Vorhersage von Energiedaten im Projekt WINNER Reloaded

Tobias Jung

Mai 2022

Inhaltsverzeichnis

1 Einleitung **2**
1.1 Tanzende Siedlung . . . 2
1.2 Alfons-Pech-Straße . . . 3
1.3 Zielstellung . . . 3
1.4 Aufbau der Arbeit . . . 4

2 Long Short Term Memory Systems **5**

3 Datenaufbereitung **6**

4 Machbarkeitsanalyse **8**
4.1 Alfons-Pech-Straße . . . 8
4.2 Tanzende Siedlung . . . 11

5 Aufbau eines LSTM-Modells **15**
5.1 Wahl der Parameter . . . 15
5.2 Vergleich der Modelle . . . 17
5.3 Modelle für die Alfons-Pech-Straße . . . 17
5.3.1 60 Minuten Vorhersage . . . 17
5.3.2 24 Stunden Vorhersage . . . 18
5.3.3 7 Tage Vorhersage . . . 18
5.4 Modelle für die Tanzende Siedlung . . . 19
5.4.1 Eine Stunde Vorhersage Einspeisung . . . 20
5.4.2 24 Stunden Vorhersage Einspeisung . . . 21
5.4.3 Sieben Tage Vorhersage Einspeisung . . . 22
5.4.4 Eine Stunde Vorhersage Verbrauch . . . 22
5.4.5 24 Stunden Vorhersage Verbrauch . . . 23
5.4.6 Sieben Tage Vorhersage Verbrauch . . . 23

6 Analyse des Vorhersagen **24**
6.1 Vorhersage über die PV-Stromproduktion in der Alfons-Pech-Straße . . . 24
6.2 Vorhersage über die Einspeisung der tanzenden Siedlung . . . 25
6.3 Vorhersage über den Verbrauch der tanzenden Siedlung . . . 27

7 Fazit **29**

8 Ausblick **29**

1 Einleitung

“WINNER” (Wohnungswirtschaftlich INtegrierte Netzneutrale Elektromobilität in Quartier und Region) ist ein Projekt eines Konsortiums unter der Führung der Chemnitzer Siedlungsgemeinschaft eG (CSg), an dem sich zudem der Verband Sächsischer Wohnungsgenossenschaften e.V. (VSWG), die GEMAG Gebäudemanagement AG, die HEOS Energy GmbH, die NSC GmbH, die Friedrich-Schiller-Universität Jena und teilAuto beteiligen. Ziel des Projektes ist es, Mietwohnungen und Ladesäulen für Elektrofahrzeuge mit eigens erzeugtem Photovolatikstrom zu versorgen und daraus ein Geschätsmodell zu entwerfen, dass anschließend in ganz Deutschland Anwendung finden kann. Der selbsterzeugte Strom wird den Mietern als “Mieterstrom” zur eigenen Nutzung sowie zum Aufladen der Elektrofahrzeuge zur Verfügung gestellt. Dieses Projekt wurde in der Alfons-Pech-Straße in Chemnitz durchgeführt, das Projekt endete Ende 2019. Auf das Quartier Alfons-Pech-Straße wird im Kapitel 1.2 näher eingegangen werden. [21][17]

Unter dem Namen “WINNER Reloaded” wurde das Projekt “WINNER” fortgeführt. Das Projekt hat sich ein ähnliches Ziel gesetzt, wie bereits “WINNER” zuvor, nämlich an vier verschiedenen Standorten (“Demonstratoren”) “bedarfsgerechter Ladeinfrastruktur, Mieterstrom und dezentraler Energieerzeugung für Mieter, Gewerbetreibende und sonstige Drittnutzer” [16] umzusetzen. Dafür werden die Demonstratoren mit Photovoltaikanlagen und öffentlicher sowie privater Ladeinfrastruktur für Elektrofahrzeuge ausgerüstet. An den vier Standorten werden “dynamische Preis- und Abrechnungsmodelle erprobt, die von der Nutzungszeit und der intern erzeugten und bereitgestellten Leistung abhängen”. Hauptdemonstrator ist die Tanzende Siedlung in Chemnitz, die im Kapitel 1.1 näher vorgestellt wird. [16]

Dabei sollen für die unterschiedlichen Beteiligten an dem Pojekt verschiedene Ziele erreicht werden. Für die Mieter soll die Möglichkeit geschaffen werden, Elektromobilität umzusetzen sowie die Nebenkosten zu begrenzen. Aus Sicht der Wohnungswirtschaft stehen natürlich vor allem wirtschaftliche Interessen im Fokus: So soll hier zunächst ein neues Geschäftsmodell entwickelt werden, in dem auch neue Geschäftsfelder erschlossen werden, während gleichzeitig steuerliche und rechtliche Fragen geklärt werden müssen. Nicht zuletzt soll auch die Mieterbindung erhöht werden. Im Fokus des Projektes stehen die Themen Energiewende und Umweltschutz: So sind weitere Ziele des Projektes die Reduktion von Emissionen und Straßenlärm, die Voranbringung der Energiewende, die Entlastung des Stromnetzes sowie die effiziente Nutzung von Energie. [16][12]

Die Stromversorgungsnetze der Demonstratoren werden als “Micro Smart Grids” bezeichnet, die von der Firma HEOS Energy GmbH geplant und errichtet wurden. Diese Smart Grids werden auch als “intelligentes, dynmaische Stromversorgungsnetz im Kleinformat” [16] bezeichnet. An verschiedenen Stellen diese Smart Grids werden dezentral Daten erfasst, welche der Universität Jena zur Auswertung zur Verfügung gestellt werden und auf deren Basis diese Projektarbeit erstellt wurde. [16]

1.1 Tanzende Siedlung

Die “Tanzende Siedlung” in Chemnitz ist ein Häuserblock, bestehend aus vier Häusern mit 40 Wohneinheiten - im weiteren als *Quartier* bezeichnet - der Chemnitzer Siedlungsgemeinschaft (CSg). Zusätzlich zum regulären “Hausverbrauch” der Mieter, also dem Strom, der in den Wohnungen verbraucht wird, versorgt das Quartier eine Reihe an Parkplätzen, die mit insgesamt 73 Ladepunkten für Elektrofahrzeuge ausgestattet sind. Das sind für die Mieter selbst die Stellplätze der Tiefgarage sowie die Stellplätze unter den Carports, sowie für die Öffentlichkeit zugängliche Parkplätze und Car-Sharing-Plätze. Darüber hinaus sind einige Parkplätze mit High Performance Chargern (HPCs) ausgestattet, also Säulen für das schnelle Laden von Elektrofahrzeugen. [16][12]

Auf den Dächern der Häuser sind Photovoltaikanlagen installiert, deren Erzeugungs - und Verbrauchsdaten jeweils durch einen Zähler erfasst werden. Für das hiesige Projekt stehen dabei aktuell nur die Daten des Hauses 4 zur Verfügung. In Zeiträumen, in denen keine Solarenergie produziert wird, verbrauchen PV-Anlagen in geringem Maße Energie, die für das Betreiben der internen Infrastruktur, wie beispielsweise der Wechselrichter benötigt wird. Da der durch die PV-Analgen produzierte Strom nicht ausreicht, um die gesamte Anlage zu betreiben, beziehungsweise nicht immer bedarfsgerecht produziert wird, ist das Quartier zusätzlich an das öffentliche Stromnetz angeschlossen. Wird von der PV-Analge mehr Strom erzeugt, als im Quartier verbraucht wird,

so wird Strom in das öffentliche Netz eingespeist. Wird hingegen von der PV-Analge weniger Strom produziert als aktuell im Quartier benötigt, so wird der fehlende Strom aus dem öffentlichen Netz bezogen. Die Einspeise- und Verbrauchsdaten des Quartiers aus dem und ins öffentlichen Stromnetz werden über einen Zwei-Richtungszähler, der sich in der Trafostation befindet, erhoben und stehen im Projekt ebenfalls zur Verfügung. Die Messwerte liegen für den Zeitraum vom 01.01.2020 bis zum 31.12.2020 mit einer zeitlichen Auflösung von fünfzehn Minuten vor. Eine Übersicht aller verfügbaren Messwerte des Quartiers "Tanzende Siedlung" bietet Tabelle 1. Abbildung 1 veranschaulicht die Architektur der Stromversorgung der tanzenden Siedlung.

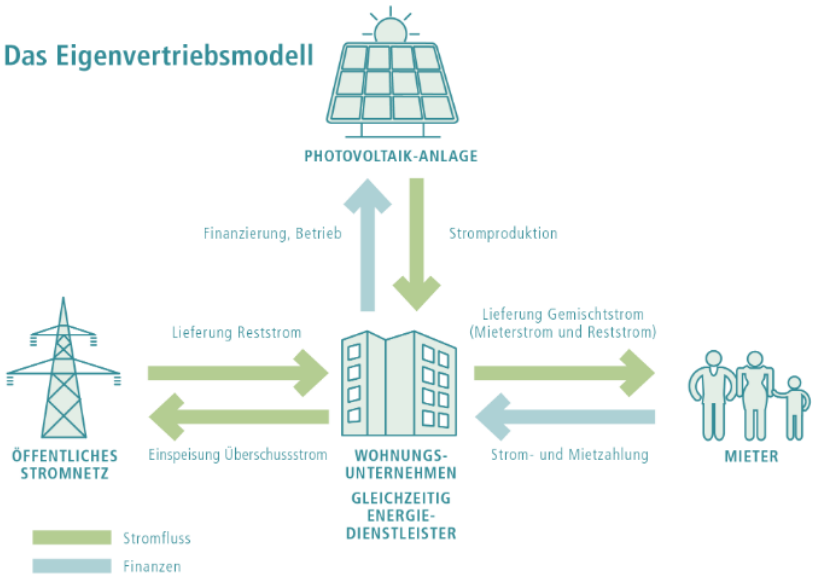


Abbildung 1: Schematische Darstellung der Verwaltung des Stromflusses in der Tanzenden Siedlung. Abbildung aus [12]

1.2 Alfons-Pech-Straße

Das Quartier an der Alfons-Pech-Straße in Chemnitz besteht aus 8 Wohnblöcken mit insgesamt 273 Wohneinheiten. Im Rahmen des Projektes "WINNER" wurde das Quartier mit einer Photovoltaikanlage mit einer Maximalleistung von 34,56 kWp (Kilowatt-Peak) auf dem südseitigen Dach sowie mit zwei Ladesäulen mit jeweils zwei Ladepunkten für Elektrofahrzeuge ausgestattet. Die Ladesäulen können dabei nicht nur von den Mietern selbst genutzt werden, sondern stehen auch "Dienstleistern, die im Wohngebiet tätig sind" [19] sowie privaten Carsharingnutzern zur Verfügung. [19][12]. Für das hiesige Projekt stehen Messreihen des Erzeugungszählers zur Verfügung, der die Produktion der Photovoltaikanlage misst. Diese Daten liegen für den Zeitraum vom 01.04.2020 bis zum 16.05.2020 mit einer zeitlichen Auflösung von einer Minute vor.

1.3 Zielstellung

Ziel dieser Projektarbeit ist die Erstellung neuronaler Netzwerke, die für die oben genannten Quartiere / Demonstratoren Vorhersagen zur Energieproduktion sowie zum Energieverbrauch treffen, aus denen sich ein optimierter Betrieb der Anlagen ableiten lässt. Zunächst soll die grundlegende Machbarkeit überprüft werden, indem die vorliegenden Messreihen bezüglich des Vorhandenseins von Periodizitäten untersucht werden. Dann sollen potentielle Einflussgrößen identifiziert und erhoben werden, für die eine Korrelation mit der entsprechenden Zielgröße angenommen werden kann.

Die Ertstellung des Modells selber orientiert sich im wissenschaftlichen Artikel "LSTM based long-term energy consumption prediction with periodicity" von Wang et al. [22] Wie der Titel

Bezeichnung des Messwertes	**Zähler**	**Beschreibung**
0010498954/Chemnitz/ Kaßbergstr.20 /Haus4/LG_Wirk+	Zwei-Richtungszähler (in der Trafostation)	Netzbezug durch das gesamete Quartier
0010498954/Chemnitz/ Kaßbergstr.20 /Haus4/LG_Wirk-	Zwei-Richtungszähler (in der Trafostation)	Netzeinspeisung durch das Quartier (PV-Erzeugung, die nicht im Quartier verbraucht wird)
0010504253/Chemnitz/ Kaßbergstr.20/Haus4/LG_Wirk+	Erzeugungszähler (im WINNER-Raum Haus 3)	Einspeisung der PV-Anlage ins Quartier
0010504253/Chemnitz/ Kaßbergstr.20/Haus4/LG_Wirk-	Erzeugungszähler (im WINNER-Raum Haus 3)	Bezug durch PV-Analge: Energie, die für den Betrieb der PV-Infrastruktur notwendig war, aber nicht durch die PV-Analge erzeugt wurde

Tabelle 1: (relevante) Messwerte, die für das Quartier "Tanzende Siedlung" zur Verfügung stehen

bereits andeutet, wird dort ein LSTM-Modell (Long short-term memory) trainiert und ausgewertet, das Energieverbrauchsdaten eines Kühlsystems vorhersagt. Für die drei Zielgrößen PV-Stromproduktion der Alfons-Pech-Straße, PV-Einspeisung der Tanzenden Siedlung und Gesamtverbrauch der Tanzenden Siedlung sollen jeweils drei Modelle erstellt werden, die kurz- mittel und langfristige Vorhersagen treffen.

Schließlich sollen die erstellten Modelle validiert werden: Zunächst soll dies auf grafische Art und Weise geschehen, indem die vorhergesagten und die tatsächlich gemessenen Werte gegeneinander visualisiert werden. Zusätzlich soll die Vorhersagekraft des Modells auch rechnerisch aufgezeigt werden, indem der RMSE (Root Mean Square Error) je Modell erhoben und ausgewertet wird.

1.4 Aufbau der Arbeit

Nachdem das Kapitel 1 eine Einführung in die grundlegende Umstände und die der Arbeit zu Grunde liegende Problematik und Aufgabe gegeben hat, wird im nächsten Kapitel, dem Kapitel 2, grob die Funktionsweise des Long short-term Systems umrissen werden. Das Kapitel 3 befasst sich mit der Datenaufbereitung, also der Thematik, welche Größen als Kandidatengrößen identifiziert und erhoben werden. Im Kapitel 4 geht es um die Analyse der Ausgangsdaten, hinsichtlich gegebener Periodizität und somit genereller Machbarkeit sowie der Untersuchung von möglichen Korrelationen zwischen den zuvor erhobenen Einfluss- und den gegebenen Zielgrößen. Im Kapitel 5 wird die Erstellung und das Training der LSTM-Modelle thematisiert. Zunächst werden dabei die Eignungen verschiedener Parameter für das jeweilige Modell untersucht (Kapitel 5.3 & 5.4), anschließend werden im Kapitel 6 die so parametrisierten Modelle bezüglich ihrer Vorhersagekraft ausgewertet.

