tageszeit) ausgelöst werden. Saisonalitäten haben immer eine fixe Frequenz.
- **Zyklen:** Ein Zyklus beschreibt wiederholte Fluktuationen mit unterschiedlicher Frequenz. Zyklen treten häufig in ökonomischen Zeitreihen auf.

Statistische Vorhersagemethoden setzen voraus, dass die Zeitreihen „stationär“ sind, d.h. dass sich die statistischen Eigenschaften wie Mittelwert und Varianz nicht mit der Zeit ändern. Sobald eine Zeitreihe einen Trend oder eine Saisonalität aufweist, muss entweder das Modell für Stationarität sorgen oder der/die Prognostiker/in zuerst entsprechende Transformationen auf die Daten anwenden. Eine einfache Möglichkeit Trends und Saisonalitäten zu entfernen ist das zeitliche Differenzieren von Zeitreihen. Durch den Differenzierungsschritt in der Datenaufbereitung sollten unsere Zeitreihen (Zapfprofile) bereits stationär sein. Sogenannte Unit-Root-Tests, wie z.B. ADF oder KPSS, ermöglichen eine automatisierte Feststellung, ob und wie oft eine Zeitreihe differenziert werden muss, um Stationarität zu erhalten.

Eine weitere gute Analysemöglichkeit stellt die sogenannte (partial) Autocorrelation-Function (ACF) dar. Mit ACF lässt sich der lineare Zusammenhang zwischen Beobachtungen leicht darstellen. Die ACF-Diagramme (auch Korrelogramme) der Zapfprofile zeigen eine starke Korrelation zu Beobachtungen mit zeitlichem Versatz von 1 bzw. Vielfachen von 24 (Saisonalität mit Periode $m=24$).

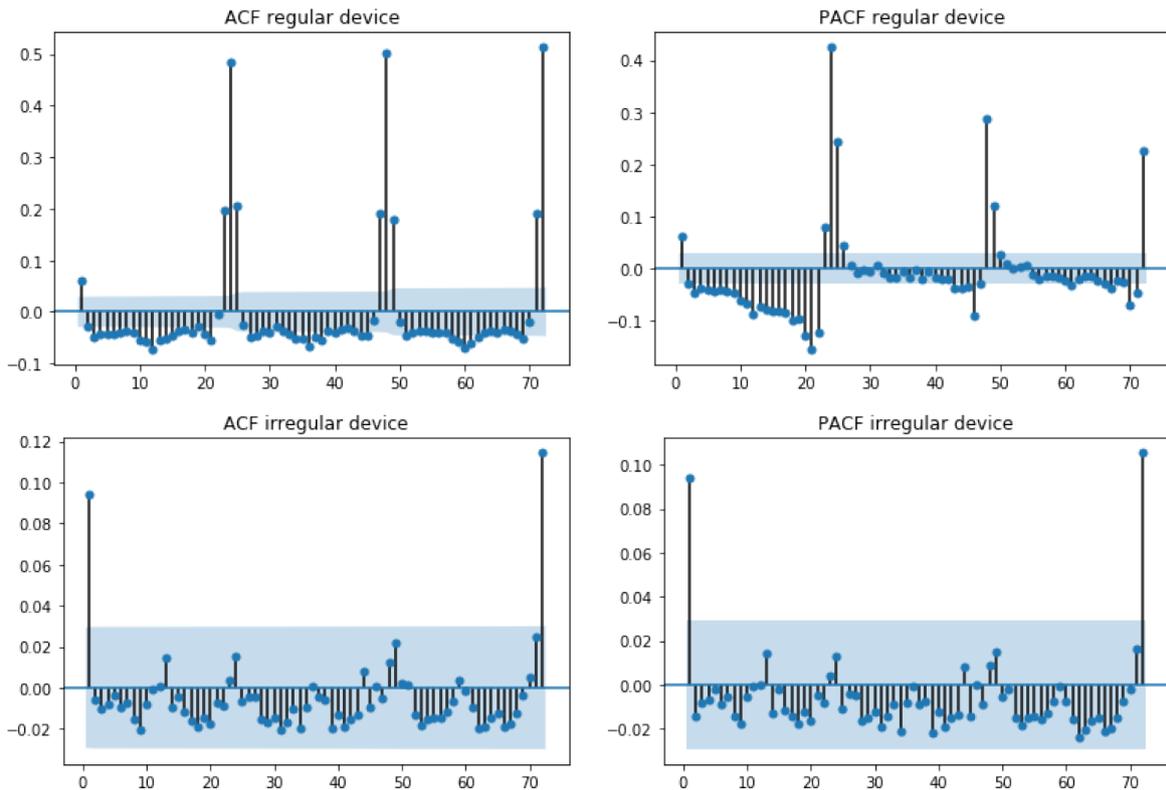


Abbildung 21 ACF und PACF Diagramme für ein regelmäßiges (obere Reihe) und unregelmäßiges (untere Reihe) Zapfprofil

5.2.3 Exponential Smoothing

Die Methode der exponentiellen Glättung funktioniert mittels gewichteten Mittelwertes von vorherigen Beobachtungen. Die Gewichtungen verkleinern sich exponentiell je weiter die Beobachtungen in die Vergangenheit reichen, wobei $0 \leq \alpha \leq 1$ der Glättungs-Parameter ist. Das sogenannte „Simple Exponential Smoothing“ (SES) lässt sich mit folgenden Gleichungen darstellen:

$$\hat{y}_{t+h} = l_t$$

$$l_t = \alpha y_t + (1 - \alpha)l_{t-1}$$

Sind die Parameter l_0 und α bekannt, lässt sich die rekursive Formel lösen und eine Vorhersage kann berechnet werden. Um diese beiden Parameter zu berechnen muss ein nicht-lineares Optimierungsproblem durch Minimierung des Sum-of-Squares Fehlers gelöst werden (Hyndman & Athanasopoulos, 2018). Holt und Winters haben die SES Methode um eine Trend- und Saisonalitätskomponente erweitert zu „Holt-Winters' additive seasonal method“ (Winters, 1960):

$$\hat{y}_{t+h} = l_t + hb_t + s_{t+h-m(k+1)} \text{ mit } k = \left\lfloor \frac{h-1}{m} \right\rfloor$$

$$l_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} - b_{t-1})$$

$$b_t = \beta^*(l_t - l_{t-1}) + (1 - \beta^*)b_{t-1}$$

$$s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m}$$

Diese Methode besitzt zwei weitere Glättungs-Parameter β^* und γ für den Trend bzw. Saisonalität.

5.2.4 SARIMA

SARIMA ist ein Akronym für „Seasonal Auto-Regressive Integrated Moving Average“.

SARIMA Modelle werden immer durch eine Ordnung $(p, d, q)(P, D, Q)_m$ bestimmt, wobei p , q , P und Q den maximal zu berücksichtigenden zeitlichen Versatz angeben und d und D den Grad der Differenzierung und die Großbuchstaben immer für den saisonalen Teil verantwortlich sind.

Um die Notation zu vereinfachen, empfiehlt es sich den sogenannten „Backshift-Operator“ einzuführen.

$$B^m y_t = y_{t-m}$$

Die vollständige Beschreibung eines SARIMA-Modelles kann durch folgende Gleichung erfolgen:

$$\underbrace{(1 - \phi_1 B - \dots - \phi_p B^p)}_{AR(p)} \underbrace{(1 - \Phi_1 B^m - \dots - \Phi_P B^{mP})}_{\text{seasonal } AR(P)} \underbrace{(1 - B)^d}_{\text{diff}} \underbrace{(1 - B^m)^D}_{\text{seasonal diff.}} y_t = c + \underbrace{(1 + \theta_1 B + \dots + \theta_q B^q)}_{MA(q)} \underbrace{(1 + \Theta_1 B + \dots + \Theta_Q B^{mQ})}_{\text{seasonal } MA(Q)} \varepsilon_t$$

Der AR-Teil in der Gleichung kann als lineare Regression von versetzten Beobachtungen und der MA-Teil von versetzten Vorhersagefehlern (weißes Rauschen = ε_t) gesehen werden. Die Parameter $c, \phi, \Phi, \theta, \Theta$ werden durch Maximum-Likelihood-Estimation bestimmt.

5.2.5 LSTM

Klassische rekurrente neuronale Netzwerke (RNNs) eignen sich sehr gut, um sequenzielle Daten wie Sätze in der Sprachverarbeitung oder Zeitreihen zu lernen. Allerdings haben RNNs das Problem der verschwindenden oder explodierenden Gradienten bei Langzeit-Abhängigkeiten. Das Long Short-Term Memory wurde von (Hochreiter & Schmidhuber, 1997) entwickelt, um dieses Problem zu beheben.

Den Kern einer sogenannten LSTM-Zelle bildet der Zustand c_t , der als „Gedächtnis“ fungiert. Dieser Zustand wird rückgeführt und durch sogenannte „Gates“ modifiziert. Ein Gate besteht aus einem einzelnen neuronalen Netz-Layer mit einer Sigmoid-Aktivierungsfunktion, gefolgt von einer elementweisen Multiplikations-Operation \otimes . Jedes Gate wirkt als Filter, da jedes Element des Inputs mit einem Faktor zwischen 0 und 1 multipliziert wird. Ein LSTM hat 3 Gates: Input, Output- und Forget-Gate. Das Input-Gate steuert wieviel von der aktuellen Information in das „Gedächtnis“ wandert, das Forget-Gate welche Information gelöscht werden soll, und das Output Gate welche Information vom internen Zustand an den Ausgang gelangen soll.

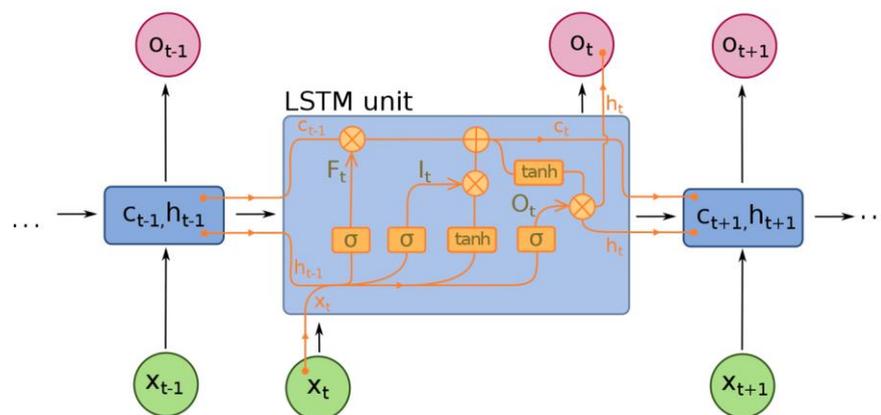


Abbildung 22 Die interne Struktur eines LSTMs. Quelle: (Deloche)

Für eine detaillierte Beschreibung und Erklärung der Funktionsweise und des Lernverfahrens verweisen wir auf den Original-Artikel (Hochreiter & Schmidhuber, 1997) und (Goodfellow, Bengio, & Courville, 2016).

5.3 Evaluierung

5.3.1 Evaluierungsstrategie

Die typische Vorgehensweise um die Leistung von Vorhersagemodellen zu ermitteln, ist das aufteilen des Datenbestandes in ein Trainings-Set und ein Test-Set. Die Modell-Parameter werden aus den Trainings-Daten ermittelt, während die Fehlermetrik anhand des auf das Test-Set angewandte Modell berechnet wird, um die Leistungsfähigkeit (=Generalisierungsfähigkeit) des Modells bei noch neuen, ungesehenen Daten zu testen. Der Anteil des Trainings-Sets von den Gesamtdaten wurde auf 80% festgelegt, die restlichen 20% für das Test-Set.

Klassische Kreuzvalidierungsverfahren mit zufälligen Aufteilungen können bei Zeitreihen nicht angewendet werden, da die chronologische Ordnung der Beobachtungen beibehalten werden muss.

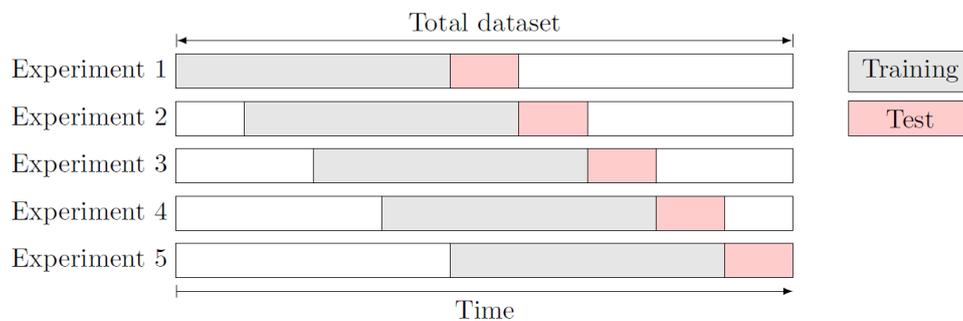


Abbildung 23 Zeitreihen-Kreuzvalidierung mit konstanter Trainings-Set-Größe und rollendem Zeitfenster

Bei den statistischen Modellen wurde die Evaluierungs-Variante aus [Abbildung 23 Zeitreihen-Kreuzvalidierung mit konstanter Trainings-Set-Größe und rollendem Zeitfenster](#) mit nicht überlappenden Training-Sets einer Länge von 8 Wochen angewandt. Das Test-Set enthält 48 Werte, was einem Zeitraum von 2 Tagen und somit dem Vorhersagezeitraum entspricht. Diese Strategie wurde für alle 20 Geräte durchgeführt und die individuellen Ergebnisse (MASE-Werte) gemittelt, um ein Gesamtergebnis zu erhalten. Für das SARIMA-Modell wurde weiters eine Rastersuche über die Ordnungs-Parameter p , P , q und Q mit möglichen Werten 0, 1 und 2 durchgeführt. Die Ordnung der Differenzierung wurde nach Ausführung von Unit-Root-Tests auf 0 gesetzt. Dadurch konnte eine beste Konfiguration $(2,0,0)(0,0,1)_{24}$ gefunden werden.

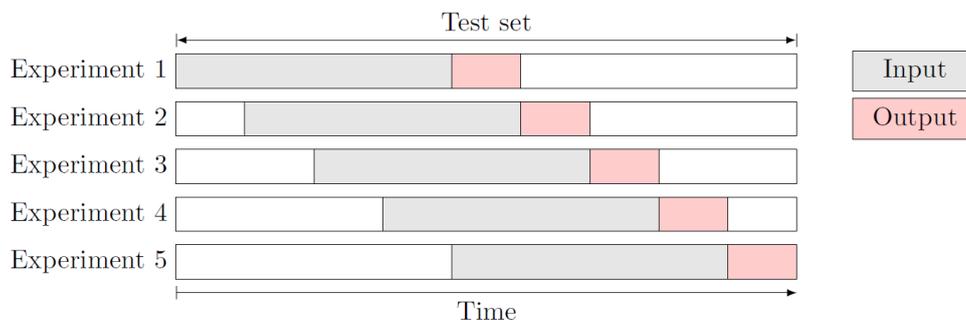


Abbildung 24 Die Zeitreihen-Kreuzvalidierung für überwachttes maschinelles Lernen wird nur auf das Test-Set angewandt

Die Evaluierung von Modellen aus der Kategorie „überwachtes maschinelles Lernen“, wie z.B. künstliche neuronale Netze, unterscheidet sich zur ersten Validierungs-Variante, da die Modelle mit sehr vielen Input/Output-Paaren aus dem Trainings-Set „gefüttert“ werden, um ein Lernergebnis zu erzielen. Nach dem Lernen werden Daten-Paare aus dem Test-Set für die Evaluierung genommen (siehe Abbildung 24).

Im Falle des LSTM-Modells untersuchten wir zwei verschiedene Ansätze:

1. Für jeden Boiler wurde ein eigenes LSTM-Netz trainiert und evaluiert und die einzelnen Ergebnisse gemittelt → „Single-Device Model“ (SD)
2. Nur ein LSTM-Modell wurde mit allen von 20 Geräten verfügbaren Daten trainiert → „Multi-Device Model“ (MD)

Weiters wurde für den Input des RNNs ein Zeitraum von 2 Wochen festgelegt, der Zeitraum des Outputs entspricht wieder dem Prognose-Horizont von 48 Stunden.