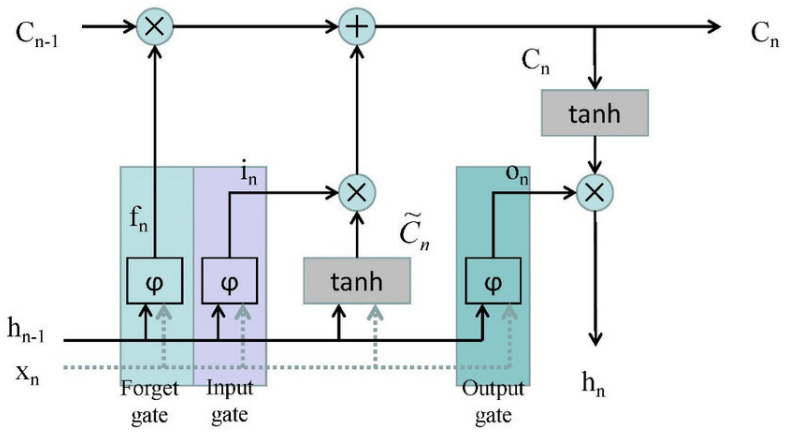


Abbildung 2: Schematische Darstellung eines Memory-Blocks in einem LSTMs. Abbildung aus [14]

2 Long Short Term Memory Systems

Long Short Term Memory Systems - oder kurz LSTMS - sind eine Sonderform der Recurrent Neural Networks (RNNs). Bei RNNs handelt es sich um neuronale Netzwerke, bei denen der Output eines Knotens als Input des selben Knotens verwendet wird. Somit wird es möglich, Informationen über vorangegangene Eingaben ins Netzwerk einzuspeisen. Diese Feedbackschleifen können Informationen, in diesem Falle als Kontext bezeichnet, jedoch lediglich über einen sehr begrenzten Zeitraum vorhalten. Daher wurden 1997 von Hochreiter und Schmidhuber [8] das LSTM vorgestellt. Dieses Modell verfügt über Mechanismen, die es erlauben, Kontext auch über sehr viel längere Zeit vorzuhalten, als das in herkömmlichen RNNs möglich ist. [13][1]

Im Gegensatz zu herkömmlichen RNNs, die aus Knoten mit jeweils einer Aktivierungsfunktion bestehen, werden Knoten in LSTMS als "Memory-Blocks" bezeichnet, wobei jeder Memory-Block aus vier Aktivierungsfunktionn besteht. Abbildung 2 zeigt einen solchen Memory-Block schematisch auf. Der Mechanismus, der es ermöglicht, Kontextinformationen über lange Zeit zu speichern ist der Zustand (Cell State), der in der Abbildung mit C gekennzeichnet ist. Im "Forget-Gate" wird zunächst entschieden, welche Informationen des Zustandes verworfen werden. Diese Entscheidung wird mithilfe einer Sigmoid-Funktion getroffen, die wiederum als Eingabe den Zustand des vorherigen Memory-Blocks sowie die aktuelle Eingabe erhält. Im Input Gate wird anschließend entschieden, welche Informationen geupdated werden sollen. Kombiniert mit der Ausgabe eines Tanh-Layers, das Kandidatenwerte ($\tilde{C}_n$) produziert, fließen diese Werte wiederum in den Status mit ein. Im Output gate wird zuletzt der Output des Memory-Blocks berechnet, der dann als Eingabe des nächsten Memeory-Blocks verwendet werden kann. [13]

Wie bereits angedeutet, verfügen LSTMs über die Fähigkeit, Kontext wesentlich länger vorzuhalten als herkömmliche RNNs. Laut [7] ist es mit LSTMs möglich "Zeitabstände von mehr als 1000 diskreten Zeitschritten zu überbrücken". Zudem sind LSTMS nicht vom Vanishing-Gradient-Problem, sowie dem umgekehrten Exploding-Gradient -Problem betroffen. Ersteres bezeichnet den Fall, dass beim Training des Modells keine nützlichen Gradienteninformationen von der Ausgabe hin zur Eingabe propagiert werden können. Mathematisch gesehen rührt dieser Fehler von der Multiplikation kleiner Werte zwischen 0 und 1 her, da die Multiplikation solcher Zahle stets ein Ergebnis liefert, dass kleiner als der kleinere der beiden Faktoren ist. In einem Netzwerk mit n hidden layers werden nun n Ableitungen miteinander multipliziert, was mit größerem n zu immer kleineren Werten beziehungsweise Änderungen führt. Für Module, die näher an der Eingabe- als an der Ausgabeschicht liegen, ändern sich Werte also nur minimal. Das Exploding-Gradient-Problem bezeichnet den umgekehrten Fall, in dem mit Werten über Eins multipliziert wird und die Größen der Werte somit über das gesamte Netz hinweg explodieren. [20][11][18]

Für eine ausführlichere Einführung sei an dieser Stelle auf entsprechende Literatur verwiesen.

3 Datenaufbereitung

Um aus den vorliegenden Daten sinnvolle und möglichst genaue Vorhersagen treffen zu können, mussten diese zunächst um eine Reihe von Kandidatenwerten erweitert werden. Da die Umsetzung des Projektes in Python erfolgte, mussten die Daten dafür erst einmal in ein Format gebracht werden, mit dem innerhalb von Python weitergearbeitet werden kann. Dazu bot sich das Array aus der Blibothek numpy an. Im ersten Schritt wurden die verfügbaren Messwerte in ein solches Array geladen. Anschließend wurden mögliche Kandidatenwerte identifiziert, die zur Berechnung eines Modells hilfreich sein können. Dabei lag zunächst nahe, Informationen aus dem Zeitstempel beziehungsweise der Datumsangabe zu extrahieren.

Da sich der Sonnenstand und damit die voraussichtliche Photovolatikleistung über den Tag hinweg ändert, wurde die Uhrzeit als erster Kandidatenwert berechnet. Um diese als sinnvolle Vorhersagegröße verwenden zu können, musste sie als kontinuierliche Größe dargestellt werden. Um dies zu gewährleisten wurde die Uhrzeit in die Einheit "Sekunden, die seit Mitternacht vergangen sind" umgerechnet und darauf wiederum eine Kosinus-Funktion berechnet. Mehr dazu in Kapitel 4. Die nächsten Kandidatenwerte, die aus der Zeit gewonnen werden konnten, waren der Wochentag und die Angabe, ob die entsprechenden Werte am Wochenende erhoben wurde. Bei der Identifizierung dieses Kandidatenwertes wurde angenommen, dass der Energieverbrauch unter der Woche geringer ist als am Wochenende, da sowohl arbeitstätige als auch schulpflichtige Menschen unter der Woche weniger Zeit zu Hause verbringen. Dieser Wert wird als einstellige Zahl zwischen 0 (Monatg) und 6 (Sonntag) angegeben, der später vom Framework in die One-Hot Vektorrepräsentation konvertiert wird. Der letzte Kandidatenwert, der direkt aus dem Datumsstempel erhoben werden konnte ist die Wochennummer. Diese wurde als mögliche Einflussgröße erachtet, da sich der Sonnenstand über das Jahr hinweg ändert, was wiederum Auswirkungen auf die Solarstromproduktion hat.

Als weiterer Werte, die Aufschluss über die An- beziehungsweise Abwesenheit der Mieter in ihren Wohnungen geben könnten und somit als Kandidatenwerte in Frage kommen wurden Schulferien und Feiertage identifiziert. Beide Informationen wurden von entsprechenden APIs bezogen, die diese Information für das entsprechende Bundesland, also in diesem Falle Thüringen, zur Verfügung stellen. Die wohl vielversprechendsten Daten bezüglich der Solarstromproduktion bieten jedoch die Wetterinformationen, vor allem Messgrößen, die in Zusammenhang mit der Sonneneinstrahlung stehen. Entsprechende Messgrößen werden vom deutschen Wetterdienst in einem Messintervall von zehn Minuten zur Verfügung gestellt: Konkret werden hier die diffuse Himmelstrahlung (Sonnenstrahlung, die die Erdoberfläche erreicht, nachdem sie durch Moleküle oder Partikel in der Atmosphäre vom direkten Sonnenstrahl gestreut wurde [4]), die Globalstrahlung (am Boden von einer horizontalen Ebene empfangene Sonnenstrahlung, [23]), die Sonnenscheindauer (tatsächliche Dauer der direkten Sonnenstrahlung an einem bestimmten Ort innerhalb eines definierten Zeitraumes [23]) sowie die langwellige Strahlung (= Wärmestrahlung [23]) zur Verfügung gestellt. Jedoch kommen nicht allein Daten über die Sonneneinstrahlung in Betracht, sondern auch Daten über den Niederschlag, da Regen den Einfall von Sonnenstrahlen auf die PV-Anlage behindert. Diese werden ebenfalls vom DWD mit einem Messintervall von zehn Minuten zur Verfügung gestellt. Hier stehen die Messwerte Niederschlagsdauer, Summe der Niederschlagshöhe der vorangegangenen 10 Minuten sowie ein Niederschlagsindikator ([Angabe,] ob im Bezugszeitraum Niederschlag aufgetreten ist [5, S. 46]) zur Verfügung.

Da für die Tanzende Siedlung sowohl Produktions- als auch Verbrauchsdaten (der gesamten Anlage, nicht der einzelnen Mieter) zur Verfügung stehen, kann hier auch der Gesamtenergieverbrauch der Anlage errechnet und nicht nur als Zielgröße, sondern auch als Kandidatenwert für die Vorhersage der Einspeisung in Betracht gezogen werden. Er wird indirekt aus dem Netzbezug des Quartiers aus dem öffentlichen Stromnetz, der Netzeinspeisung durch das Quartier (Energie, die die PV-Anlage produziert, die aber im Quartier nicht benötigt wird und somit ins öffentliche Netz eingespeist wird) sowie dem Bezug der PV-Anlage errechnet. Zudem wurden hier bereits Messwerte erhoben, bevor alle Wohnungen des Quartiers belegt waren. Die Belegung des Hauses, beziehungsweise die Angabe darüber, wie viele der verfügbaren Wohnungen vermietet sind, wird daher als zusätzlicher Kandidatenwert aufgenommen. Eine Übersicht über diese beschriebenen Werte inklusive ihrer Bezeichnungen, die im weiteren auch in den aufgezeigten Diagrammen Verwendung finden werden, bietet Tabelle 2.

Name	**Beschreibung**	**Datenquelle**
Zeit	Kosinus der Sekunden, die seit Mitternacht des Tages vergangen sind	Zeitstempel
Wochentag	einzelne Ziffer, 0 für Montag, 6 für Sonntag	Zeitstempel
Wochenende	Liegt der betreffende Tag an einem Wochenende	Zeitstempel
Wochennummer		Zeitstempel
Feiertag	ist der betreffende Tag in Thüringen ein Feiertag, 1 oder 0	https://feiertage-api.de/api/
Schulferien	sind am betreffenden Tag Schulferien in Thüringen, 1 oder 0	https://ferien-api.de/api/v1/holidays/
DS_10	diffuse Himmelstrahlung	Deutscher Wetterdienst
GS_10	Globalstrahlung joule	Deutscher Wetterdienst
SD_10	Sonnenscheindauer	Deutscher Wetterdienst
LS_10	Langwellige Strahlung	Deutscher Wetterdienst
RWS_DAU_10	Niederschlagsdauer 10 Minuten; wie lange (in Minuten) regnete es in den letzten zehn Minuten	Deutscher Wetterdienst
RWS_10	Summe der Niederschlagshöhe der vorangegeenen 10 Minuten	Deutscher Wetterdienst
RWS_IND_10	Niederschlagsindikator; gibt an, ob es in den letzten zehn Minuten geregnet hat (1) oder nicht (0)	Deutscher Wetterdienst
Verbrauch (nur Tanzende Siedlung)	Gesamtenergieverbauch des Quartiers	Erzeugungs- und Verbrauchsdaten
Mietverträge (nur Tanzende Siedlung)	Anzahl abgeschlossener Mietverträge	Chemnitzer Siedlungsgemeinschaft (CSg)

Tabelle 2: Kandidatenwerte, die bei der Dateaufbereitung den Messdaten hinzugefügt werden

4 Machbarkeitsanalyse

Um zu eruieren, ob sich aus den vorliegenden Daten überhaupt sinnvolle Vorhersagen treffen lassen, sollen die Daten in diesem Kapitel auf verschiedene Weise untersucht werden. Zunächst soll dazu untersucht werden, ob sich Muster oder sogar Periodizitäten in den Daten wiederfinden. Hierfür werden zunächst die Messwerte selbst veranschaulicht, um offenkundige Muster ausmachen zu können. Dann wird eine Autokorrelation auf die verschiedenen Messreihen angewendet werden, die die Korrelation der Messreihe mit sich selbst untersucht. Sie gibt Aufschluss darüber, wie stark sich aufeinander folgende Tage ähneln. Zuletzt werden die Messreihen bezüglich ihrer Differenzdaten, also im Grunde hinsichtlich der ersten Ableitung untersucht, wodurch der Variationstrend und insbesondere die Periodizität reduziert werden kann. (vgl. [22]) Anschaulich gesprochen ist dieser Wert also ein Indikator für starke Änderungen in den Daten, die keiner Periodizität unterliegen.

Anschließend sollen die zuvor aufgestellten Kandidatengrößen hinsichtlich der Stärke ihrer Korrelation mit der Zielgröße untersucht werden. Dazu werden die Kendall Ranks für die verschiedenen Einflussgrößen berechnet: Diese sind ein Maß für die Stärke der Korrelation zweier Größen und lassen so Rückschlüsse darauf zu, wie gut ein Kandidatenwert als Einflussgröße für das zu trainierende Modell geeignet ist. Laut [22] sind diejenigen Einflussgrößen geeignet, die einen Kendall-Wert mit einem Betrag von mehr als 0,3 aufweisen.

4.1 Alfons-Pech-Straße

In diesem Kapitel werden die Daten der Alfons-Pech-Straße auf die oben beschriebene Weise untersucht. Um sowohl einen Blick ins Detail als auch eine Übersicht über das Große Ganze zu erlangen, werden die Analysen jeweils auf einem Subdatensatz, der die ersten drei Tage der vorliegenden Messreihe umfasst, also den Zeitraum vom 5. bis zum 7. April 2020, sowie auf dem gesamten Datensatz (5. April bis 15. Mai 2020) angewendet.

Abbildung 3a zeigt die Erzeugungsdaten der PV-Anlage in der Alfons-Pech-Straße eben jene ersten drei Tage. Hier lässt sich ganz klar eine Periodizität erkennen, die so zu erwarten gewesen war: Während nachts über keine PV-Energie produziert wird, steigt die Produktionskurve am Vormittag an, bevor die ihren Peak etwa um die Mittagszeit erreicht um dann am Nachmittag wieder langsam gegen die Nulllinie zu sinken. Die Produktionskurve folgt somit grob dem Sonnenstand. Die nebenstehende Abbildung 3b zeigt die gleichen Daten, für den gesamten Zeitraum auf. Auch hier lassen sich klar die Tagespeaks erkennen, wenngleich auch Tage zu sehen sind, für die nur ein sehr geringer Peak verzeichnet wurde. Im Detail betrachtet sehen allerdings nicht alle Tage so mustergültig aus, wie die in Abbildung 3a. Andere Daten weisen wesentlich unstetigere Verläufe auf. Dies ist wohl darauf zurückzuführen, dass es an den drei hier aufgezeigten Tagen keine oder eine durchgehende Bewölkung gab. Dem entgegen könnten die unstetigeren Verläufe der anderen Tage darauf zurückzuführen sein, dass die Sonne immer wieder von Wolken verdeckt wurde.

Abbildung 3c zeigt die normalisierte Autokorrelation über die ersten drei Tage. Hier lassen sich Peaks der Kurve bei x-Werten von 1440 und 2880 erkennen, die genau das Intervall eines Tages widerspiegeln. (Die Daten liegen in einer zeitlichen Auflösung von einer Minute vor, ein Tag hat 1440 Minuten). Sowohl in dieser wie auch in der nebenstehenden Abbildung 3d lässt sich jedoch über mehrere Tage hinweg eine sinkende Korrelation beobachten. Auch dies erscheint sinnvoll, da sich über einen längeren Zeitraum hinweg verschiedene Einflussgrößen wie beispielsweise das Wetter ändern.

Die Abbildungen 3e und 3f zeigen die Differenzdaten für die beiden angesprochenen Zeiträume auf. Mit Differenzdaten wird dabei jeweils die Subtraktion eines Wertes mit seinem unmittelbaren Vorgänger bezeichnet, die auch als erste Ableitung beschrieben werden kann. Für die ersten beiden Tage lässt sich dabei eine nur sehr geringe Änderung erkennen, während am dritten Tag ein deutlich erkennbarer Sprung vorliegt. Wie die Abbildung 3f zeigt, passieren solche Sprünge auch an anderen Tagen und haben teilweise einen sehr viel höheren Ausschlag. Diese Sprünge könnten auf das Hervorkommen der Sonne hinter den Wolken oder im umgekehrten Fall das Verschwinden der Sonne hinter einer Wolkendecke zurückzuführen sein.

Zusammenfassend lässt sich also sagen, dass sowohl aus den reinen Visualisierungen der Produktionsdaten, als auch aus denen der Autokorrelationen eine klare Periodizität ablesen lässt. Die Differenzdaten hingegen deuten darauf hin, dass diese Periodizität von schwerenwiegenden Aus-

reißern geprägt ist, da die Peaks dieser Verlaufskurve nahezu an die Peaks der Produktionskurve heranreichen.

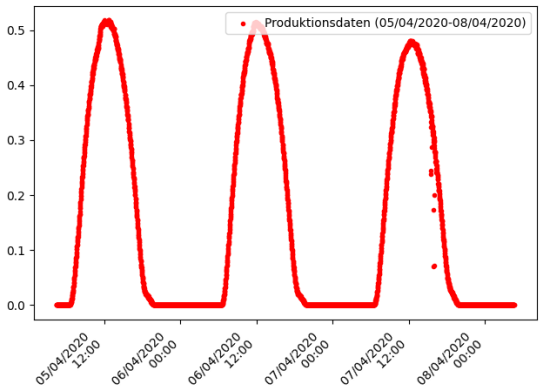
Abbildung 5 zeigt die Auswertung der Analyse bezüglich der Stärke der Korrelation zwischen der PV-Produktion und den aufgeführten Größen. Wie nicht anders zu erwarten war, bestehen die stärksten Korrelationen mit Größen, die das Verhalten der Sonne beschreiben. Diese sind: Die diffuse Himmelstrahlung (0,58), die Globalstrahlung in Joule (0,7) sowie die Sonnenscheindauer (0,57). Die Langwellige Strahlung hingegen weist absolut keine Korrelation mit der PV-Stromproduktion auf (0).