In Erwartung auf möglicherweise bessere Ergebnisse wurden dem Modell zusätzliche „Features“ zum Warmwasserverbrauch zugeführt:

- Anzahl der Haushaltsmitglieder: Natürlicherweise sollte die Anzahl der Bewohner eines Haushalts eine große Auswirkung auf das Zapfprofil haben. Leider besteht keine Garantie, dass die verfügbaren Daten über die Bewohneranzahl aktuell und richtig sind.
- Wochentag: Speziell an Wochenenden darf ein anderes Verhalten erwartet werden.
- Durchschnitt des Verbrauchs der letzten 3 Stunden (MA3): Aufgrund der ACF Diagramme wurde ermittelt, dass der aktuelle Wert stark von den letzten 1-2 Stunden abhängt. Evtl. sollte das Modell explizit darüber informiert werden.
- Geräte-ID (nur Multi-Device Modell): Im Falle des einzelnen Modells, das mit den Profilen von allen 20 Geräte trainiert wird, könnte es Sinn machen dem Modell mitzuteilen, von welchem Gerät die Trainingsdaten stammen.

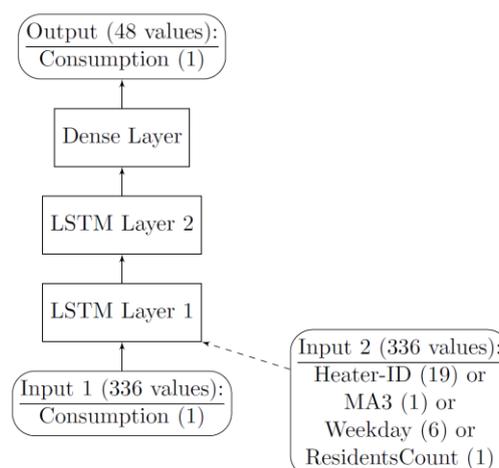


Abbildung 25 Die Architektur des evaluierten LSTM-Modells

Um Overfitting zu vermeiden, wurden die Regularisierungs-Techniken „Dropout“ und „Early-Stopping“ verwendet. Als Optimierer wurde „Adam“ verwendet, da dieser als schnell und zuverlässig gilt (Goodfellow, Bengio, & Courville, 2016).

Es wurde eine Rastersuche über die Hyper-Parameter Batch-Size (BS), Learning-Rate (LR), Anzahl der Neuronen (U) und Dropout-Rates (DO) vorgenommen. Wir evaluierten die Batch-Größen 64 und 259, Learning-Rates 0.01 und 0.001, 32 und 62 Neuronen und Dropout-Raten von 0.0 und 0.2.

5.3.2 Ergebnisse

[Tabelle 4 Ergebnisse der Evaluierung in Form der Fehlermetriken Mean Absolute Scaled Error \(MASE\) und Mean Absolute Error \(MAE\).](#) ~~Tabelle 4~~ die Ergebnisse der Evaluierung mit SARIMA als Modell mit dem kleinsten Fehler (MASE = 0.623). Die Fehlermetrik MAE ist ohne Benchmark-Modell wenig aussagekräftig und vergleichbar, wurde aber trotzdem inkludiert, um sich eine Vorstellung über die absolute Größenordnung machen zu können. Alle MASE-Werte sind kleiner als 1, was bedeutet, dass zumindest alle Modelle besser als das Benchmark-Methode „Saisonal-Naiv“ sind.

Tabelle 4 Ergebnisse der Evaluierung in Form der Fehlermetriken Mean Absolute Scaled Error (MASE) und Mean Absolute Error (MAE).

Modell	Beste Konfiguration	MAE	MASE
Exponential Smoothing	Kein Trend	0.124	0.826
SARIMA	(2,0,0)(0,0,1) ₂₄	0.091	0.623
SD LSTM	BS 64, LR 0.01, U 64, DO 0.2	0.129	0.994
SD LSTM + MA3	BS 64, LR 0.01, U 64, DO 0.0	0.120	0.934
SD LSTM + Wochentag	BS 64, LR 0.01, U 64, DO 0.2	0.099	0.795
SD LSTM + #Bewohner	BS 64, LR 0.01, U 64, DO 0.2	0.096	0.782
MD LSTM	BS 64, LR 0.01, U 32, DO 0.2	0.082	0.969
MD LSTM + Geräte-ID	BS 64, LR 0.01, U 32, DO 0.0	0.080	0.868
MD LSTM + MA3	BS 64, LR 0.01, U 32, DO 0.2	0.082	0.939
MD LSTM + Wochentag	BS 64, LR 0.01, U 64, DO 0.0	0.073	0.822
MD LSTM + #Bewohner	BS 64, LR 0.01, U 32, DO 0.2	0.078	0.891

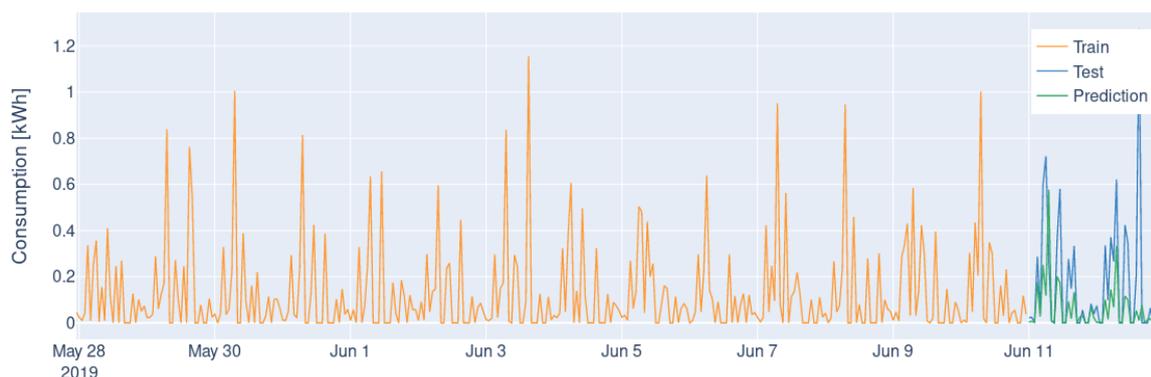


Abbildung 26 Beispiel-Vorhersage des SARIMA-Modells

Das Multi-Device LSTM Modell war bei einigen Konfigurationen nicht imstande zu lernen, was zur Folge hatte, dass die Vorhersage eine konstante 0 war. Das resultierte in einem niedrigem MAE und somit auch zu einem kleinem MASE, da ein großer Teil der Zeitreihen ebenfalls 0-Werte sind.

Das Single-Device LSTM-Modell zeigte bessere Ergebnisse bei Inkludierung der Features „Wochentag“ bzw. „Anzahl Bewohner“, allerdings entstehen die niedrigeren Fehlermetriken teilweise wieder durch Vorhersagen von langen 0-Werte-Sequenzen bei unregelmäßigen Profilen, die schwer zu prognostizieren sind. Dabei ist nicht festzustellen, ob das LSTM gelernt hat die 0-Sequenzen zu erzeugen, um den Fehler möglichst gering zu halten oder „Nichts“ bzw. das „Falsche“ gelernt hat. Eine Beispiel-Prognose für dieses Verhalten in [Abbildung 27 Die Prognose des SD-LSTMs liefert konstant "0" bei einem unregelmäßigem Profil.](#) [Abbildung 27](#) sehen.