Für die Einflussgröße "Zeit" wurde angenommen, dass diese ebenfalls einen starken Einfluss auf die PV-Stromproduktion haben würde, was die Abbildung 3a bestätigt. Zunächst wurde die Zeit im Format "Sekunden, die seit Mitternacht vergangen sind" angegeben. Für dieses Format ergab sich allerdings lediglich ein Kendall-Wert von -0,1, also eine sehr schwache bis gar keine Korrelation. Da in KI-Systeme typischerweise mit Werten zwischen 0 und 1 gerechnet wird, wurde die Zeit nun auf diesen Wertebereich geclippt, was allerdings ebenfalls einen Kendall-Wert von 0,08, also eine noch schwächere Korrelation ergab. Laut [6] ist das Problem, dass nahe beieinander liegende Werte, also beispielsweise 23:59 und 00:01 in ihrer Datenrepräsentation keine Nähe zueinander aufweisen, wie Abbildung 4a veranschaulicht. Um dieses Problem zu lösen wird dort vorgeschlagen, die Zeit auf eine Kosinus-Funktion abzubilden, wie es in Abbildung 4b veranschaulicht ist. Dir Formel dazu lautet

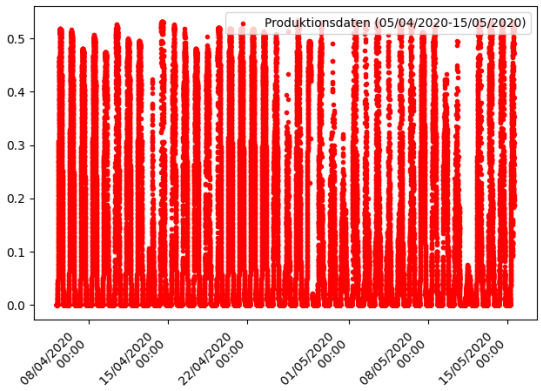
$cos(2 * \pi * s/d)$

wobei s für den Zeitstempel in Sekunden seit Mitternacht und d für die Sekunden eines vollen Tages steht. Für die so errechnete Einflussgröße ergab sich ein Kendall-Wert von -0,46, also eine nennenswerte Korrelation. Für alle anderen Größen hingegen ergeben sich lediglich sehr geringe Korrelationsstärken, womit diese als Einflussgrößen für das neuronale Netzwerk nicht in Betracht kommen.

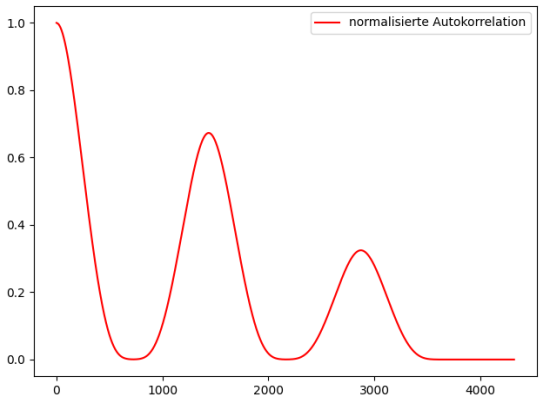
Die gleiche Methodik wurde testweise auch auf die Wochennummer angewendet, denn hier ergibt sich das gleiche Problem: Die Woche 53, also die letzte Woche im Dezember, und die Woche 1, also die erste Woche im Januar, folgen zwar zeitlich gesehen direkt aufeinander, liegen in der reinen "Indexrepäsentation" aber um 52 Schritte voneinander entfernt. Um diese Werte aneinander anzunähern, wurde die oben erwähnte Kosinus-Funktion auch auf die Wochennummer angewandt. Damit konnte allerdings keine Steigerung des Kendall Wertes erzielt werden: Für die Daten der Alfons-Pech-Straße blieb dieser mit jeweils -0,03 für die Wochennummer und den Kosinus der Wochennummer sogar gleich.



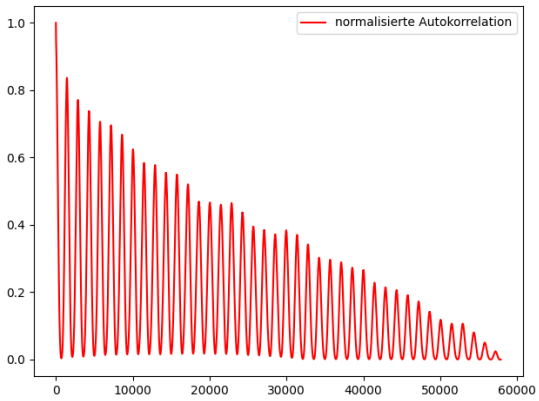
(a) PV-Produktionsdaten über die ersten drei Tage der vorliegenden Daten



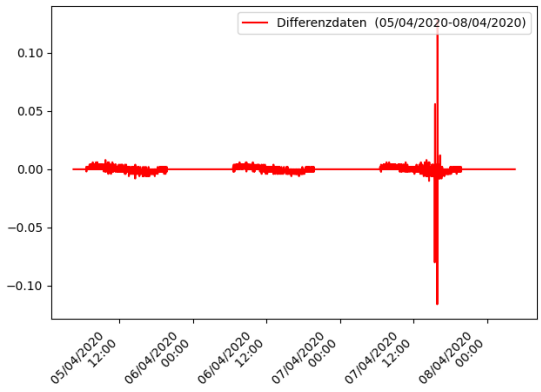
(b) PV-Produktionsdaten über alle vorliegenden Daten



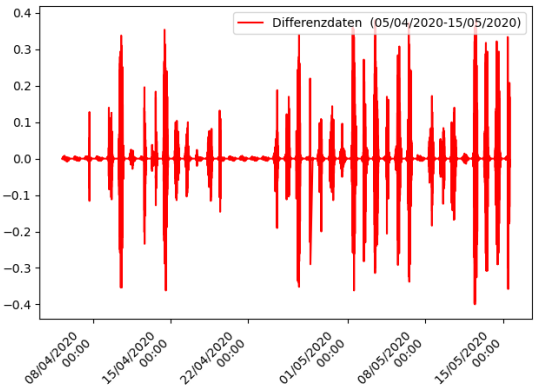
(c) Normalisierte Autokorrelation über die ersten drei Tage der vorliegenden Daten



(d) Normalisierte Autokorrelation über alle vorliegenden Daten

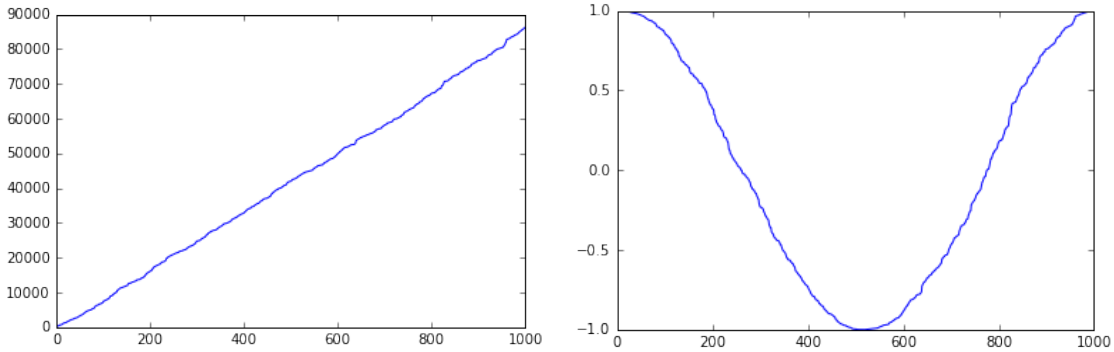


(e) Differenzdaten über die ersten drei Tage der vorliegenden Daten



(f) Differenzdaten über alle vorliegenden Daten

Abbildung 3: Machbarkeitsanalyse auf den Daten der Alfons-Pech-Straße



(a) Zeitstempel als “Sekunden seit Mitternacht”

(b) Kosinus des Zeitstempels aus Abbildung 4a

Abbildung 4: Abbildungen aus [6]

4.2 Tanzende Siedlung

Für die Tanzende Siedlung sollen für zwei Werte Vorhersagen getroffen werden: Einerseits der Gesamtenergieverbrauch des Hauses und andererseits die PV-Einspeisung in das Quartier. Abbildung 6a zeigt den Verlauf der PV-Einspeisung über den gesamten Messzeitraum, also des gesamten Jahres 2021 auf: Während im Winter relativ wenig Solarstrom produziert wird, nimmt dieser Wert über den Frühling bis in den Sommer hinein immer weiter zu, bis der Höchstwert etwa Anfang Juli erreicht wird. Über den Herbst und in den Winter hinein sinkt die Erzeugungskurve wieder auf den ursprünglichen Wert ab. Betrachtet man kleinere Ausschnitte dieser Kurve, die einen oder wenige Tage zeigen, so lässt sich beobachten, dass die Stromproduktion in den Nachtstunden bei null liegt, morgens beginnt anzusteigen, den Peak etwa zur Mittagszeit erreicht und schließlich zum Abend hin wieder abfällt, wobei an unterschiedlichen Tagen unterschiedlich hohe Peaks zu beobachten sind. Abbildung 6c zeigt diesen Verlauf. Es lassen sich also zweierlei Trends erkennen: Einerseits der untertägige Trend mit dem Peak zur Mittagszeit, sowie der jährliche Trend mit dem Peak im Sommer.

Abbildung 6b zeigt den Gesamtverbrauch des Quartiers über das Jahr 2021 hinweg. Negative Werte bedeuten dabei, dass die PV-Anlage mehr Strom produziert hat, als im Quartier benötigt wurde und somit Strom ins öffentliche Netz eingespeist werden konnte. Dies ist verstärkt in den Sommermonaten zu beobachten und spiegelt die Werte der nebenstehenden Abbildung 6a wieder: Etwa Anfang Juni, also auf dem Höhepunkt der Produktionskurve, erreicht die Verbrauchskurve ihren Tiefpunkt. Abbildung 6d zeigt für einen ausgewählten Zeitraum von vier Tagen im Sommer (30.06-04.07) einen Ausschnitt dieser Kurve. Hieran kann man deutlich erkennen, dass der Gesamtverbrauch tagsüber ins negative sinkt, während über alle Tage hinweg größere und kleinere Lastspitzen zu erkennen sind. Besonders auffällig sind die beiden Lastspitzen am ersten Tag, die jeweils auf 20 und weit über 30 kWh anwachsen. Diese Spitzen lassen sich auf Ladevorgänge von Elektro-Fahrzeugen zurückführen, im ersten Falle an einer normalen Ladesäule und im Falle der Laststpitze mit weit über 30 kWh an einer Schnellladesäule.

Abbildung 7 zeigt die Differenzdaten des Gesamtverbrauchs (links) sowie der PV-Einspeisung (rechts). Differenzdaten meinen dabei wiederum die Subtraktion eines Wertes mit seinem unmittelbaren Vorgänger, oder anders ausgedrückt: Die Steigung beziehungsweise erste Ableitung der Kurve. Bei der Kurve des Gesamtverbrauchs sind dabei immer wieder starke Sprünge zu erkennen, die teilweise bis zu 40 beziehungsweise -40 kWh erreichen. Die höchsten dieser Sprünge rühren wohl vom Starten und Beenden eines Ladevorgangs für E-Fahrzeuge her. Diese Unstetigkeit ist jedoch über das Jahr hinweg relativ ausgeglichen verteilt. Dementgegen steht die Abbildung 7b, die Differenzdatenkurve für die PV-Einspeisung: Hier lassen sich ebenfalls starke Sprünge in der Einspeisekurve feststellen, diese reichen aber nur bis auf minus beziehungsweise plus 7,5 kWh. Allerdings ist klar ein Trend erkennbar, dass die Sprünge zum Sommer hin immer größer werden und danach durch den Herbst und in den Winter hinein wieder zurückgehen. Dies geht darauf zurück, dass im Sommer tagsüber einfach mehr PV-Strom produziert wird und Produktionskurven mit einem höheren Maximum auch eine größere Steigung aufweisen. Insgesamt sind die Differenzdaten für die Tanzende Siedlung also sehr viel sprunghafter als die der Alfons-Pech-Straße. Dieser Ver-

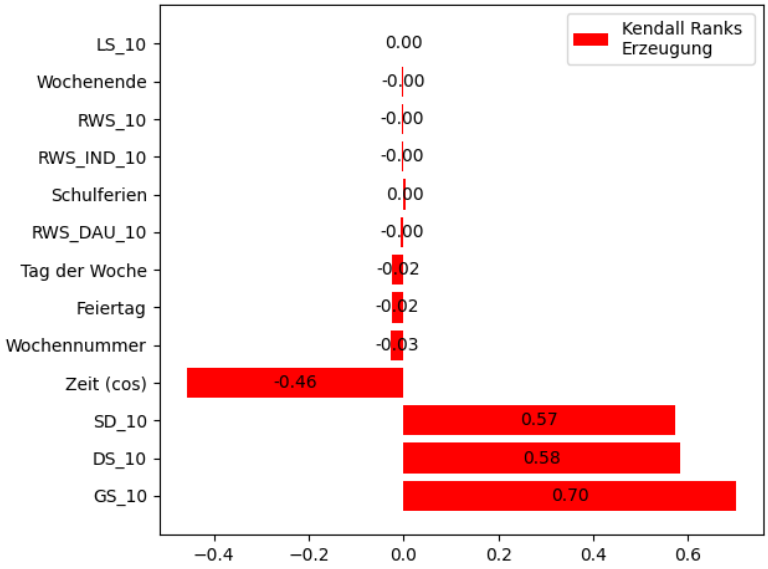


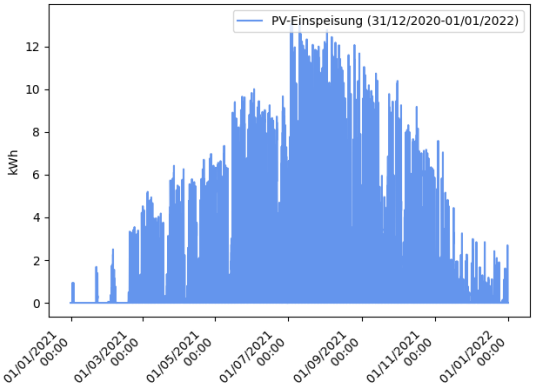
Abbildung 5: Kendall Koeffizienten für die PV-Produktion der Alfons-Pech-Straße. Die angezeigten Werte sind auf zwei Nachkommastellen gerundet. Für eine genaue Beschreibung der einzelnen Werte siehe Tabelle 2

gleich ist allerdings trügerisch, da die Daten der Alfons-Pech-Straße zeitlich gesehen 15-mal höher aufgelöst sind als die der tanzenden Siedlung.

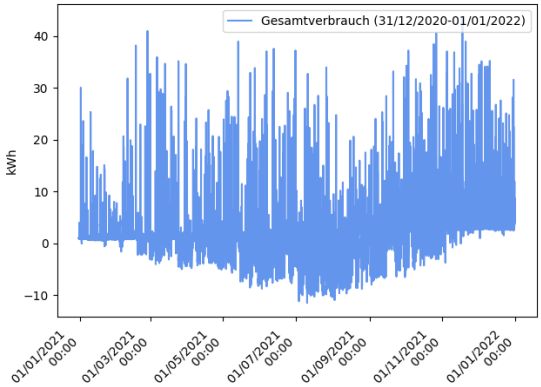
Abbildung 8a zeigt eine (normalisierte) Autokorrelation der PV-Einspeisung über den gesamten Zeitraum, wobei nur die ersten 720 Werte im Diagramm zu sehen sind. Da die Werte in einer Zeitauflösung von fünfzehn Minuten vorliegen, besteht ein Tag folgerichtig aus 96 Messungen. Die Autokorrelationsfunktion zeigt deutliche Peaks bei Vielfachen von 96, was für eine große Ähnlichkeit zwischen den Tagen spricht. Selbst bei einer Verschiebung der Kurve gegen sich selbst um sechs Tage, also bei einem x-Wert von 576, ergibt sich noch ein Autokorrelationswert von fast 0,8.

Die Autokorrelationskurve für den Gesamtenergieverbrauch (Abbildung 8b) zeigt einen ähnlichen Verlauf: Auch hier lassen sich lokale Höhepunkte der Kurve bei den Vielfachen von 96 erkennen. Allerdings sind diese wesentlich schwächer ausgeprägt als die der Autokorrelationskurve der PV-Erzeugung mit Werten um etwa 0,5. Vor allem an den Stellen der lokalen Minima der Kurve, bei einer Verschiebung der Kurve gegen sich selbst um zwölf Stunden lassen sich große Unregelmäßigkeiten erkennen. Insgesamt lässt sich aus dieser Grafik jedoch klar eine Periodizität ablesen.