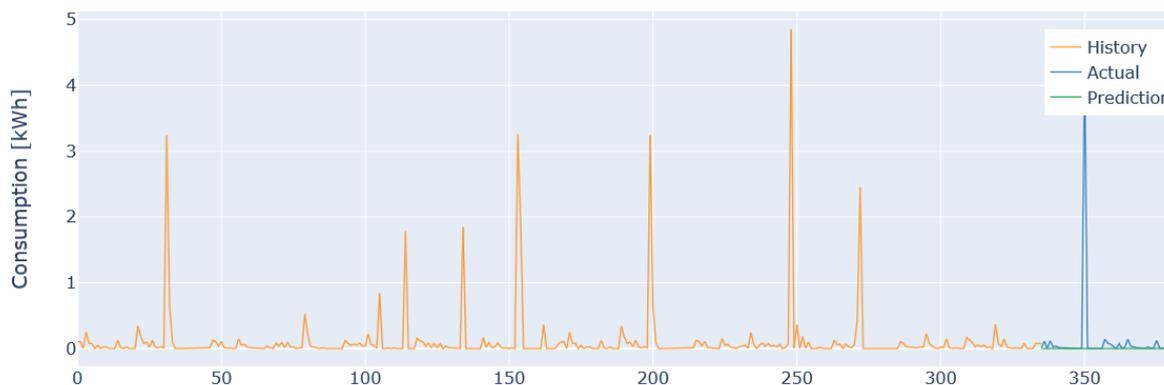


Abbildung 27 Die Prognose des SD-LSTMs liefert konstant "0" bei einem unregelmäßigem Profil.

Generell kann gesagt werden, dass die Prognose von Warmwasserzapfungen ein schwierig zu lösendes Problem ist, da die Profile zu einem gewissen Grad einem stochastischem Prozess entsprechen.

Die Einfachheit der statistischen Modelle ist ein großer Vorteil gegenüber den aufwändig zu trainierenden neuronalen Netzen. Weiters ist die Suche nach geeigneten Hyper-Parametern beim maschinellen Lernen ein zeitaufwändiger Prozess.

Die rasterbasierte Suche von Hyper-Parametern könnte in weiteren Experimenten durch Techniken wie Bayes'sche oder gradientenbasierte Optimierung ausgetauscht werden, um eventuell bessere Ergebnisse mit LSTM-Modellen zu erhalten.

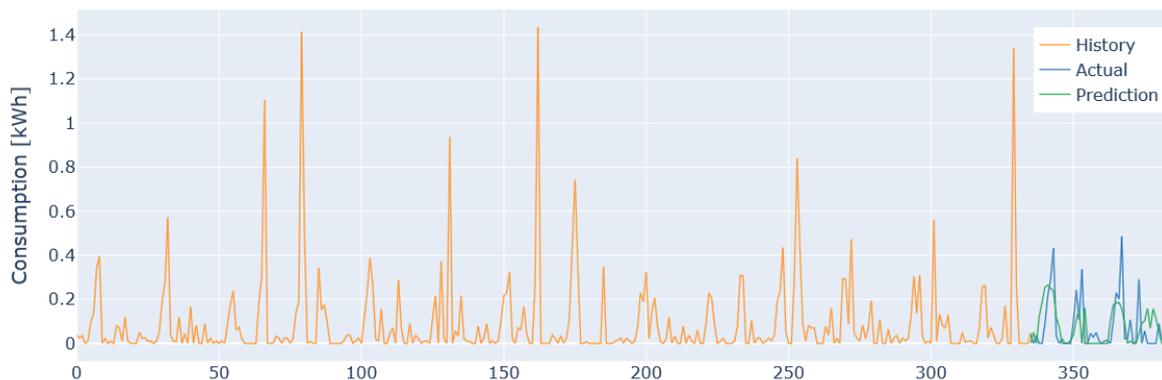


Abbildung 28 Gute Vorhersage des SD LSTMs mit zusätzlichem Feature "Wochentag"

5.4 Erweiterung der Methodik: LSTM und FFNN

Parallel zu den zuvor beschriebenen Methoden wurde unterschiedliche Ausprägungen von LSTM und FFNN Modellen entwickelt und mit unterschiedlichen Konfigurationen getestet. Das Ziel hinter dieser Entwicklung war es, ein Modell zu entwickeln, das sowohl für den Warmwasserverbrauch als auch für den Verbrauch von anderen Energieformen (z.B. elektrische Energie) ohne weitere Anpassungen genutzt werden kann.

5.4.1 Prognose – 24-Stunden Warmwasserverbrauch

Für die Prognose des Warmwasserverbrauchs in einem zeitlichen Horizont von 24 Stunden wurden unterschiedliche Modelle auf Basis von künstlichen neuronalen Netzwerkarchitekturen mit verschiedenen zeitlichen Auflösungen (15 Minuten, 1 Stunde, 4 Stunden) implementiert:

- Long Short Term Memory (LSTM) networks. Dieses Modell wird einmal mit Zeitreihen des Warmwasserverbrauchs trainiert. Ein zweites Modell gleicher Architektur wird neben den Zeitreihen des Warmwasserverbrauchs mit zusätzlichen Features „ausgestattet“ und trainiert. Das Ziel dieses Ansatzes war die Ermittlung des Einflusses der zusätzlichen Features auf die Qualität der Prognosen. Die zusätzlichen Features umfassen: die Tageszeit, den Wochentag, ob es sich um einen Feiertag handelt und autoregressive Informationen (zeitlichen Verschiebung der Zeitreihen).

Grundsätzlich können LSTM networks als Erweiterung zu Recurrent Neural Networks (RNN) gesehen werden. Beide Modelle können Informationen in einem inneren Zustand (Gedächtnis – siehe Beschreibung LSTM oben) speichern. Im Vergleich zu RNN können mittels LSTM aufgrund ihrer inneren Struktur auch lange sequentielle Zeitreihen verarbeitet werden.

- Feedforward Neuronales Netz mit Warmwasserverbrauchs-Zeitreihen. Eine Ebene in einem Neuronales Netz wird durch eine Vielzahl von Neuronen gebildet, wobei jedes Neuron einer Ebene dieselben Input-Daten erhält. Die Ausgangswerte jeder Ebene stellt ein Vektor dar, der die Ausgangsdaten jedes einzelnen Neurons dieser Ebene enthält. Ein Vektor wird von einer Ebene (als Eingangssignal) verarbeitet, der Ausgangsvektor dient wiederum der nächsten Ebene als Eingang. Werden mehrere dieser Ebenen gestapelt, wird dieser Vorgang als Deep Learning bezeichnet. Abbildung 29 zeigt ein einfaches Beispiel eines Neuronales Netzes, in dem fünf Neuronen eine Ebene bilden. Durch die Stapelung mehrerer solcher Ebenen kann ein beliebig großes Netz gebildet werden. Unter einem vollständig-

verbundenen Netz versteht man ein Netzwerk, bei dem der Ausgang einer Ebene als Eingang jedes einzelnen Neurons der Folgebene dient.

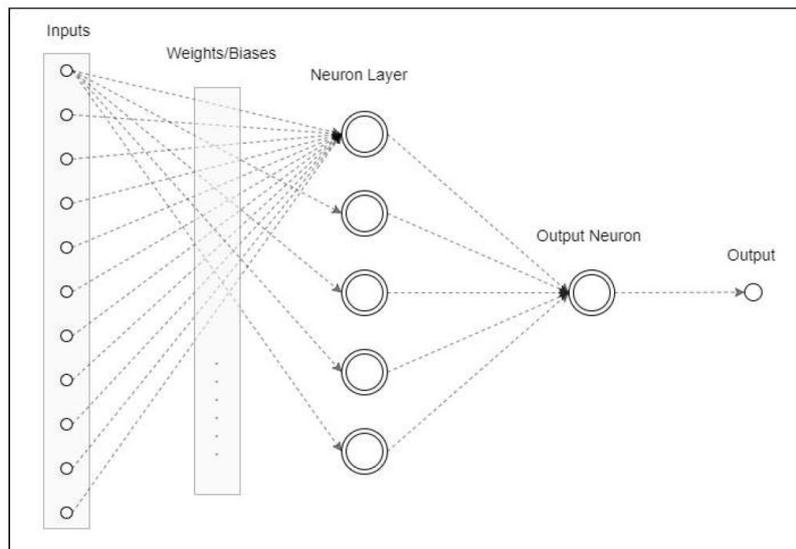


Abbildung 29 Vollständig-verbundenes Feedforward Neural Network.