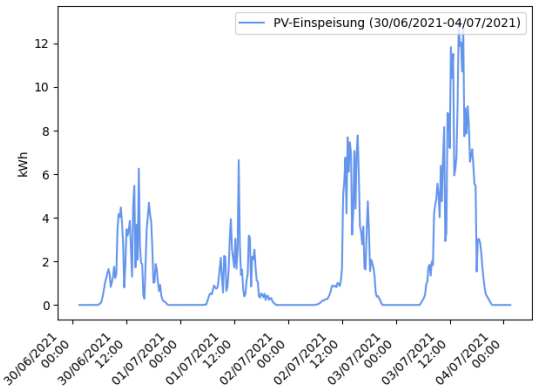
Bei den Kendall-Ranks der PV-Einspeisung (Abbildung 8a) ergaben sich gleich sieben Größen mit einem Wert größer gleich 0,3. Die stärkste Korrelation wies dabei der Bezug der PV-Anlage mit 0,48 auf. Die erscheint einleuchtend, da die PV-Anlage ihren Eigenbedarf nur dann extern decken muss, wenn intern nicht genug Energie zu diesem Zweck produziert werden kann. Die zweitstärkste Korrelation nach Kendall-Ranks bieten der Gesamtverbrauch und der Netzbezug mit jeweils 0,43. Der Grund für die Korrelation mit dem Gesamtverbrauch könnte der Tag-Nacht-Rythmus sein: Während am Tag einerseits viel Energie produziert wird, weil die Sonne scheint, wird auch viel Energie benötigt, da die Mieter wach sind und Strom verbrauchen. Nachts, wenn die Mieter schlafen und keine Sonne am Himmel steht wird weder viel Strom produziert noch verbraucht. Außerdem wird der Gesamtverbrauch derart berechnet, dass Einspeisung ins öffentliche Netz als negativer Verbrauch gewertet wird, was diese Korrelation ebenfalls erklärt. Die starke Korrelation mit dem Netzbezug lässt sich dadurch erklären, dass Energie nur aus dem öffentlichen Stromnetz bezogen werden muss, wenn dieser Bedarf nicht durch die eigene PV-Anlage gedeckt werden kann. Mit Kendall-Werten von -0,38 und -0,34 weisen die Globalstrahlung Joule (GS_10) sowie die diffuse Himmelstrahlung (DS_10) ebenfalls starke Korrelationen mit der PV-Einspeisung auf. Auch dies



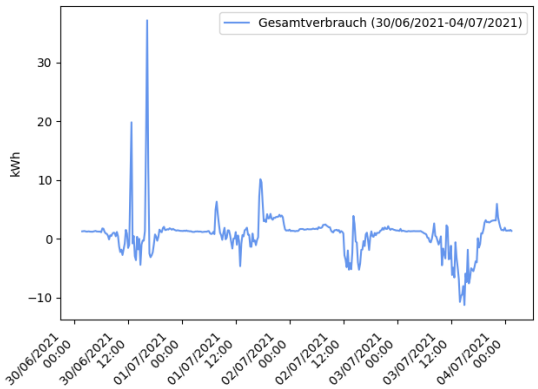
(a) Einspeisung der PV-Anlage ins Quartier über den gesamten Zeitraum



(b) Gesamtverbrauch der tanzenden Siedlung über den gesamten Zeitraum



(c) Einspeisung der PV-Anlage ins Quartier über einen Zeitraum von vier Tagen im Sommer

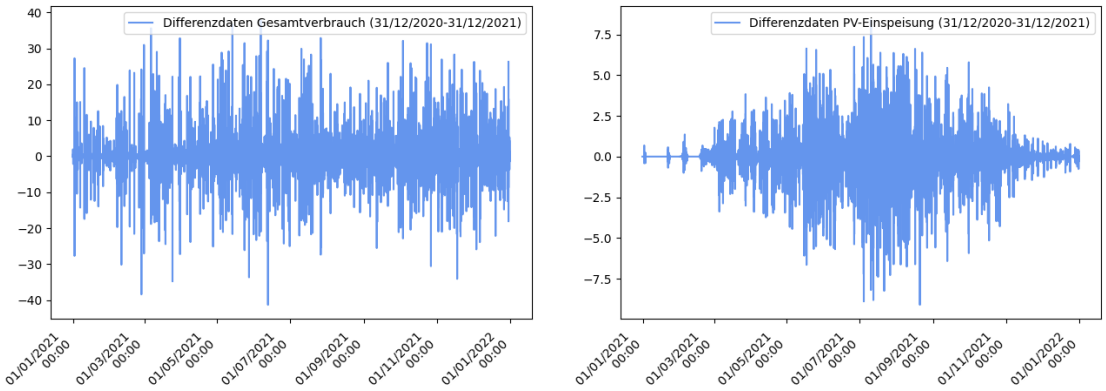


(d) Verbrauch über vier Tage hinweg

Abbildung 6: Verlaufskurven über Messwerte der Tanzenden Siedlung, jeweils über den gesamten Zeitraum (oben) und vier ausgewählte Tage im Sommer (unten)

erscheint plausibel, beschreiben diese beiden Faktoren doch die auf der Erde auftreffenden Sonnenstrahlen und somit die Möglichkeit für die PV-Anlage, diese in Strom umzuwandeln. Auch die Netzeinspeisung stellt mit einem Kendall-Wert von -0,31 eine verwendbaren Einflussgröße dar. Diese Korrelation lässt sich folgendermaßen erklären: Wenn viel PV-Strom produziert wird, kann dieser unter Umständen nicht im Quartier genutzt werden und fließt somit ins öffentliche Stromnetz. Die Kosinus-Funktion angewendet auf auf die Zeit stellt mit einem Kendall-Wert von -0,3 eine gerade noch verwendbaren Einflussgröße dar. Wie bereits ausgeführt verläuft die Produktionskurve tagesperiodisch mit einem Peak in etwa um die Mittagszeit. Die Sonnenscheindauer, die Wochennummer sowie die Anzahl der Mietverträge weisen jeweils nur eine schwache Korrelation mit der Zielgröße “PV-Einspeisung” auf. Für die Wochennummer war dies recht überraschend, zeigt doch Abbildung 6a einen deutlich erkennbaren Trend über das Jahr hinweg auf. Wie ebenfalls zuvor schon beschrieben wurde, konnte auch die Anwendung einer Kosinus-Funktion auf die Wochennummer diese Korrelation nicht verstärken. Dies könnte darauf zurückzuführen sein, dass im Messzeitraum immer mehr Mieter eingezogen sind und somit das Ergebnis verfälscht wird. Für die restlichen Kandidatenwerte ergaben sich sehr niedrige bis gar keine Korrelationen, sodass sie im Weiteren keine Beachtung mehr finden werden.

Abbildung 8d zeigt die Berechnungen der Kendall-Werte für den Gesamtverbrauch gegen alle zuvor erhobenen Einflussgrößen. Von den 18 berechneten Einflussgrößen weisen insgesamt fünf einen Kendall-Wert von über 0,3 auf, der laut [22] die Schwelle für eine relevante Korrelation bildet. Die stärkste Korrelation mit dem Gesamtverbrauch des Quartiers stellt der Wert des Netzbezugs



(a) Differenzdaten des Gesamtverbrauchs über den gesamten Zeitraum

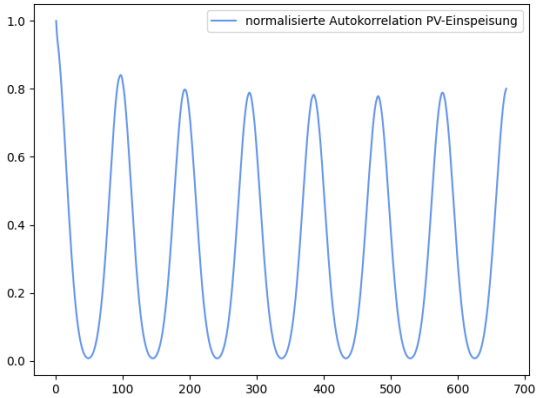
(b) Differenzdaten der PV-Einspeisung über den gesamten Zeitraum

Abbildung 7: Differenzdaten der Tanzenden Siedlung

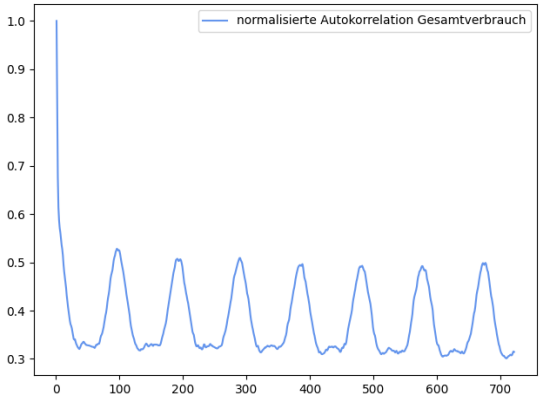
des Quartiers mit einem Wert von 0,96 dar. Dies deutet darauf hin, dass die PV-Anlage nur einen Teil des Energiebedarfs decken kann, während gleichzeitig oft Strom aus dem öffentlichen Netz bezogen werden muss. Die zweitgrößte Korrelation weist die Netzeinspeisung mit einem Wert von -0,47 auf. Das liegt darin begründet, dass tagsüber mehr Energie verbraucht wird, da die Menschen wach sind, gleichzeitig scheint die Sonne tagsüber und die PV-Anlage liefert Strom. Auch die PV-Einspeisung weist mit einem Kendall-Wert von 0,43 eine nennenswerte Korrelation auf. Dies kann damit erklärt werden, dass Energie nur dann ins öffentliche Stromnetz eingespeist werden kann, wenn dieser Strom aktuell nicht für den Eigenbedarf des Quartiers benötigt wird. Die Einflussgrößen, die die auf der Erde auftreffende Sonne charakterisieren, kommen mit Kendall-Werten von -0,32 (Globalstrahlung Joule), -0,28 (diffuse Himmelstarhlung) und -0,27 (Sonnenscheindauer) daher. Das könnte darauf zurückzuführen sein, dass der Verbrauch nicht den reinen Verbrauch der Mieter meint, sondern aufgrund seiner Berechnung auch Überschuss an Strom im Quartier (und daraus resultierende Einspeisung ins Revier) als negativer Verbrauch zu Buche schlägt. Scheint die Sonne also sehr intensiv (was hoher Globalstrahlung Joule, Sonnenscheindauer und diffuser Himmelstrahlung entspricht) wird viel Strom produziert, der wiederum nicht in Gänze benötigt wird und somit ins öffentliche Netz fließt. Auch wenn die Sonnenscheindauer mit -0,27 und die diffuse Himmelstrahlung mit -0,28 unter der Marke von 0,3 liegen, können sie aufgrund des nur geringen Abstandes zu dieser Marke trotzdem testweise in ein Modell aufgenommen werden. Gleiches gilt für die Wochennummer, die mit einem Kendall-Wert von 0,29 sogar noch knapper unter dieser Grenze liegt. Diese Kendall-Werte scheinen also plausibel erklärbar zu sein.

Überraschend war hingegen, dass Einflussgrößen, die als Indikatoren für die An- und Abwesenheit der Mieter dienen sollten, etwa Wochentag/Wochenende, Schulferien und Feiertage, nur sehr geringen bis keinen Einfluss auf den Gesamtenergieverbrauch zu haben scheinen. Allerdings handelt es sich bei den Daten um Werte für das Jahr 2020, in dem aufgrund der Corona-Pandemie viel Arbeit und Unterricht aus dem Homeoffice verrichtet wurde. Auch der Kosinus der Zeit erreicht nur einen Kendall-Wert von 0,13. Hier war erwartet worden, dass es einen Tag-Nacht-Zyklus gibt, was offenbar nicht zutreffend ist. Überraschend war zudem, dass die Anzahl der Mieter nur einen geringen Einfluss auf den Stromverbrauch hat (Kendall-Rank von 0,2)

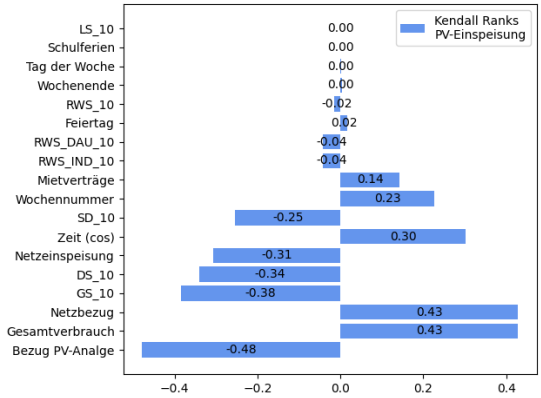
Insgesamt lässt sich also festhalten, dass die Daten der PV-Einspeisung einer starken Periodizität unterliegen und es eine ganze Reihe an Einflussgrößen gibt, die nachvollziehbar einen Einfluss auf diese Zielgröße haben. Auch beim Gesamtverbrauch des Quartiers lässt sich mittels der Autokorrelation eine Tagesperiodizität und mittels des reinen Plottings der Daten ein jährlicher Trend erkennen. Hier ergeben sich eine Reihe von Einflussgrößen, die laut ihrer Kendall-Ranks gut als Einflussgrößen für eine Vorhersage geeignet sind. Besonders hervorzuheben ist hier der Netzbezug mit einem fast maximalen Wert. Die Differenzdaten hingegen weisen auf starke Sprünge innerhalb der Daten hin, die wiederum keiner Periodizität unterliegen. Trotzdem erscheinen beide Zielgrößen geeignet zu sein, Vorhersagen über ihren zukünftigen Verlauf treffen zu können.



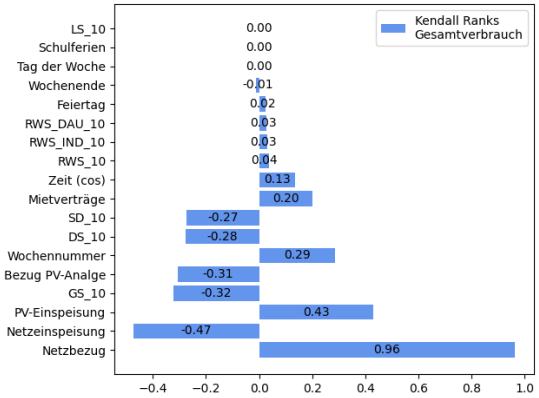
(a) Autokorrelation der PV-Einspeisung, gerechnet über den gesamten Messzeitraum, geplottet sind die ersten 7,5 Tage



(b) Autokorrelation des Gesamtverbrauchs über den gesamten Zeitraum



(c) Kendall Koeffizienten für die PV-Einspeisung. Werte auf zwei Nachkommastellen gerundet. Für eine genaue Beschreibung der einzelnen Werte siehe Tabelle 2



(d) Kendall Koeffizienten für den Gesamtverbrauch. Werte auf zwei Nachkommastellen gerundet. Für eine genaue Beschreibung der einzelnen Werte siehe Tabelle 2

Abbildung 8: Machbarkeitsanalyse auf den Daten der Tanzenden Siedlung

5 Aufbau eines LSTM-Modells

Um aus den zuvor berechneten Daten Prognosen zu erstellen, wurde ein neuronales Netzwerk, genauer ein LSTM (Long short-term memory), trainiert. Dazu wurde die entsprechende Implementierung aus der Keras Python-Bibliothek verwendet. Für die drei Zielgrößen der beiden Demonstratoren sollten je drei verschiedene Modelle mit jeweils unterschiedlicher zeitlicher Auflösung trainiert werden, die kurz- mittel und langfristige Vorhersagen treffen.

5.1 Wahl der Parameter

Da keine Regeln ausgemacht werden konnten, auf Basis derer die Parameter eines Modells festgelegt werden konnten, wurden diese teils aus dem wissenschaftlichen Artikel von Wang et al. [22] übernommen, teils experimentell ermittelt. Für die experimentell ermittelten Parameter wurden durch Recherchen eine Reihe von typischen Werten zusammengetragen und schließlich im Modell ausprobiert um daraus die beste Wahl im Sinne der RMSE treffen zu können. Aufgrund der stochastischen Berechnung des LSTMS variieren die Ergebnisse, sodass hier jeweils drei Testdurchläufe ausgeführt wurden, die im Folgenden angegebenen RMSEs sind jeweils als gerundete Durchschnittswerte der Ergebnisse der verschiedenen Testläufe zu verstehen.

Wie auch im wissenschaftlichen Artikel von Wang et al. wird das Netzwerk aus zwei Hidden

Layers aufgebaut. Die Anzahl der LSTM-Units beschreibt die "Anzahl der hidden units, die der Länge des Zustandsvektors entsprechen". [24] Die Wahl dieser Größe hat damit Einfluss darauf, "die strukturellen und semantischen Merkmale der Eingabesequenz zu erfassen." [24] Daher wird in [24] empfohlen, dass diese Zahl nicht zu klein gewählt werden sollte. Bei Wang et al. werden 100 LSTM-Units verwendet. An anderer Stelle wurden hierfür allerdings Werte von 256, 512 und 1024 vorgeschlagen. Daher wurde jedes Modell mit 100, 256, 512 und 1024 LSTM-Units berechnet.

Die *Batch Size* beschreibt die Anzahl der Proben, die bearbeitet werden, bevor die Parameter des Modells upgedatet werden. Diese Einstellung kann sich einerseits auf die Dauer auswirken, die das Modell benötigt, trainiert zu werden. Andererseits kann eine hohe Batch Size aber auch dazu führen, dass lokale Minima schwerer gefunden werden. Der allgemeine Konsens besagt, dass 32, 64 und 128 geeignete Werte für die Batch Size sind. (vgl. z.B. [3]).

Die Eingabematrix eines LSTM-Modells enthält in jeder Spalte die aktuellen Messwerte, sowie weitere Spalten, in denen die Eingabewerte jeweils um 1 bis m Schritte in die Vergangenheit und um 1 bis n Schritte in die Zukunft verschoben werden. Im Lernprozess wird das Modell also trainiert aus m Schritten in die Vergangenheit n Zukunftsschritte vorherzusagen. Die Angabe, wie groß n dabei ist wurde semantisch festgelegt und beträgt jeweils eine Stunde, einen Tag und eine Woche (mehr dazu später). Um zu eruieren, wie das m idealerweise zu wählen ist, wurde jeweils n, 2n und 3n Schritte in die Vergangenheit als Parameter eingestellt und entsprechend ausprobiert.