Die Modelle wurden zur Analyse der Prognosequalität mit einer Benchmark (Tag zuvor, Wochen-Durchschnitt) verglichen. Zusätzlich wurde ein „Allgemeines Modell“ mit einem „Individuellen Modell“ in Bezug auf die Prognosequalität verglichen. Beim Allgemeinen Modell erfolgt das Training des Neuronalen Netzes auf Basis historischer Verbrauchsdaten aller Haushalte, beim „Individuellen Modell“ wird jeweils nur der Haushalt-spezifische historische Datensatz für das Training verwendet.

5.4.2 Error- und Performance-Metric

Um die Qualität von Prognosemodellen erheben zu können, werden zwei Error-Metriken angewendet:

- Error-Metric: Mittels Mean-Absolute-Error (MAE) kann die Abweichung zwischen Prognose-Zeitreihen und den tatsächlichen Messwerten angegeben werden:

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

wobei n für die Länge der Zeitreihe steht (24 für eine stündliche Auflösung oder 96 für eine Auflösung von 15 Minuten). Die berechnete Abweichung zeigt die Abweichung zwischen den Prognosewerten \hat{y}_t und den tatsächlichen Werten y_t gemittelt über alle Zeitschritte t . Jedoch führt dieses Verfahren aufgrund des punktwisen Vergleichs aufgrund eines möglicherweise auftretenden „Double Penalty Errors“ zu hohen Abweichungen: Der „Double Penalty Error“ beschreibt eine Situation, in der der tatsächliche Bedarf und die Prognose-Zeitreihe einen ähnlichen Verlauf aufweisen, jedoch um wenige Zeitschritte versetzt (siehe Abbildung 29). Im Extremfall sind beide Zeitreihen identisch, jedoch um einen einzelnen Zeitschritt versetzt. Grundsätzlich würde es sich in diesem Fall um eine qualitativ hochwertige Prognose handeln – vor allem in jenen Anwendungen, in denen der zeitliche Aspekt vernachlässigt werden kann. Durch die punktweise Berechnung der Abweichung zwischen Prognose- und Messdaten in

jedem Zeitschritt, kann sich eine hohe Abweichung zum Zeitschritt t sowie im nachfolgenden Zeitschritt ergeben (wodurch sich die Bezeichnung des Double Penalty Errors ergibt).

- Performance Metric: Aufgrund der zuvor beschriebenen Schwierigkeit der punktwisen Error-Metric, wurde ein zweites Verfahren eingesetzt. Zusätzlich zur mittleren Abweichung MAE wird die Abweichung der gesamten Energiemenge berücksichtigt, wodurch sich eine Verbesserung in der Prognose-Qualität ergibt:

$$EnergyDeviation = \left| \sum_{t=1}^n \hat{y}_t - \sum_{t=1}^n y_t \right|$$

wobei sich n wiederum auf die Länge der Zeitreihen und t auf die Zeitschritte bezieht. Der erste Term der rechten Seite der angegebenen Gleichung enthält die prognostizierte Gesamtenergie an einem Tag und der rechte Term die tatsächlich gemessene Energie.

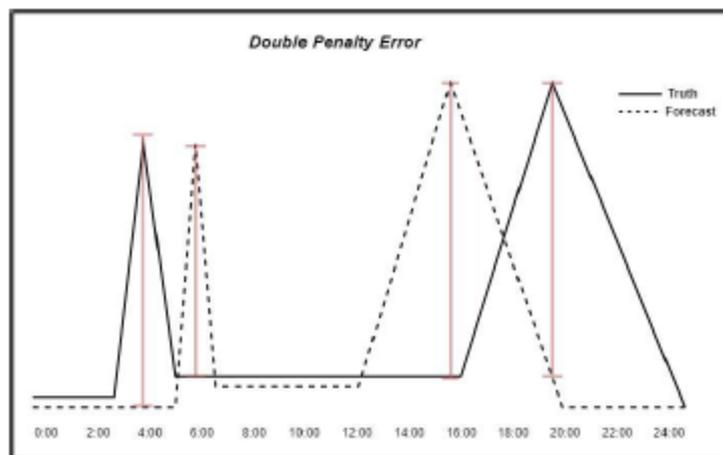


Abbildung 30 Grafische Darstellung des Double Penalty Errors

5.4.3 Trainings- und Testdaten

Die vorhandenen Daten über den Warmwasserverbrauch wurden in ein Set von Trainingsdaten sowie in ein Set von Testdaten eingeteilt. Die Modelle lernen anhand der Trainingsdaten die optimale Einstellung der Parameter, die den Zusammenhang zwischen Eingangs- und Ausgangsdaten beschreiben. Anschließend werden die Modelle mittels Testdaten validiert um zu prüfen, ob auch unvorhersehbare Eingangsdaten zu verwendbaren Ergebnissen führen.

Die Aufteilung der Daten erfolgte in einem Verhältnis von 3:1, das bedeutet, dass 75 % der Daten für den Lernprozess und 25 % der Daten für die Validierung der Modelle verwendet wurden.

5.4.4 Ergebnisse

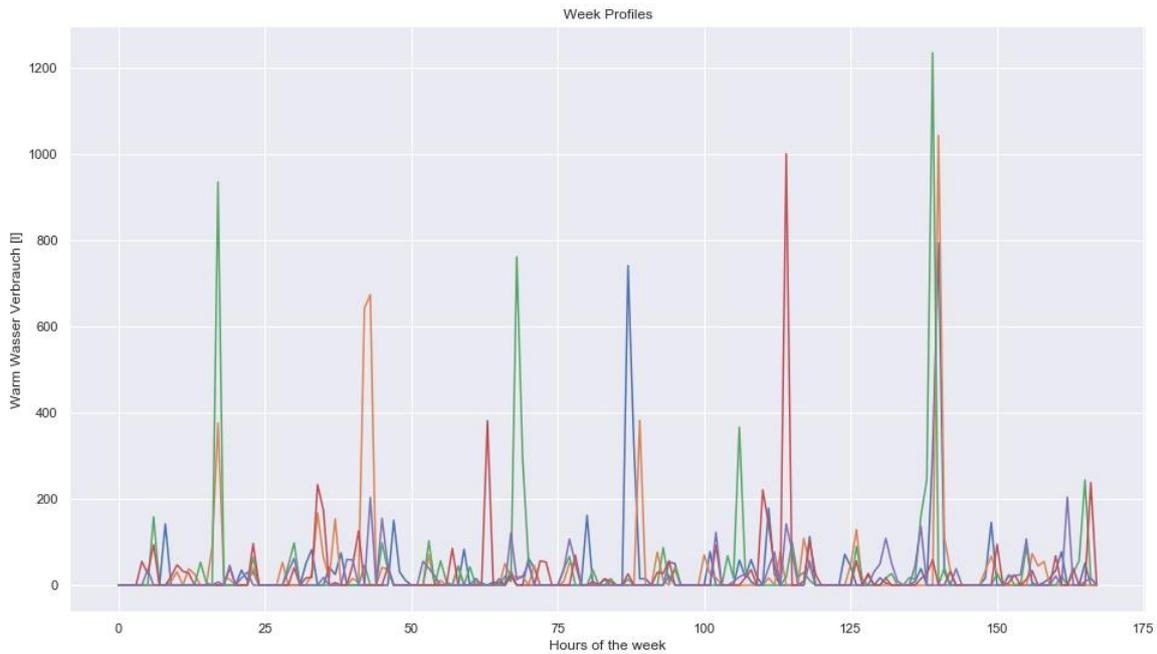


Abbildung 31 Warmwasser-Verbrauch einer Woche (5 Haushalte)

Abbildung 31 zeigt den Warmwasser-Verbrauch von 5 Haushalten einer Woche in einer Auflösung von 15 Minuten. Anhand dieser Grafik ist der dynamische und sehr volatile Wasserverbrauch erkennbar. Dies führt zu einer enormen Herausforderung an die Modelle zur Vorhersage von zukünftigen Warmwasserentnahmen und zu hohen Fehlern zwischen den prognostizierten Zeitreihen und den tatsächlichen Entnahmen.