Über den Parameter *Dropout* kann angegeben werden, wie viel Prozent der vorhandenen Neuronen zufällig bei einem Trainingsdurchlauf ausgelassen werden, wodurch Overfitting (Auswendiglernen von Daten anstelle von Erkennen und Lernen von Mustern) vorgebeugt werden soll. Eine typische Größe, die dafür empfohlen wird sind 20 % Dropout (vgl. z.B. [3]). Vor allem bei den Modellen für die Alfons-Pech-Straße liegen allerdings recht wenig Eingabedaten vor. Einerseits basieren diese Modelle auf lediglich vier Einflussgrößen, andererseits sind vor allem für Modelle, in denen Daten gruppiert werden, um etwa Stundenwerte zu erhalten auch wenig Messwerte vorhanden. Daher wurde befürchtet, dass die Verwendung eines Dropouts die Vorhersagekraft des Modells wesentlich verschlechtern könnte. Um die zu testen, wurden ebenfalls Modelle mit null beziehungsweise zehn Prozent Dropout getestet.

Jedes der kombinierten Modelle wurde mit jeweils 100 Epochen trainiert, um gleiche Voraussetzungen für den Test zu schaffen. Für die Modelle der Alfons-Pech-Straße war dieser Wert ausreichend, allerdings stagnierten die Test-Loss (vgl. Kapitel 5.2) Kurven für die Tanzende Siedlung nach 100 Durchläufen noch nicht. Daher wurde dafür ein anderes Vorgehen gewählt: Das verwendete Framework Keras bietet die Möglichkeit, das Training vor Erreichen der angegebenen Anzahl an Epochen abzubrechen, sobald sich eine angegebene Metrik nicht weiter verbessert. Zudem kann dabei angegeben werden, dass das Training nicht unmittelbar gestoppt werden soll, wenn dieser Fall eintritt, sondern das noch einige Epochen weiter trainiert werden soll. Daher wurden die Modelle für die Tanzende Siedlung jeweils mit 1000 Epochen trainiert. Das Training wurde vorzeitig abgebrochen, wenn sich der Test-loss über 15 Epochen nicht weiter verbesserte.

Mittls der *Loss Function* werden beim Training des Modells die Parameter optimiert. Für verschiedene Zwecke eignen sich dabei verschiedene Funktionen, in [9] sind insgesamt sieben Funktionen angegeben, die sich als Loss Function für Regressionsprobleme eignen: Mean Square Error, Mean Absolute Error, Cosine Similarity, Huber Loss, Mean Absolute Percentage Error, Mean Squared Logarithmic Error und Log Cosh. Für jede dieser Funktionen wurde ein eigenes Modell trainiert, zum Testen der anderen Parameter wurde stets der Mean Average Error als Loss Function verwendet.

Eine weitere Frage war, welcher Prozentsatz der Daten als Test- beziehungsweise Trainingsdaten verwendet werden. Viele im Internet verfügbare Tutorials arbeiten mit 67 % Traininsgdaten und entsprechenden 33 % Testdaten. In anderen Fällen wurden jedoch auch 80 % aller Daten als Trainingsdaten verwendet (z.B. [15]). Da vor allem für die Alfons-Pech-Straß ohnehin recht wenig Daten vorliegen, wurden daher alle Modelle mit 80 % Traininsgdatenanteil trainiert.

Wie die Experimente zeigten, ergaben sich beim Testen der unterschiedlichen Parameter oft sehr ähnliche Modelle im Sinne der resultierenden RMSE. Um Effekten wie Auswendiglernen der Daten durch das Modell anstelle von Lernen von Mustern (Overfitting) vorzubeugen, wurden im Zweifelsfall stets die Faktoren gewählt, die ein Modell einfacher oder stabiler gestalten.

5.2 Vergleich der Modelle

Die Eignung dieser Parameter für das Modell wurde anhand des RMSE, des Root-Mean-Square Error berechnet. Dieser beschreibt die Standardabweichung der Differenz zwischen Originalem und vorhergesagtem Datum. Vorhersagen meint in diesem konkreten Sinne lediglich die Vorhersagen der Zielgröße, nicht jedoch die Vorhersagen für die Einflussgrößen.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}}$$

Dabei ist P_i die Vorhersage der i-ten Beobachtung, die in dieser Formel wiederum durch O_i verkörpert wird. n bezeichnet die Summe aller Beobachtungen.

Zur Evaluierung des Trainings eines Modells stellt Keras den Loss, sowie den Val Loss zur Verfügung. Der Loss beschreibt dabei das Ergebnis der Loss Function auf den Traininsgdaten nach dem Training, während der Val Loss die selben Daten angewandt auf die Testdaten beschreibt. [10] Wird dies als Kurve geplottet, lassen sich daraus Schlüsse über die Performance des Modells ableiten: Idealerweise nähern sich beide Kurven gleichermaßen immer weiter der Nulllinie an, bis dieses negative Wachstum schließlich stagniert. Allerdings kann es ebenso passieren, dass sich beide Kurven zunächst der Nulllinie annähern, bis schließlich die Val Loss Kurve wieder ansteigt. Das ist ein Zeichen dafür, dass Overfitting vorliegt: Das Modell lernt dann nicht mehr Muster in den Daten zu erkennen, sondern lernt lediglich die Traininsgdaten auswendig. Daher schrumpft der Loss-Wert immer weiter, während das Modell immer schlechter für die Testdaten (und somit auch für reale Daten) geeignet ist. [2]

5.3 Modelle für die Alfons-Pech-Straße

Für die Alfons-Pech-Straße liegen Daten für etwa 1,5 Monate vor, die in einer zeitlichen Auflösung von einer Minute vorliegen. Zielgröße der Vorhersage ist die Stromproduktion durch die PV-Anlage. Für das Modell für die kurzfristigen Vorhersagen schien es angemessen, die Stromproduktionsdaten der nächsten Stunde minutengenau vorherzusagen. Für die Modelle für mittel- und langfristige Vorhersagen wurden jeweils die nächsten 24 Stunden in einer 1-Stunden-Auflösung sowie die nächste Woche in einer 1-Tages-Auflösung vorhergesagt. Da es sich bei diesen Daten um Leistungswerte (unbekannter Einheit) handelt, müssen für die Weiterverarbeitung der Daten die jeweiligen Durchschnitte der einzelnen Einflussgrößen berechnet werden.

Als Einflussgrößen wurden der Messwert der PV-Stromproduktion, der Kosinus der Zeit, sowie die drei die Sonne betreffenden Wettergrößen diffuse Himmelstrahlung, Globalstrahlung und Sonnenscheindauer verwendet. Da die Summe aller Kosinuswerte über den gesamten Tag hinweg jedoch immer 0 ist, fiel diese Größe bei dem 7-Tage-Modell, in dem ganze Tage aufsummiert werden weg. Dafür wurden hier drei andere Einflussgrößen relevant: die Niederschlagsdauer, die Summe der Niederschlagshöhe sowie der Niederschlagsindikator. Eine Auflistung dieser Größen zeigt Tabelle 5.

5.3.1 60 Minuten Vorhersage

Wie Abbildung 9a zeigt, haben die meisten Parameter kaum einen Einfluss auf die durchschnittliche RMSE des Modells. Die einzige Ausnahme hierbei bietet die verwendete Loss Function: Werden hier Cosine Similarity oder Mean Absolute Percentage Error verwendet, so ergibt sich eine RMSE von 0,25 beziehungsweise 0,19. Zwischen den anderen untersuchten Funktionen lässt sich bezüglich der RMSE hingegen kaum ein Unterschied erkennen: Die Funktion mit der niedrigsten RMSE ist jedoch der Mean Absolute Error. Werden hingegen die Vorhersagen geplottet und mit den "wahren"Daten verglichen, so fällt bei den Funktionen Mean Squared Error, Mean Squared Logarithmic Error und LogCosh auf, dass Modelle mit dieser Loss Function auch negative Werte vorhersagen (siehe Abbildung 9b). Da die Stromproduktion jedoch nicht negativ sein kann, sind diese Funktionen für die Vorhersage nicht geeignet. Damit erscheint der Mean Average Error als am besten für dieses Modell geeignet.

Überraschenderweise steigt die RMSE, je mehr LSTM-Units im Modell verwendet werden. Allerdings legen die zugehörigen Verlaufskurven des Loss- und Val Loss nahe, dass dies auf Overfitting

1-Stunde	**24-Stunden**	**7-Tage**
PV-Stromproduktion	PV-Stromproduktion	PV-Stromproduktion
diffuse Himmelstrahlung	diffuse Himmelstrahlung	diffuse Himmelstrahlung
Globalstrahlung	Globalstrahlung	Globalstrahlung
Sonnenscheindauer	Sonnenscheindauer	Sonnenscheindauer
Kosinus der Zeit	Kosinus der Zeit	Summe der Niederschlagshöhe
		Niederschlagsindikator
		Niederschlagsdauer

Tabelle 3: Einflussgrößen der unterschiedlichen Modelle für die Alfons-Pech-Straße

zurückzuführen ist. Daher erscheint es angemessen, lediglich 100 LSTM-Units zu verwenden, wie dies auch von Wang et al. [22] durchgeführt wurde. Auch beim Dropout war befürchtet worden, dass ein hoher Wert aufgrund der spärlichen Datenlage zu einem schlechteren Modell führt. Allerdings sind auch hierbei kaum Unterschiede zu erkennen. Um das endgültige Modell möglichst stabil zu gestalten, soll hierbei der übliche Wert von 20 % als Dropout angesetzt werden.

Ebenfalls überraschend war, dass das Modell, welches seine Vorhersagen auf Basis der Daten der letzten 60 Minuten berechnet besser abschnitt, als diejenigen, die 120 oder 180 Minuten als Eingabedaten erhielten. Hier war erwartet worden, dass auf breiterer Datenbasis, also in diesem Falle mehr Informationen über die vergangenen Stunden, fundiertere Vorhersagen getroffen werden können. Daher soll das endgültige Modell 60 Minuten Vergangenheitsschritte als Eingabe erhalten.

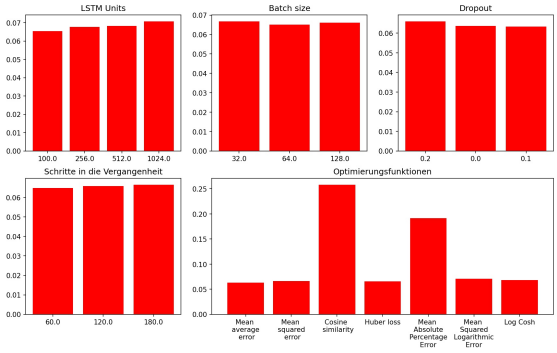
5.3.2 24 Stunden Vorhersage

Für das 24 Stunden Modell für die Alfons-Pech-Straße macht die Verwendung unterschiedlich vieler LSTM Units kaum einen Unterschied auf die resultierende RMSE. Der Unterschied zwischen der RMSE des Besten und des schlechtesten Modells betrug dabei lediglich etwa 0,0013. Logischerweise wirkt sich die Verwendung von mehr Units negativ auf die Traininsgdauer aus. In diesem Beispiel benötigte das 100-LSTM-Unit Modell nur ein Drittel der Trainingszeit, die das 1024 LSTM Units Modell benötigte. Um das aktuelle Modell möglichst einfach zu halten und Effekte wie das Ausweniglernen der Daten durch das Modell zu verhindern, sollen daher 100 LSTM-Units verwendet werden.

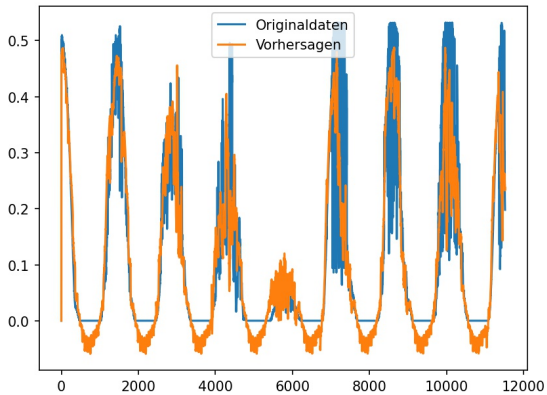
Auch beim Vergleich verschiedener Batch Sizes konnte kaum ein Unterschied bezüglich der resultierenden RMSE festgestellt werden. Die Verwendung eines Dropouts hingegen verbessert das Modell im Sinne der RMSE: Mit 0 % Dropout liegt die RMSE bei 0,079, mit 20 % hingegen bei 0,076. Bezüglich der Anzahl an Daten aus der Vergangenheit, die dem Modell zum Training zur Verfügung gestellt werden ergab sich, dass ein Modell, dass die nächsten 24 Stunden vorhersagen soll idealerweise die letzten 24 Stunden als Eingabedaten erhält. Bezüglich der besten Loss Function eignen sich gleich drei Funktionen nahezu gleich gut: Der Mean Squared Error, der Huber Loss und der Log Cosh. Allerdings besteht bei allen Funktionen die Schwierigkeit, die Nullproduktion in der Nacht richtig vorherzusagen zu können: Die Vorhersagen fluktuieren um den Nullpunkt herum, sodass auch negative Produktionsdaten vorhergesagt werden. Am geringsten ausgeprägt ist dieses Phänomen jedoch bei den Modellen, die mithilfe des Huber Loss trainiert wurden. Daher soll dieser zur Vorhersage verwendet werden.

5.3.3 7 Tage Vorhersage

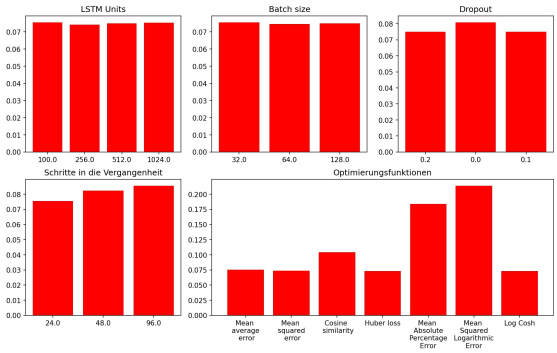
Für das Modell, dass die PV-Stromproduktion der nächsten sieben Tage vorhersagen soll zeigen sich vergleichsweise große Unterschiede zwischen den Modellen.



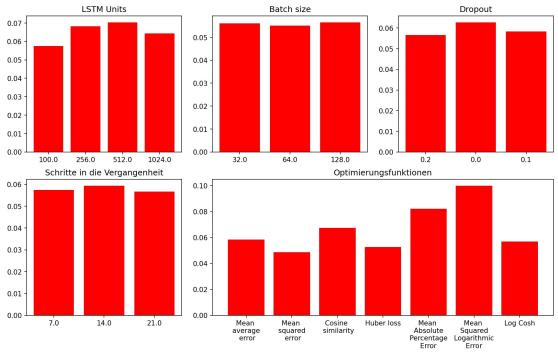
(a) 60-Minuten Modell



(b) 60 Minuten Vorhersagen auf Basis des Mean Squared Logarithmic error



(c) 24 Stunden Modell



(d) 7 Tage Modell

Abbildung 9: Visualisierung der durchschnittlichen RMSEs für die unterschiedlichen Modelle zur Vorhersage des PV-Stromproduktion in der Alfons-Pech-Straße

Die Auswertung der Tests bezüglich der Anzahl der LSTM-Units ergab, dass die Verwendung von 100 LSTM-Units am besten geeignet ist. Auch die Verlaufskurven der vier Modelle für jeweils 100, 265, 512 und 1024 verwendete LSTM-Units deuten an, dass eine Erhöhung der LSTM Units den Effekt des Overfittings begünstigt. Wiederum verbessert die Verwendung eines Dropouts von 10 % oder 20 % die RMSE des Modells leicht, wobei zwischen diesen beiden Größen kaum ein Unterschied zu erkennen ist. Daher sollen im resultierenden Modell 20 % Dropout verwendet werden. Die bestgeeignete Loss Function ist der Mean Squared Error.

5.4 Modelle für die Tanzende Siedlung

Für die Tanzende Siedlung wurden für die zwei Zielgrößen Verbrauch und PV-Einspeisung jeweils drei Modelle unterschiedlicher Zeitauflösung erstellt. Wiederum wurden dabei die Werte der nächste Stunden, der nächsten 24 Stunden und der nächsten 7 Tage vorhergesagt. Im Gegensatz zu den Daten der Alfons-Pech-Straße wurden diese lediglich aufsummiert und auf den Summen die Modelle berechnet, die Durchschnittsberechnung erübrigte sich hierbei.

In allen Modellen, die den Verbrauch vorhersagten wurden die Einflussgrößen Netzbezug, Netzeinspeisung, Bezug PV-Anlage, PV-Einspeisung sowie die Anzahl der in der Siedlung wohnenden Mieter verwendet. Für die mittel- und langfristigen Modelle wurden jeweils zusätzlich die Wettergrößen diffuse Strahlung, Globalstrahlung und Sonennscheindauer als Einflussgrößen verwendet. Das 1-Stunde Modell basiert zusätzlich auf der Wochennummer.

Für alle Modelle, die die Einspeisung der tanzenden Siedlung ins öffentliche Netz vorhersagen sollen, waren die Einflussgrößen Netzbezug, Netzeinspeisung, PV-Einspeisung, Globalstrahlung, Sonnenscheindauer und Gesamtverbrauch relevant. Für das kurz- und das mittelfristige Modell (also 1 und 24 Stunden) war zudem der Kosinus der Zeit als Einflussgröße laut Kendall-Ranks

1-Stunde	**24-Stunden**	**7-Tage**
Netzbezug	Netzbezug	Netzbezug
Netzeinspeisung	Netzeinspeisung	Netzeinspeisung
Bezug PV-Analge	Bezug PV-Analge	Bezug PV-Analge
PV-Einspeisung	PV-Einspeisung	PV-Einspeisung
Anzahl Mieter	Anzahl Mieter	Anzahl Mieter
Wochennummer	Diffuse Strahlung	Diffuse Strahlung
	Globalstrahlung	Globalstrahlung
	Sonennscheindauer	Sonennscheindauer

Tabelle 4: Einflussgrößen der unterschiedlichen Modelle zur Vorhersage des Verbrauchs der Tanzenden Siedlung

1-Stunde	**24-Stunden**	**7-Tage**
Netzbezug	Netzbezug	Netzbezug
Netzeinspeisung	Netzeinspeisung	Netzeinspeisung
Bezug PV-Analge	Bezug PV-Analge	Niederschlagsdauer 10 Minuten
PV-Einspeisung	PV-Einspeisung	PV-Einspeisung
Kosinus der Zeit	Kosinus der Zeit	Niederschlagsindikator
diffuse Himmelstrahlung	diffuse Himmelstrahlung	diffuse Himmelstrahlung
Globalstrahlung	Globalstrahlung	Globalstrahlung
Sonnenscheindauer	Sonnenscheindauer	Sonnenscheindauer
Gesamtverbrauch	Gesamtverbrauch	Gesamtverbrauch

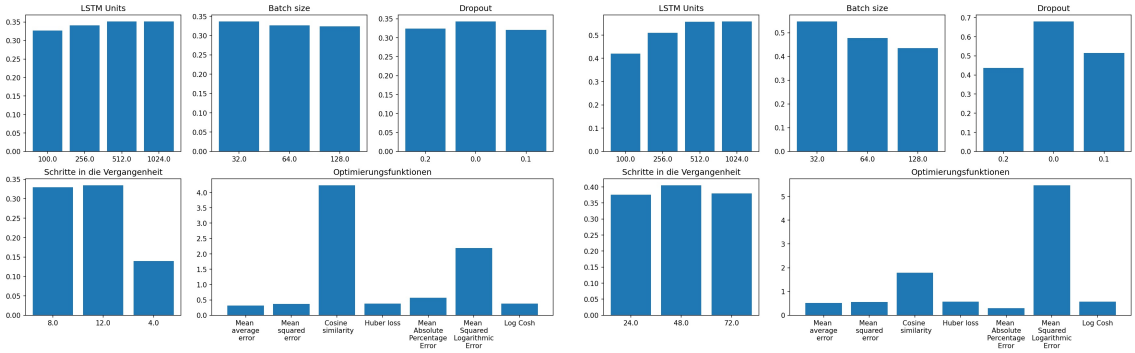
Tabelle 5: Einflussgrößen der unterschiedlichen Modelle zur Vorhersage der Einspeisung der Tanzenden Siedlung

relevant, beim 7-Tage Modell summierte sich dieser erneut für jeden Tag zu 0 auf, sodass daraus keine Informationen gewonnen werden konnten. Dagegen wurden beim 7-Tage Modell die Größen Niederschlagsdauer 10 Minuten und Niederschlagsindikator relevant.

5.4.1 Eine Stunde Vorhersage Einspeisung

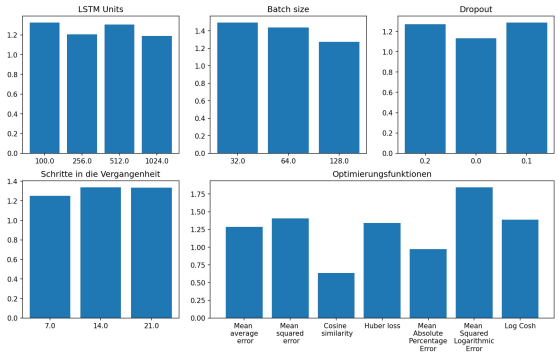
Da die Daten für die Tanzende Siedlung in 15-Minuten Schritten vorliegen, müssen für die Vorhersage der nächsten Stunde lediglich vier Werte ermittelt werden. Dies wurde auf Basis der Daten der letzten ein, zwei und drei Stunden vorhergesagt. Modelle mit zwei oder drei Stunden Eingabedaten unterschieden sich kaum voneinander, wurden jedoch lediglich eine Stunden als Eingabedaten gewählt, halbierte sich die RMSE beinahe. Daher soll das endgültige Modell Vorhersagen auf Basis der Daten der letzten Stunde, also auf Basis der letzten 4 Werte treffen.

Für die 1-Stunde Vorhersage der Einspeisung lässt sich der deutlichste Unterschied auf die resultierende RMSE in der Wahl der Loss Function feststellen. Wiederum ist der Cosine Loss mit Abstand am schlechtesten geeignet. Am besten geeignet hingegen ist nach RMSE die Verwendung des Mean Average Errors als Loss Function mit einem Wert von lediglich 0,32. Auch in diesem



(a) 1-Stunde Modell

(b) 24 Stunden Modell



(c) 7 Tage Modell

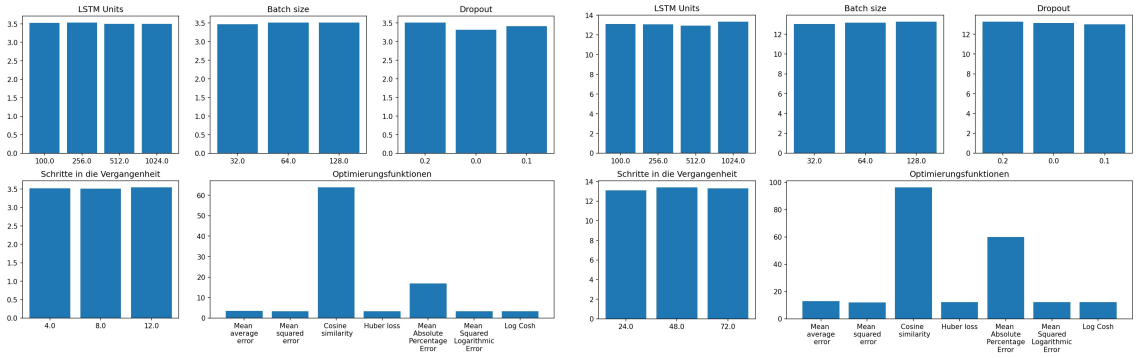
Abbildung 10: Visualisierung der durchschnittlichen RMSEs für die unterschiedlichen Modelle zur Vorhersage der PV-Einspeisung in der Tanzende Siedlung

Modell ist auffällig, dass alle verwendeten Funktionen auch negative Werte vorhersagen, obwohl die Eingabedaten keine negativen Werte enthalten. Besonders deutlich lässt sich dieser Effekt beim Mean Suqared Logarithmic Error erkennen, der folgerichtig die schlechtesten Modelle nach RMSE trainiert.

Bei den LSTM Units lässt sich erkennen, dass die Güte des Modells sinkt, je mehr Units verwendet werden. Ergibt sich bei 100 Units noch eine RMSE von 0,32, so sind es bei 1024 verwendeten Units bereits 0,35. Bezüglich des Dropouts lässt sich hier beobachten, dass die Verwendung eines Dropouts das Modell nicht nur stabiler sondern auch nach RMSE besser macht: Bei Verwendung eines Dropouts von 10 % beziehungsweise 20 % ergeben sich RMSEs von jeweils 0,32, wird kein Dropout angegeben ein RMSE von 0,34.

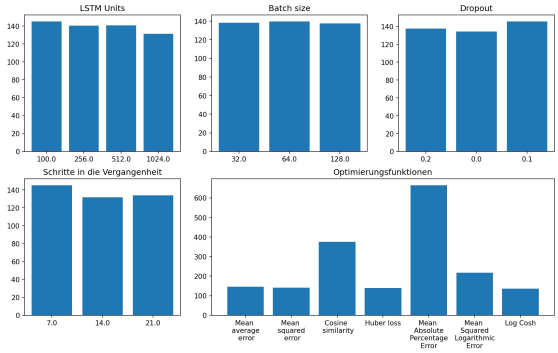
5.4.2 24 Stunden Vorhersage Einspeisung

Für die Vorhersage der Einspeisung der nächsten 24 Stunden auf den Daten der tanzenden Siedlung fällt deutlich ins Auge, dass die Verwendung von mehr LSTM-Units zu schlechteren RMSEs führt: Bei 100 verwendeten Units sind es 0,42, bei 1024 verwendeten Units bereits 0,55. Die Größe der Batch Size hingegen hat einen umgekehrten Einfluss: Je größer diese gewählt wird, desto kleiner wird die RMSE: Bei einer Batch Size von 128 beträgt die RMSE dann lediglich 0,43. Auch die Verwendung eines größeren Dropouts hat hier einen spürbaren Effekt: Ohne Dropout rangiert die RMSE bei etwa 0,67, bei 20% Dropout sind es nur noch 0,43. Auch hier lässt sich also beobachten, dass ein Dropout das Modell insgesamt nicht nur stabiler, sondern auch besser macht. Die Loss Function, die für diese Aufgabe am besten geeignet zu sein scheint ist der Mean Absolute Percentage Error mit einer RMSE von nur 0,3. Ein Blick auf den entsprechenden Vorhersageplot enthüllt allerdings, dass die Vorhersage hier fast ausschließlich auf der Nulllinie verläuft, was in diesem Falle zu einer sehr geringen RMSE führt. Die beiden nächstbestgeeigneten Loss Functionen sind der Mean Average Error und der Log Cosh. Bei Verwendung des Log Cosh werden allerdings recht viele negative Werte vorhergesagt. Da diese aber nicht vorkommen können, scheint der Mean



(a) 60-Minuten Modell

(b) 24 Stunden Modell



(c) 7 Tage Modell

Abbildung 11: Visualisierung der durchschnittlichen RMSEs für die unterschiedlichen Modelle zur Vorhersage des Gesamtverbrauchs in der Tanzende Siedlung

Average Error die Funtion zu sein, auf deren Basis das Modell trainiert werden sollte. Wiederum erscheint es am effektivsten, die Daten der letzten 24 Stunden als Eingabe heranzuziehen, sollen die nächsten 24 Stunden vorhergesagt werden.

5.4.3 Sieben Tage Vorhersage Einspeisung

Soll der Verlauf der Einspeisung über die nächsten sieben Tage vorhergesagt werden, dann erscheint die Verwendung von mehr LSTM-Units angemessen: Die kleinste RMSE beim Vergleich der verschiedenen Anzahlen verwendeter Units ergab das Modell mit 1024 Units mit einer RMSE von 1,19. Auch bei der Batch Size ergab die größte ausprobierte Eingabe von 128 die beste RMSE. Beim dropout hingegen war das Modell, das ohne Dropout trainiert wurde das beste. Dies könnte darauf zurückzuführen sein, dass auch für die tanzenden Siedlung nach Aufsummieren der Tage relativ wenige Trainingsdaten zur Verfügung stehen.

Wiederum war zu erkennen, dass die Verwendung von ebensovielen Daten in die Vergangenheit wie in die Zukunft die beste RMSE resultierte, in diesem Falle also sieben Tage. Bei den Loss Functions hingegen schien die Cosine Similarity am besten für diese Vorhersage geeignet zu sein.

5.4.4 Eine Stunde Vorhersage Verbrauch

Wie bereits in Kapitel 5.4.1 beschrieben, mussten in diesem Modell lediglich vier Werte vorhergesagt werden. Die Wahl der verschiedenen Parameter hatte in diesem Modell kaum einen Einfluss auf die resultierende RMSE. So unterschieden sich Modelle mit 100 und 1024 LSTM-Units bezüglich ihres RMSE-Werts lediglich um 0,038. Da die Trainingsdauer des Modells mit 100 LSTM-Units jedoch wesentlich geringer ausfällt als die des 1024 LSTM Units Modells, werden im resultierenden Modell 100 Units verwendet werden.

Auch bezüglich der Batch Size lässt sich kaum ein Unterschied im Bezug auf die Güte der resultierenden Modelle erkennen. Beim Dropout ist im Gegensatz dazu eine kleinere Abweichung

zu erkennen: Ein Modell mit 0 % Dropout ist zwar das Beste im Sinne der RMSE, jedoch kommen auch Modelle mit 10 % oder 20 % Dropout auf vergleichbare RMSEs. Um einerseits die Güte des Modells zu optimieren, das Modell andererseits aber auch stabiler zu gestalten, sollte daher ein Dropout von 10 % gewählt werden.

Bei den Loss Functions fällt auf, dass die Cosine Similarity denkbar schlecht für diese Aufgabe geeignet zu sein scheint. Ein Blick auf die entsprechenden Vorhersagekurve zeigt, dass die Daten zwar recht gut vorhergesagt werden, allerdings sind die Vorhersagen auf der y-Achse um etwa 45 Einheiten in positive Richtung (nach oben) verschoben, was zu einen enormen Fehler führt. Auch wenn viele der Funktionen nach RMSE nahezu eine gleiche Eignung für die Verwendung in diesem Modell aufweisen, ist am Ende der Mean Squared Logarithmic Error die beste Funktion für dieses Modell.

Auch bezüglich der Anzahl der Schritte in die Vergangenheit, auf deren Basis die nächsten vier Viertelstunden vorhergesagt werden, lässt sich kaum ein Unterschied bezüglich der RMSEs erkennen. Daher soll hier wiederum das einfachste Modell gewählt werden und das resultierende Modell mit vier Schritten in die Vergangenheit trainiert werden.

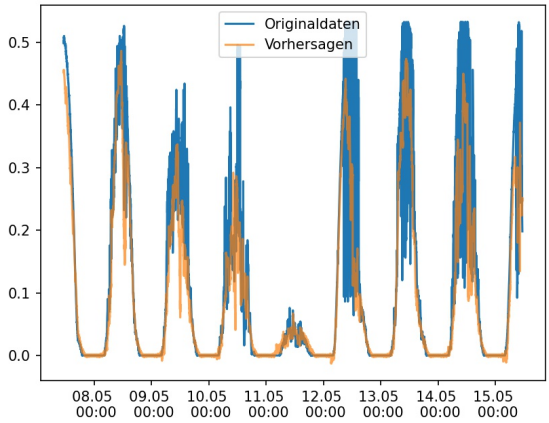
5.4.5 24 Stunden Vorhersage Verbrauch

Für diese Daten ergibt sich bezüglich der Anzahl von LSTM Units kaum ein Unterschied auf die resultierende RMSE. Das beste Modell ist in diesem Sinne jedoch das mit 512 Units mit einer durchschnittlichen RMSE von etwa 12,94. Jedoch sind auch hier die Unterschiede so marginal, dass diesen keine Relevanz zugesprochen werden kann. Im Sinne der Einfachheit des Modells sollen daher trotzdem 100 LSTM-Units verwendet werden.

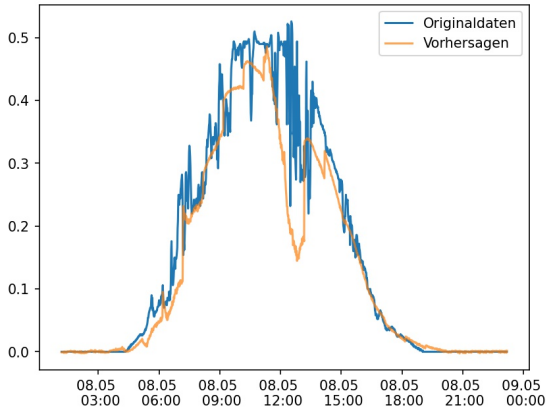
Bei der Batch Size lässt sich ein leichter Trend dahingehend erkennen, dass die Verwendung von größeren Batch Sizes zu marginal schlechteren Modellen führt. Eine Batch Size von 32 erscheint für diese Daten am besten geeignet zu sein. Bezüglich des Dropouts ergeben sich die besten Modelle bei einer Verwendung 10 % Dropout. Bei der Betrachtung der Vergangenheitsschritte, auf deren Basis das Modell Vorhersagen generiert, fällt wiederum auf, dass das Modell, das ebenso viele Schritte in die Vergangenheit wie in die Zukunft rechnet - in diesem Falle 24 - die besten Resultate erzielt. Wiederum sind die Abweichungen gegenüber den anderen Modellen jedoch sehr gering. Die bestgeeignete Loss Function ist der Huber Loss. Modelle, die mit dieser Loss Function trainiert wurden erreichten durchschnittlich eine RMSE von 12,14.