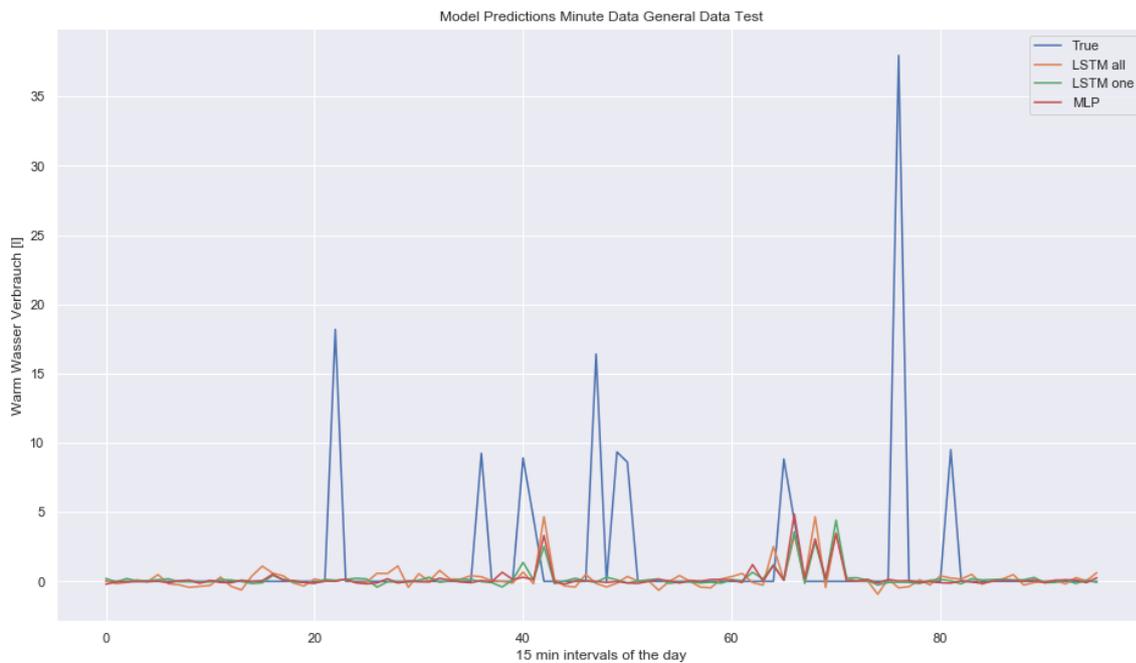


Abbildung 32 Vergleich zwischen dem tatsächlichen Warmwasser-Verbrauch und den Ergebnissen der Prognosemodelle in einer Auflösung von 15 Minuten.

Abbildung 32 zeigt den tatsächlichen Warmwasser-Verbrauch eines ausgewählten Haushaltes (blau „True“) in einer zeitlichen Auflösung von 15 Minuten sowie die Ergebnisse

der Prognose. Hierbei wurde ein generelles Modell trainiert (auf Basis der historischen Verbrauchsdaten aller Haushalte), das als Basis für die Prognose verwendet wurde.

- „LSTM all“ (Long short-term memory with several model input parameters) zeigt die Vorhersage des Verbrauchs unter Verwendung zusätzlicher Features (Wochentag, Feiertagsinformation, Stunde des Tages, Monat, Wohnungsgröße) und der historischen Verbrauchsdaten.
- „LSTM one“ (Long short-term memory with one model input parameter) zeigt die Vorhersage des Verbrauchs ohne Berücksichtigung zusätzlicher Informationen – es wurde lediglich der historische Verbrauch als Eingangsgröße für das Training und die Prognose verwendet.
- „MLP“ zeigt die Vorhersage des Verbrauchs auf Basis eines Feed Forward Neural Networks

Offensichtlich verhalten sich die drei Modelle sehr ähnlich – nach Zeitschritt 40 wird in den Modellen eine signifikante Entnahme prognostiziert, ebenso werden mehrere Entnahmen zwischen Zeitschritt 60 und 80 in allen Modellen prognostiziert. Es kann festgehalten werden, dass sich die Prognosen und der tatsächliche Verlauf weder in der Menge noch in den Zeitpunkten des Verbrauchs beim generellen Modell mit einer zeitlichen Auflösung von 15 Minuten decken.

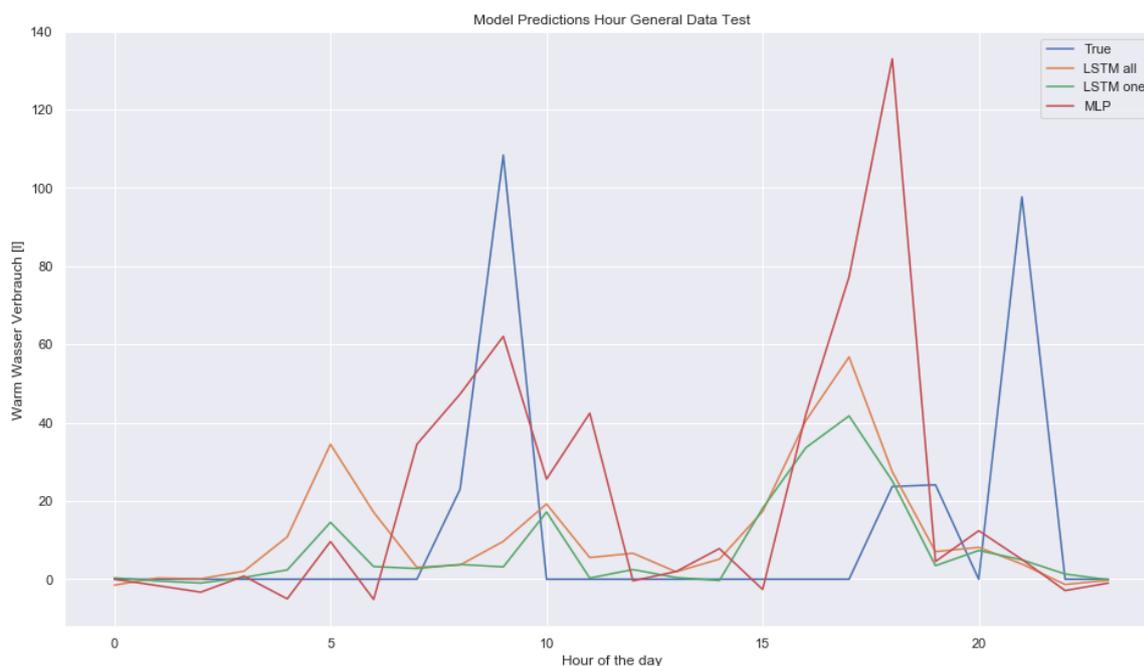


Abbildung 33 Vergleich zwischen dem tatsächlichen Warmwasser-Verbrauch und den Ergebnissen der Prognosemodelle in einer Auflösung von 60 Minuten.

Abbildung 33 zeigt die Ergebnisse der drei zuvor erwähnten Prognosemodelle im Vergleich mit dem tatsächlichen Verlauf in einer zeitlichen Auflösung von einer Stunde (im Vergleich zu Abbildung 32, bei der eine Auflösung von 15 Minuten verwendet wurde). In diesem Fall werden die Unterschiede zwischen LSTM und MLP deutlich sichtbar, wobei bei MLP höhere Entnahmen prognostiziert werden. Grundsätzlich ist aber auch hier eine signifikante Abweichung zwischen Prognose und tatsächlichem Verbrauch erkennbar.