5.4.6 Sieben Tage Vorhersage Verbrauch

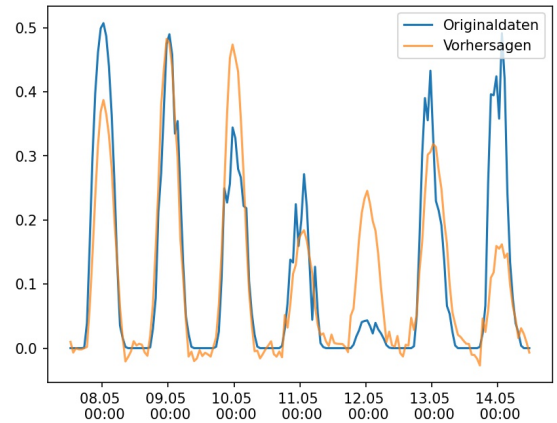
Für die Vorhersage des Verbrauchs der nächsten sieben Tage sinkt die RMSE weiter, je mehr LSTM-Units verwendet werden. Somit ergibt sich die beste RMSE für das Modell mit 1024 LSTM-Units. Wiederum hat das Dropout kaum einen nennenswerte Einfluss auf die resultierende RMSE, weswegen im Sinne der Stabilität des Modells gegen Ausreißer 20 % Dropout im resultierenden Modell verwendet werden sollen. Bezüglich der Anzahl der Vergangenheitsschritte, aus denen die nächsten sieben Tage vorhergesagt werden erscheint ein Modell mit 14 Tage Input bei sieben Tagen Output mit einer RMSE von 131,7 am besten geeignet zu sein. Bezüglich der verwendeten Loss Function gilt selbiges für die Log Cosh Function.



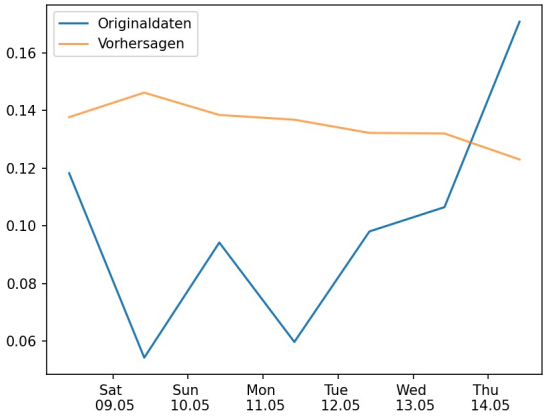
(a) 60-Minuten Vorhersagen über alle Testdaten



(b) 60-Minuten Vorhersagen über den ersten Tag der Testdaten



(c) 24 Stunden Vorhersagen der PV-Stromproduktion in der Alfons-Pech-Straße



(d) 7 Tage Vorhersagen der PV-Stromproduktion in der Alfons-Pech-Straße

Abbildung 12: Vorhersagen der PV-Stromproduktion in der Alfons-Pech-Straße

6 Analyse des Vorhersagen

In diesem Kapitel sollen die endgültigen Modelle, deren Parameter im vorherigen Kapitel experimentell ermittelt wurden, bezüglich ihrer Fähigkeit die Realität vorherzusagen bewertet werden. Dies geschieht einerseits auf Basis des bereits erläuterten RMSE, andererseits anhand des Vergleichs der Verlaufskurven der Vorhersagen mit den Verlaufskurven der Realität.

Die grafische Aufbereitung der Vorhersagen gestaltete sich allerdings problematisch. Dies soll hier am Beispiel eines 60-Minuten Vorhersagemodells für die Alfons-Pech-Straße erläutert werden: Für jede einzelne Zeile der Testdaten, und somit für jede einzelne Minute werden die nächsten 60 Minuten vorhergesagt. Daher existieren für jede Minute in den Testdaten (ausgenommen die ersten und letzten 60 Minuten) insgesamt 60 Vorhersagewerte. Diese alle in einer Grafik darzustellen würde in einer wenig aussagekräftigen Grafik resultieren, da eine Vielzahl unterschiedlicher Kurven sich gegenseitig überlagern würde. Daher wurde nur jede 60. Minute herangezogen und die jeweiligen Vorhersagen (für die nächsten 60 Minuten) zu einer einzigen Kurve zusammengefügt. Dies ist als “Vorhersagen”-Kurve in den nachfolgend präsentierten Diagrammen zu sehen.

6.1 Vorhersage über die PV-Stromproduktion in der Alfons-Pech-Straße

Abbildung 12 zeigt die Vorhersagen zur PV-Stromproduktion in 60-Minuten Schritten für die Alfons-Pech-Straße. Abbildung 12a stellt dabei Vorhersagen für sämtliche Testdaten dar, während Abbildung 12b den ersten Tag dieser Testdaten als vergrößerten Ausschnitt darstellt. Insgesamt

lässt sich aus diesen Grafiken ablesen, dass die Vorhersagen den Verlauf der realen Daten recht gut treffen. Unstetigkeiten der Produktionskurve, die auf Einflüsse wie Wolken, die sich vor die Sonne schieben zurückzuführen sind, sind in diesem Modell nicht vorhersagbar. Somit können die Produktionsdaten nicht exakt vorhergesagt werden. Insbesondere Abbildung 12b lässt jedoch erkennen, dass die durchschnittliche Steigung des jeweiligen Kurvenabschnitts durch die Vorhersagekurve relativ gut angenähert wird. Dies lässt sich auch in der durchschnittlichen RMSE erkennen, das Modell erreicht dabei einen Wert von 0,071.

Bei Vorhersage der PV-Stromproduktion der nächsten 24 Stunden lässt sich erkennen, dass beide Kurven in etwa gleich breite Peaks aufweisen. Die Ausschläge der Peaks hingegen stimmen nur teilweise überein und weichen in manchen Fällen (etwa der 14.05.) deutlich voneinander ab. Ebenfalls auffällig ist die offenkundige Unfähigkeit des Modells, die Nullproduktion in der Nacht vorherzusagen. Stattdessen schwankt die Vorhersagekurve dort um die Nulllinie herum, somit werden auch negative Produktionswerte vorhergesagt. Da es sich bei den Eingabedaten jedoch um die reine Nettostromproduktion durch die PV-Anlage handelt, können keine negativen Werte entstehen.

Besonders auffällig sind die Vorhersagen für den 12.05., da an diesem Tag kaum Strom produziert wurde. Das Modell hingegen sagt eine Produktion, die die des Vortages sogar übersteigt vorher. Allerdings muss dazu erwähnt werden, dass der 12.05. der einzige Tag in den Daten mit einer derart niedrigen Produktion ist. Da sich dieser Tag in den Testdaten befindet, hat das Modell im Training keine solche Eingabe verarbeiten müssen. Daher erscheint es naheliegend, dass die Vorhersage dieses Verlaufs schwer fällt.

Wie bereits zuvor angedeutet, liegen für die Alfons-Pech-Straße nach Summation der Werte nach Stunden relativ wenige Daten vor, anhand derer das Modell trainiert werden kann. Daher erscheint es logisch, dass die Vorhersagen entsprechend schlechter sind, als die des 60-Minuten Modells. Trotz allem erreicht das Modell eine RMSE von 0,074.

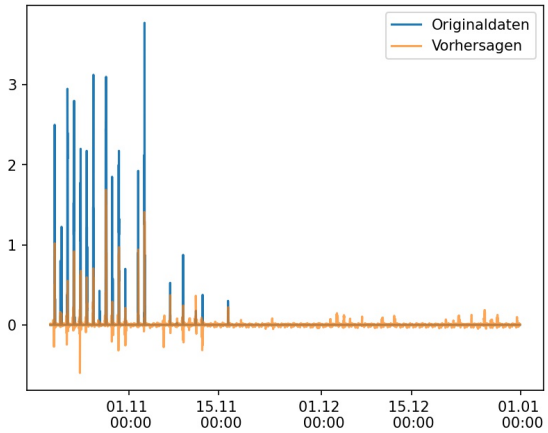
Das Problem, die jeweiligen Modelle nur aus Basis weniger Daten trainieren und evaluieren zu können verschlimmert sich, je größer die Gruppen werden, nach denen Aufsummiert wird. Nach Summation der Daten der Alfons-Pech-Straße nach Tagen mussten sogar weniger Daten als Trainingsdaten (67 % statt 80 %) verwendet werden, damit in den Testdaten überhaupt eine Folge von sieben Tage enthalten ist, die (optisch) mit den Vorhersagen verglichen werden kann. Abbildung 12d zeigt diesen Vergleich. Wie dort zu sehen ist, liegen die Vorhersagen weit entfernt von den Daten der Realität. Dies muss jedoch nicht bedeuten, dass diese Daten prinzipiell nicht vorhersagbar sind. Vielmehr sollte dieses Modell erneut trainiert werden, wenn mehr Daten verfügbar sind um eine generelle Machbarkeit zu evaluieren.

Insgesamt nähert das 60-Minuten Modell die Realität recht gut an, auch wenn Unstetigkeiten in den Verlaufskurven nicht vorhergesagt werden können. Da diese jedoch auf (hier) unvorhersehbare Wetterereignisse wie kurzfristige Verdunklung der Sonne durch Wolken zurückzuführen sind, konnte dies auch nicht erwartet werden. Auch im 24-Stunden Modell kann die Realität noch einigermaßen gut angenähert werden, das 7 Tage Modell ist allerdings in dieser Form für Vorhersagen nicht zu gebrauchen. Hier ist zudem anzumerken, dass nur 1,5 (Frühlings-) Monate Teil der Daten sind. Sollen auf dieser Datenbasis Stunden oder Tage im Sommer, Herbst oder Winter vorhergesagt werden, so ist zu erwarten, dass diese Vorhersagen schlechter sind als diejenigen für den Frühling.

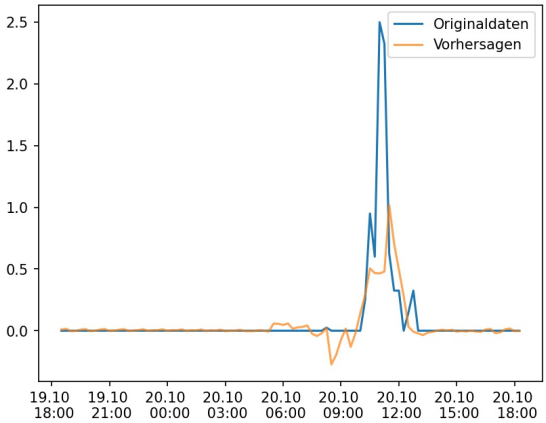
Die Modelle erscheinen in ihrer aktuellen Form schon recht gut zu funktionieren (wenigstens das 60-Minuten und das 24 Stunden Modell), würden sie mit mehr Daten von ganzen oder sogar mehreren Jahren trainiert werden, wären allerdings fundiertere und bessere Vorhersagen zu erwarten.

6.2 Vorhersage über die Einspeisung der tanzenden Siedlung

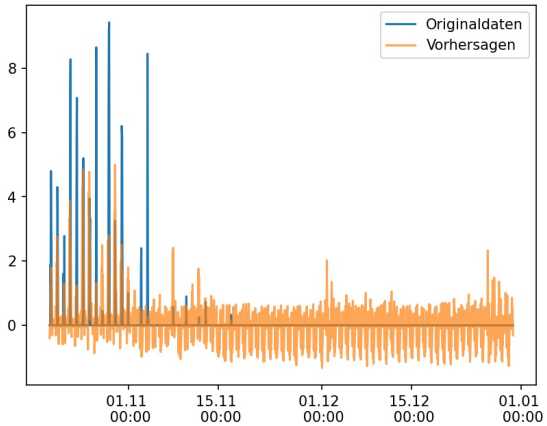
Bei den Vorhersagen der Stromeinspeisung in der tanzenden Siedlung fällt auf, dass die Spitzen der realen Daten zwar zeitlich gut getroffen werden, allerdings wird die Höhe der Spitze nur ein einziges Mal überschätzt, alle anderen Schätzungen bleiben weit hinter der realen Kurve zurück. Abbildung 13b zeigt die Vorhersagen für den ersten (vollständigen) Tag der Testdaten vom 19. auf den 20.10. an. Hierbei ist auffällig, dass die Kurve der Vorhersagen ins negative "schwingt", bevor sie in den positiven Wertebereich übergeht. Auch für die Einspeisedaten der tanzenden Siedlung können allerdings eigentlich keine negativen Werte vorkommen. An diesem Tag lässt sich



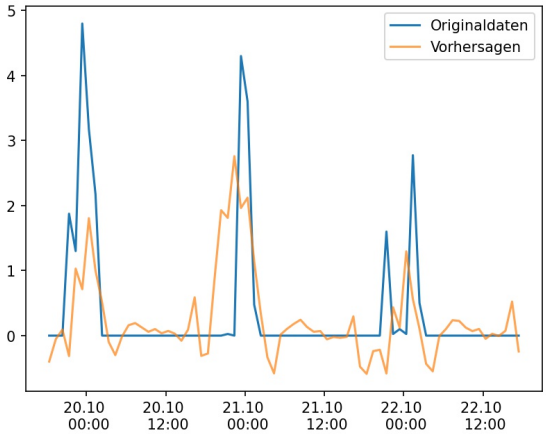
(a) Vorhersagen der jeweils nächsten 4 Viertelstunden über alle Testdaten



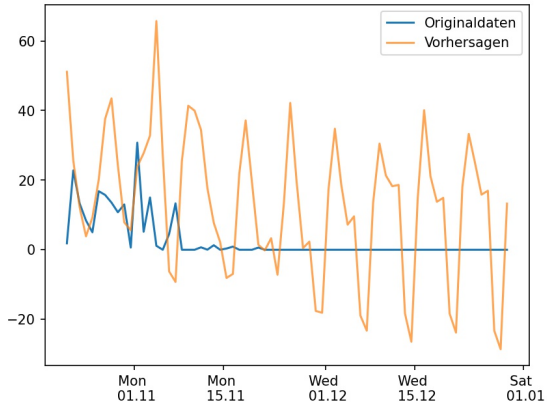
(b) Vorhersagen der jeweils nächsten 4 Viertelstunden über den ersten Tag der Testdaten



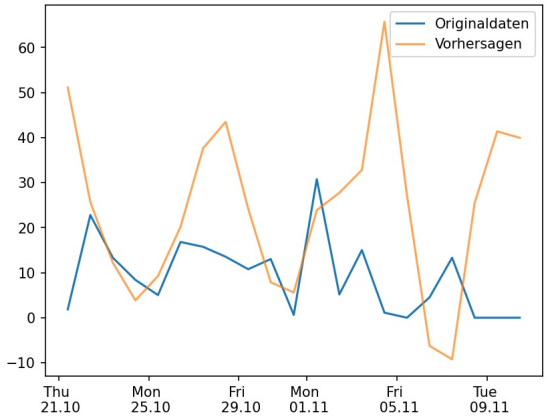
(c) Vorhersagen ganzer Stunden über alle Testdaten



(d) Vorhersagen ganzer Stunden über die ersten drei Tage der Testdaten



(e) Vorhersagen ganzer Tage über alle Testdaten



(f) Vorhersagen ganzer Wochen über die ersten drei Wochen der Testdaten

Abbildung 13: Vorhersagen der Einspeisung der Tanzenden Siedlung

beobachten, dass der Peak der Vorhersage zwar in etwa an dem Zeitpunkt liegt, an dem auch der reale Peak liegt, jedoch wird auch hier die Höhe dieses Peaks bei Weitem unterschätzt. Insgesamt kommt das Modell so auf eine RMSE von 0,14.

Ein recht ähnliches Bild zeichnet sich bei der Vorhersage der nächsten 24 Stunden ab: Hier werden einige Einspeisepeaks vom Modell stark über-, andere hingegen stark unterschätzt (vgl. Abbildung 13c). Die Abbildung 13d zeigt die Vorhersagen für die ersten drei Tage. Hier fällt noch stärker als im 1-Stunden Modell auf, dass die Nulllinie der Stromproduktion nachts nicht vorhergesagt werden kann. Zudem lässt sich beobachten, dass die Vorhersage vor und nach der eigentlichen Stromproduktionskurve erneut ins negative "schwingt". Dieser Effekt scheint sich zu verstärken, je geringer die Peaks der Stromproduktionskurve werden.

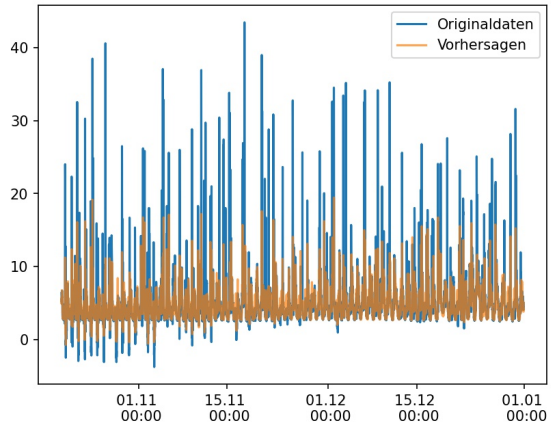
Auch beim 7-Tage Modell lässt sich beobachten, dass tagesscharfe Vorhersagen ganzer Wochen erheblich von der tatsächlichen Produktionskurve abweichen. Diese ist in weiten Teilen der Testdaten gleich 0, da die Testdaten nur den November und Dezember enthalten und es in diesen Wintermonaten kaum noch genug Sonennschein gibt, um PV-Strom zu produzieren. Das Modell hingegen sieht eine starke Fluktuation vorher. Das könnte darauf zurückzuführen sein, dass diese Monate dem Modell im Training fehlten und der oben geschilderte Zusammenhang dementsprechend nicht erlernt werden konnte. In der aktuellen Form ist das Modell zu Vorhersagezwecken nicht geeignet. Dies könnte sich allerdings ändern, wenn es mit Daten mehrerer Jahre trainiert würde.