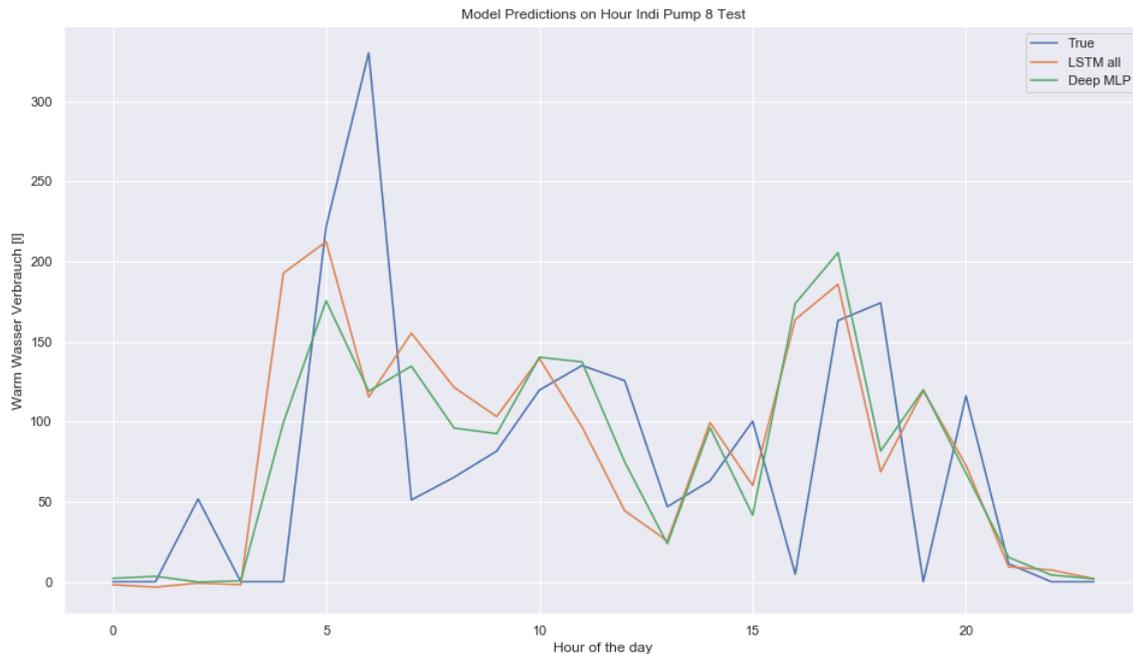


Abbildung 34 Warmwasser-Entnahme einer Wärmepumpe in einer zeitlichen Auflösung von einer Stunde.

Abbildung 34 zeigt die Warmwasser-Entnahme einer ausgewählten Wärmepumpe („True“ illustriert die tatsächliche Entnahme, orange und grün zeigen die prognostizierten Zeitreihen des Verbrauchs). In diesem Fall konnten deutlich bessere Ergebnisse im Vergleich zu den vorherigen Beispielen erzielt werden – sowohl die Entnahme-Mengen als auch deren Zeitpunkte konnten in beiden Modellen relativ gut prognostiziert werden.

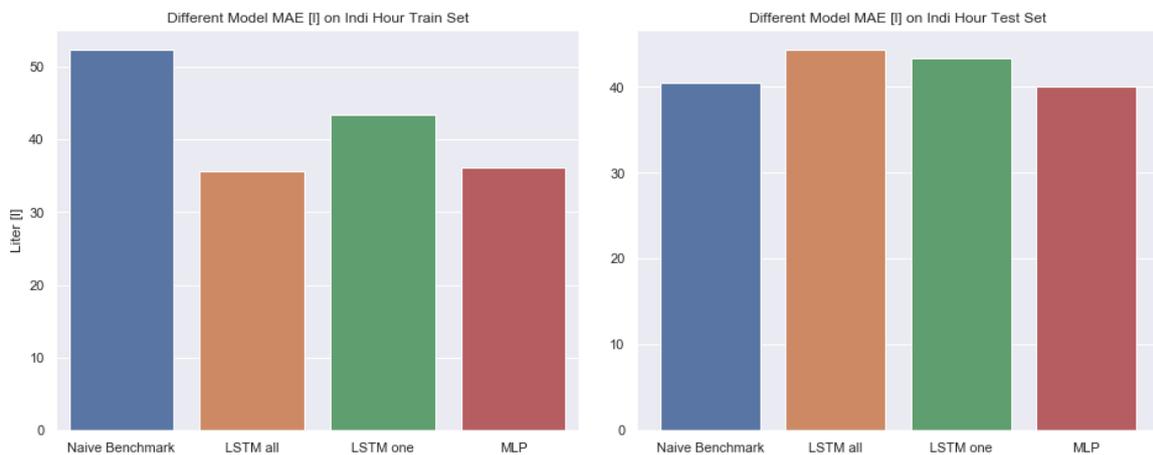


Abbildung 35 Mittlerer absoluter Fehler der Modelle für die Trainings- und Testdaten.

Abbildung 35 zeigt den mittleren absoluten Fehler (MAE) in Liter für die Benchmark und die drei implementierten Prognosemodelle sowohl für die Trainings- (links) als auch die Testdaten (rechts). Aufgrund dieser Ergebnisse ist erkennbar, dass sich die LSTM-Prognosemodelle unwesentlich unterscheiden.

Zusammenfassung der Ergebnisse:

- Die Prognose des Warmwasser-Verbrauchs auf Einzelhaushaltsebene ist aufgrund der hohen Dynamik sowie des volatilen Verhaltens äußerst schwierig zu prognostizieren und führt zu hohen Abweichungen

zwischen dem tatsächlichen Verlauf und den prognostizierten Zeitreihen.

- Die Zeitreihen des Warmwasser-Verbrauchs enthalten an vielen Zeitpunkten den Wert 0 (da an diesen Zeitpunkten keine Entnahme stattgefunden hat). Zusätzlich variiert der zeitliche Verlauf der Entnahmen sehr deutlich. Das führt letztendlich dazu, dass sich durch die gewählte Fehler-Metrik die Abweichung zwischen Prognose und Realdaten doppelt auswirkt (Double Penalty Fault).
- Die LSTM Prognosemodelle liefern keine besseren Ergebnisse (in Bezug auf die Abweichung zwischen Prognose- und Realdaten) als einfache Feed Forward Neural Networks.
- Die Verwendung von zusätzlichen Features bei LSTM (Wochentag, Feiertag, Wohnungsgröße, etc.) führte zu keiner signifikanten Verbesserung der Prognosequalität.

Ausblick/Verbesserungspotential

- Zusammenfassung/Gruppierung (Clustering) von Kunden mit einem ähnlichen Verbrauchsprofil sowie Training der Modelle und Prognose des Verbrauchs pro Kundengruppe.
- Reduktion der zeitlichen Auflösung um einen Glättungseffekt zu erzielen.
- Vorab-Filterung der Verbrauchsdaten um außergewöhnliche (Urlaubszeit) und fehlerhafte (Messfehler) Zeitreihen aus den Trainingsdaten zu entfernen, um die Beeinflussung der Modelle durch unerwünschte Daten zu minimieren.
- Entwicklung und Nutzung einer angemessenen Fehler-Metrik für Zeitserien.

6 Referenzen

- Cesa-Bianchi, N., & Lugosi, G. (2006). *Prediction, Learning, and Games*. Cambridge, New York: Cambridge University Press.
- Corinaldesi, C., Lettner, G., Schwabeneder, D., Fleischhacker, A., Dallinger, B., Eibl, E., . . . Knapp, M. (2019). *Deliverable D5: Spezifikation der Geschäftsmodelle und deren Rahmenbedingungen*. Wien: Flexplus Projekt der Österreichische Forschungsförderungsgesellschaft.
- Deloche, F. (kein Datum). *Long Short-Term Memory, licensed under CC BY-SA 4.0*. Von https://commons.wikimedia.org/wiki/File:Long_Short-Term_Memory.svg abgerufen
- Feixiang, X. (2011). Research on electric vehicle charging station load forecasting. *IEEE*.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Grubinger, T., Chasparis, G., & Natschlaeger, T. (2017). Generalized Online Transfer Learning for Climate Control in Residential Buildings. *Energy and Buildings*, 63-71.
- Hagan, M., & Behr, S. (1987). The time series approach to short term load forecasting. *IEEE Transactions on Power Systems*, (S. 785-791).
- Hochreiter, S., & Schmidhuber, J. (11 1997). Long short-term memory. *Neural Comput.*, S. 1735-1780.
- Hyndman, R., & Athanasopoulos, G. (2018). *Forecasting: principles and practice, 2nd edition*. Melbourne, Australia: Otexts.com.
- Hyndman, R., & Koehler, A. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, S. 679-688.
- Ljung, L. (1999). *System Identification*. Prentice Hall PTR.
- Majid, R., & Mir, S. (12 2018). Advances in statistical forecasting methods: An overview. *Economic Affairs*, S. 815-831.
- Notton, G., & Voyant, C. (2018). Forecasting of Intermittent Solar Energy Resource. In I. Yahyaoui, *Advances in renewable energies and power technologies: Solar and wind energies* (S. 77-114). Elsevier Inc.
- Pollok, T., Hille, C., & Schnettler, A. (2009). *Netzeinbindung von Elektrofahrzeugen – Möglichkeiten flexibler Ladekonzepte*. Düsseldorf, Deutschland: Internationaler ETG-Kongress 2009.
- Sayed, A. (2003). *Fundamentals of Adaptive Filtering*. New Jersey: John Wiley & Sons, Inc.
- Sokorai, P., & A. Fleischhacker, G. L. (2018). Stochastic Modeling of the Charging Behavior of Electromobility. *World electric vehicles journal*, 9.
- Szmit, M., Szmit, A., Slawomir, A., & Bugala, S. (2012). Implementation of Brutlag's algorithm in Anomaly Detection 3.0. *Proceedings of the Federated Conference on Computer Science and Information Systems* (S. 685-691). Wroclaw: IEEE.
- Winters, P. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, S. 324-342.
- Zhuowei, L. (2011). Forecasting charging load of plug-in electric vehicles in China. *IEEE*.

7 Anhang

7.1 Abbildungsverzeichnis

Abbildung 1 Wahrscheinlichkeitsdichtefunktionen von Plug-In-Zeiten.	10
Abbildung 2 Ladezeiten der unterschiedlichen Verbraucher	11
Abbildung 3 Unflexibler Lastverbrauch eines Haushalts in einem Monat (von 01.10.2016 bis 31.10.2016). ...	12
Abbildung 4 Unflexibler Lastverbrauch eines Haushalts an einem Tag (22.10.2016).	12
Abbildung 5 Schematische Lastprognose unflexibler Verbraucher für den nächsten Tag.	16
Abbildung 6 Persistenzmodell ("Copy-last-days") mit N=3 letzten (gleichen) Tagen, die über einen Monat ausgewertet wurden (vom 01.10.2016 bis 31.10.2016). Der resultierende RMSE=189. Wir verwenden die Abkürzung CLD zur Bezeichnung des "copy-last-days"-Persistenzmodells.	19
Abbildung 7 Persistenzmodell ("Copy-last-days") mit N=3, das über einen Tag (22.10.2016) ausgewertet wurde. Der resultierende RMSE=153.	19
Abbildung 8 Kombiniertes AR mit Persistenzmodell ("Copy-last-days") mit N=3 letzten (gleichen) Tagen, die über einen Monat ausgewertet wurden (vom 01.10.2016 bis 31.10.2016). Der resultierende RMSE=159.	20
Abbildung 9 Kombiniertes AR mit Persistenzmodell ("Copy-last-days") mit N=3 letzten (gleichen) Tagen, die über einen Tag (22.10.2016) ausgewertet wurden. Der resultierende RMSE=174.	20
Abbildung 10 Holt-Winters Auswertung über einen Monat (vom 01.10.2016 bis 31.10.2016) für eine saisonale Komponente von 1 Woche. Der resultierende RMSE=288.	21
Abbildung 11 Holt-Winters Auswertung über einen Tag (22.10.2016) für eine saisonale Komponente von 1 Woche. Der resultierende RMSE=305.	22
Abbildung 12 Holt-Winters Auswertung über einen Monat (vom 01.10.2016 bis 31.10.2016) und für eine saisonale Komponente von 1 Tag. Der resultierende RMSE=271.	22
Abbildung 13 Holt-Winters Auswertung über einen Tag (22.10.2016) und für eine saisonale Komponente von 1 Tag. Der resultierende RMSE=366.	23
Abbildung 14 SARIMA-Modell vom Typ (1,1,1)(1,1,1)96 mit einer Auswertung über einen Monat (vom 01.10.2016 bis 31.10.2016). Der resultierende RMSE=163.	23
Abbildung 15 SARIMA-Modell vom Typ (1,1,1)(1,1,1)96 mit einer Auswertung über einen Tag (22.10.2016). Der resultierende RMSE=211.	24
Abbildung 16 Feature-Extraktionsmethode über einen Monat (vom 01.10.2016 bis 31.10.2016). Der resultierende RMSE=82.	24
Abbildung 17 Feature-Extraktionsmethode mit einer Auswertung über einen Tag (22.10.2016). Der resultierende RMSE=103.	25
Abbildung 18 Die Schritte der Datenaufbereitung dienen zum Erstellen von passenden Eingabe-Daten für die Prognosemodelle.	26
Abbildung 19 Der Boiler modelliert als offenes thermodynamisches System.	26
Abbildung 20 Beispiele von unregelmäßigem und regelmäßigem Warmwasserverbrauch in zwei verschiedenen Haushalten	27
Abbildung 21 ACF und PACF Diagramme für ein regelmäßiges (obere Reihe) und unregelmäßiges (untere Reihe) Zapfprofil.	30
Abbildung 22 Die interne Struktur eines LSTMs. Quelle: (Deloche, kein Datum)	31
Abbildung 23 Zeitreihen-Kreuzvalidierung mit konstanter Trainings-Set-Größe und rollendem Zeitfenster... ..	32
Abbildung 24 Die Zeitreihen-Kreuzvalidierung für überwachtes maschinelles Lernen wird nur auf das Test-Set angewandt	32
Abbildung 25 Die Architektur des evaluierten LSTM-Modells.	33
Abbildung 26 Beispiel-Vorhersage des SARIMA-Modells.	34
Abbildung 27 Die Prognose des SD-LSTMs liefert konstant "0" bei einem unregelmäßigem Profil.	35
Abbildung 28 Gute Vorhersage des SD LSTMs mit zusätzlichem Feature "Wochentag"	35
Abbildung 29 Vollständig-verbundenes Feedforward Neural Network.	37
Abbildung 30 Grafische Darstellung des Double Penalty Errors	38
Abbildung 31: Warmwasser-Verbrauch einer Woche (5 Haushalte)	39

Abbildung 32: Vergleich zwischen dem tatsächlichen Warmwasser-Verbrauch und den Ergebnissen der Prognosemodelle in einer Auflösung von 15 Minuten.	39
Abbildung 33: Vergleich zwischen dem tatsächlichen Warmwasser-Verbrauch und den Ergebnissen der Prognosemodelle in einer Auflösung von 60 Minuten.	40
Abbildung 34: Warmwasser-Entnahme einer Wärmepumpe in einer zeitlichen Auflösung von einer Stunde.	41
Abbildung 35: Mittlerer absoluter Fehler der Modelle für die Trainings- und Testdaten.	41

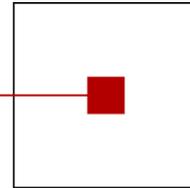
7.2 Tabellenverzeichnis

Tabelle 1 Feature-Kategorien im Feature-Extraktionsmodell	17
Tabelle 2 Merkmale, die im Feature-Extraktionsmodell als Regressoren betrachtet werden und zugeordnete Gewichte.	17
Tabelle 3 RMSE-Vergleich der untersuchten Lastprognosemethoden.	25
Tabelle 4 Ergebnisse der Evaluierung in Form der Fehlermetriken Mean Absolute Scaled Error (MASE) und Mean Absolute Error (MAE).	34

Kontakt

s c c h

software competence center
hagenberg



Hauptautor

Georgios Chasparis, Ph.D.

Software Competence Center Hagenberg GmbH

Softwarepark 21, 4232 Hagenberg, Austria

Georgios.Chasparis@scch.at

T +43 50 343 857



Konsortialleitung

Tara Esterl MSc. MSc.

Center for Energy – AIT Austrian Institute of Technology GmbH

Giefinggasse 2 | 1210 Vienna | Austria

T +43 50550-6077 | M +43 664 8157810 | F +43 50550-6390

tara.esterl@ait.ac.at | www.ait.ac.at