6.3 Vorhersage über den Verbrauch der tanzenden Siedlung

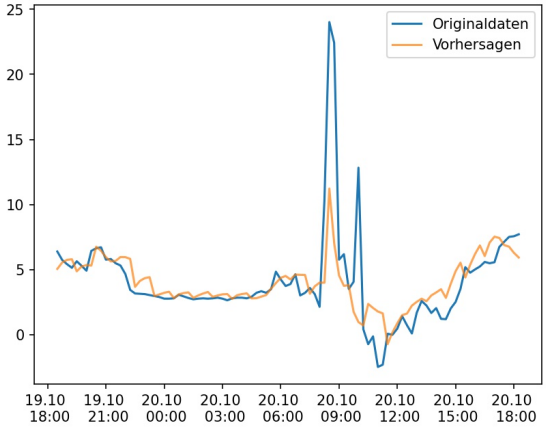
Das Modell, dass die nächste Stunde des Verbrauchs auf Basis von Viertelstunden vorhersagt schafft es recht gut, den Verlauf der realen Kurve anzunähern: Der nahezu konstante Verlauf von etwa 22 bis 6 Uhr sowie das Ansteigen der Kurve von etwa 11 bis 18 Uhr werden zwar nicht exakt getroffen, jedoch recht gut angenähert. Sogar der Ladevorgang eines Autos wird zumindest annäherungsweise richtig vorhergesagt. Dies ist in Abbildung 14b für den 10.10 etwa um 8 Uhr zu sehen. Damit solch eine Vorhersage getroffen werden kann, muss der Ladevorgang allerdings schon gestartet worden sein, ansonsten hat das Modell keine Chance, dieses Ereignis vorauszuahnen, denn die Startzeitpunkte der Ladevorgänge sind völlig zufällig. Hierbei kommt es also auch darauf an, auf Basis welcher Stunde die Vorhersagen getroffen werden. Die Vorhersagekurve über alle Testdaten (Abbildung 14a) zeigt an, dass viele dieser Lastspitzen zwar erkannt, deren Peak allerdings stets unterschätzt werden. Bezüglich des Trainings kommt hier erschwerend hinzu, dass zwischen normalen und Schnellladevorgängen unterschieden werden muss, die jeweils unterschiedlich hohe Peaks in der Kurve erzeugen und es so zusätzlich erschweren, die zukünftige Last richtig vorherzusagen.

Auch bei der stundengenauen Vorhersage der nächsten 24 Stunden lässt sich eine gute Annäherung der Vorhersage- an die tatsächliche Produktionskurve erkennen. Dort ist zu erkennen, dass der Verbrauch der Siedlung nachts sinkt und meist gegen Mitternacht sein Minimum erreicht. Gegen Morgen steigt der Verbrauch dann rapide an, um über den Tag verteilt wieder abzusinken. Diesen Zyklus bildet das Modell gut ab. Die einzigen Ausnahmen dieses Verlaufes stellen erneut die Ladevorgänge von Elektrofahrzeugen dar. Um die aus den Ladevorgänge resultierenden Lastkurven vorherzusagen, müssen diese zu der Zeit, auf der die Vorhersage beruht, bereits gestartet und noch nicht abgeschlossen worden sein. In den hier gezeigten Grafiken ist dies jedoch nicht der Fall, sodass die beiden Ladevorgänge nicht in der Vorhersagekurve zu sehen sind.

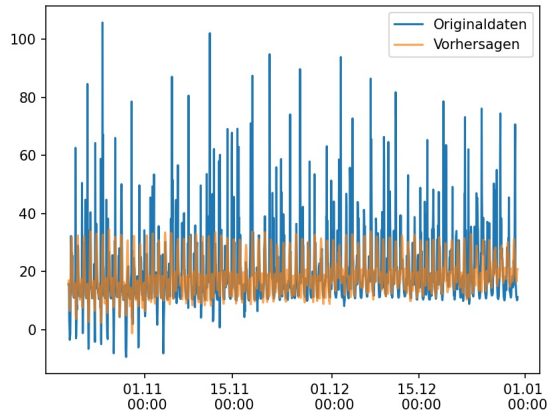
Die Vorhersage ganzer Tage hingegen funktioniert allenfalls näherungsweise. Wie in Abbildung 14e zu sehen ist, steigt der Stromverbrauch über die Tage fluktuierend an, während das Modell weiterhin einen relativ konstanten (und ebenfalls fluktuierenden) Stromverbruach prognostizierte. Dies könnte darauf zurückzuführen sein, dass es sich hierbei um einen jahreszeitlichen Effekt handelt. Da jedoch lediglich ein Jahr als Datengrundlage vorliegt, können solcherlei Effekte noch nicht erlernt worden sein. Auch hier verspricht die Verwendung von mehr Eingabedaten (und insbesondere Trainingsdaten) ein besseres Modell.



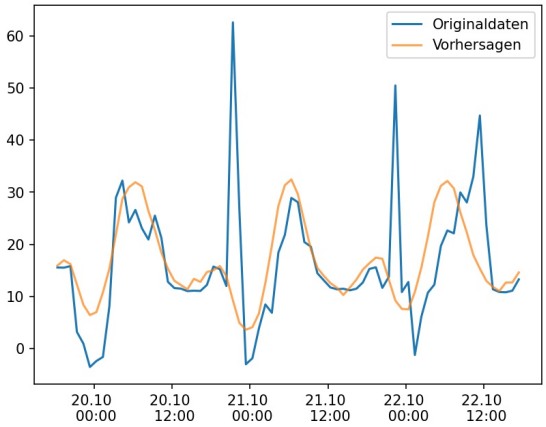
(a) Vorhersagen der jeweils nächsten 4 Viertelstunden über alle Testdaten



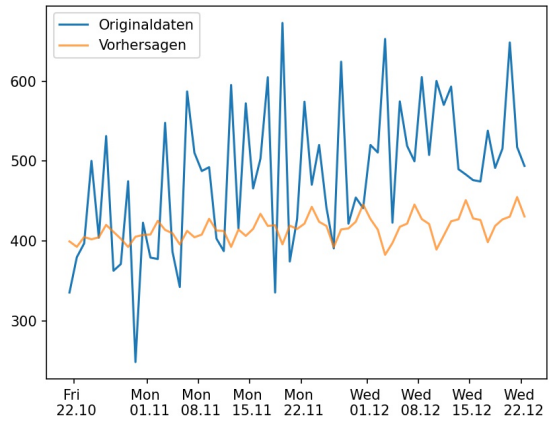
(b) Vorhersagen der jeweils nächsten 4 Viertelstunden über den ersten Tag der Testdaten



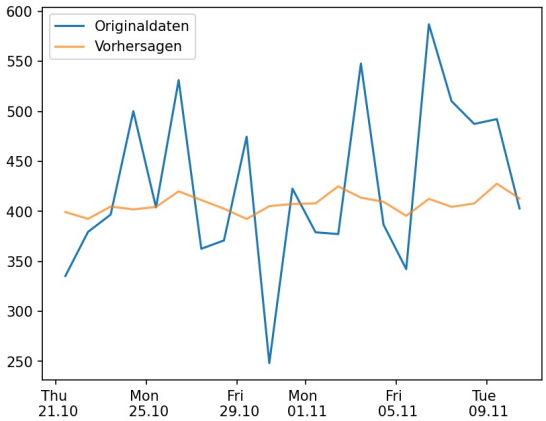
(c) Vorhersagen ganzer Stunden über alle Testdaten



(d) Vorhersagen ganzer Stunden über die ersten drei Tage der Testdaten



(e) Vorhersagen ganzer Tage über alle Testdaten



(f) Vorhersagen ganzer Tage über die ersten drei Wochen der Testdaten

Abbildung 14: Vorhersagen des Verbrauchs der tanzenden Siedlung

7 Fazit

Zusammenfassend lässt sich also sagen, dass die Vorhersage der Stromproduktion der Alfons-Pech-Straße für das 60 Minuten Modell bereits gut funktioniert. Die Vorhersage der nächsten 24 Stunden weicht schon stärker von der Realität ab, nähert die Realität aber noch so weit an, dass es sinnvoll erscheint, dieses Modell auch in der Praxis einzusetzen. Die Vorhersage der nächsten 7 Tage hingegen hat nicht mehr viel mit der Realität zu tun, hier scheint eine praktische Nutzung auf Basis der aktuell verfügbaren Daten nicht sinnvoll.

Der Abfall der Vorhersagekraft mit dem größer werdenden Vorhersageintervall liegt wohl darin begründet, dass für die Alfons-Pech-Straße mit etwa 1,5 Monaten ohnehin wenig Daten vorliegen, welche nach Aufsummation weiter zusammenschrumpfen. Dadurch müssen die Modelle mit immer weniger Traininsgdaten trainiert werden und verlieren somit an Vorhersagekraft. Außerdem beinhalten die Daten lediglich zwei Frühlingsmonate, sodass Vorhersagen für Monate anderer Jahreszeiten wohl weniger genau getroffen werden könnten. Um dieses Problem zu lösen, sollten die Modelle mit mehreren Jahren an Eingabedaten trainiert werden, um auch jahreszeitliche Effekte zu beachten.

Soll die Einspeisung der tanzenden Siedlung vorhergesagt werden, so schaffen die kurz- und mittelfristigen Modelle den Verlauf der Einspeisekurve der Realität einigermaßen anzunähern. Allerdings sind diese zeitlich oft leicht verschoben, der Peak der Einspeisung wird in den meisten Fallen stark unterschätzt. Ebenfalls ist auffällig, dass insbesondere beim Modell für die mittelfristigen Vorhersagen starke Schwankungen für die Nachtstunden vorhergesagt werden, obwohl die Einspeisung dort stets bei 0 liegt. Eine Eignung für die praktische Anwendung dieses Modells (mit diesen Trainingsdaten) ist hier also nur begrenzt gegeben. Zudem ist fraglich, ob sich diese Modelle mittels einer Erhöhung der Menge der Trainingsdaten verbessern ließen, oder ob solche Verläufe mit LSTMs generell nicht gut vorhergesagt werden können.

Der Verbrauch der tanzenden Siedlung lässt sich mit den vorgestellten kurz-und mittelfristigen Modellen sehr gut vorhersagen, das langfristige Modell erscheint wenig aussagekräftig. Werden lediglich die nächsten vier Viertelstundenwerte vorhergesagt ist es mitunter sogar möglich, mit der Vorhersagekurve die Lastkurve von Ladevorgängen von Elektroautos anzunähern, sofern der Ladevorgang zu dem Zeitpunkt, auf dem die Vorhersage basiert bereits gestartet wurde. Auch bei Vorhersagen der Werte der nächsten Stunde kann die Verlaufskurve sehr gut angenähert werden, bezüglich der Ladevorgänge gilt hier das gleiche: Diese können wohl nur vorhergesagt werden, wenn sie bereits gestartet wurden. Sollen allerdings ganze Wochen vorhergesagt werden, so weichen die Vorhersagen stark von der Realität ab. Auch hier könnte ein Modell, dass mit den Daten mehrere Jahre trainiert wird Abhilfe schaffen, denn dabei könnten auch jahreszeitliche Effekte gelernt werden.

8 Ausblick

Um die Modelle weiter zu verbessern, können noch einige Maßnahmen ergriffen werden: Wie bereits mehrfach ausgeführt wurde, sollten zunächst mehr Eingabedaten gesammelt werden, auf deren Basis die Modelle trainiert werden können. Idealerweise sollte es sich dabei um die Daten mehrerer Jahre handeln, sodass jahreszeitliche Effekte, wie eine geringere PV-Stromeinspeisung der Tanzenden Siedlung ins öffentliche Netz in den Wintermonaten gelernt und entsprechend vorhergesagt werden können.

Weiterhin sollten zusätzliche Einflussgrößen erhoben werden, die etwa die baulichen Umstände der Solarkraftwerke in Betracht ziehen. Denkbar wäre etwa, dass der Einfallwinkel der Sonnenstrahlen auf die PV-Anlage eine wertvolle Information für die Vorhersage der Stromproduktion ergeben würden.

Zudem sollte eruiert werden, ob es möglich ist, Vorhersagen anderer Modelle in die hiesigen Modelle zu integrieren. So wird etwa die Vorhersage des Wetters durch den DWD genauer sein als das in den Modellen vorhergesagte Wetter. Ein Modell, dass sich zusätzlich auf diese Information stützt, wird daher genauer sein, als eines, dass das Wetter selbst vorhersagt.

Literatur

[1] J. Brownlee. *A gentle Introduction to Long Short-Term Memory Networks by the Experts*. https://machinelearningmastery.com/gentle-introduction-long-short-term-memory-networks-experts/. 11.02.2022.

[2] J. Brownlee. *Long Short-Term Memory Networks With Python - Jason Brownlee*. https://questioneurope.blogspot.com/2020/08/long-short-term-memory-networks-with.html. 28.04.2022.

[3] K. Chowdhury. *10 Hyperparameters to keep an eye on for your LSTM model — and other tips*. https://medium.com/geekculture/10-hyperparameters-to-keep-an-eye-on-for-your-lstm-model-and-other-tips-f0ff5b63fcd4. 24.02.2022.

[4] *Diffuse Himmelsstrahlung - Diffuse sky radiation*. https://de.abcdef.wiki/wiki/Diffuse_sky_radiation. 24.01.2021.

[5] *Einheitliche Beobachteranleitung für nebenamtliche Stationen*. https://kunden.dwd.de/ankonda/docs/Beobachterhandbuch-Ankonda.pdf. 24.01.2021.

[6] *Encoding cyclical continuous features - 24-hour time – Ian London's Blog*. https://ianlondon.github.io/blog/encoding-cyclical-features-24hour-time/. 26.01.2021.

[7] F. A. Gers, J. Schmidhuber und F. Cummins. "Learning to Forget: Continual Prediction with LSTM". In: *Neural Computation* 12.10 (Okt. 2000), S. 2451–2471. ISSN: 0899-7667. DOI: 10.1162/089976600300015015. eprint: https://direct.mit.edu/neco/article-pdf/12/10/2451/814643/089976600300015015.pdf. URL: https://doi.org/10.1162/089976600300015015.

[8] S. Hochreiter und J. Schmidhuber. "Long Short-Term Memory". In: *Neural Computation* 9 (1997), S. 1735–1780.

[9] *Keras Loss Functions – Types and Examples*. https://data-flair.training/blogs/keras-loss-functions/. 30.03.2022.

[10] *Keras Unterschied zwischen val_lossundlosswährenddesTrainings*. https://qastack.com.de/datascience/25267/keras-difference-beetween-val-loss-and-loss-during-training. 28.04.2022.

[11] *Long short-term memory*. https://de.wikipedia.org/wiki/Long_short-term_memory. 12.02.2022.

[12] *Mieterstrom E-Mobilität – Erfahrungen aus den WINNER-Projekten*. https://www.durchblick-energiewende.de/sites/default/files/2021-01/201201-01-CSG-VZSH-WINNER-WINNER-Reloaded.pdf. 24.01.2021.

[13] C. Olah. *Understanding LSTM Networks – colah's blog*. https://colah.github.io/posts/2015-08-Understanding-LSTMs/. 11.02.2022.

[14] Y. Ping, C. Chen, L. Wu, Y. Wang und M. Shu. "Automatic Detection of Atrial Fibrillation Based on CNN-LSTM and Shortcut Connection". In: *Healthcare* 8 (Mai 2020), S. 139. DOI: 10.3390/healthcare8020139.

[15] *Predicting Sequential Data using LSTM: An Introduction*. https://towardsdatascience.com/time-series-forecasting-with-recurrent-neural-networks-74674e289816. 28.04.2022.

[16] *Project WINNER reloaded*. https://www.winner-projekt.de/project.html. 20.01.2021.

[17] *Projekt WINNER - Mieterstrom macht mobil: Erster Mobility Point bei der Chemnitzer Siedlungsgemeinschaft eG*. https://wohnungswirtschaft-heute.de/wp-content/uploads/2018/03/Energie-Projekt-WINNER.pdf. 20.01.2021.

[18] K. Pykes. *The Vanishing/Exploding Gradient Problem in Deep Neural Networks*. https://towardsdatascience.com/the-vanishing-exploding-gradient-problem-in-deep-neural-networks-191358470c11. 12.02.2022.

[19] *Strom und Carsharing – Neues Modell bringt Solarenergie für alle Mieter*. https://www.sachsen-fernsehen.de/strom-und-carsharing-neues-modell-bringt-solarenergie-fuer-alle-mieter-459297/. 24.01.2021.

[20] *The Vanishing Gradient Problem*. https://www.mygreatlearning.com/blog/the-vanishing-gradient-problem/. 12.02.2022.

[21] *Videozusammenfassung: Erfolgreicher Abschluss des Projektes WINNER*. https://www.winner-projekt.de/winner-2016/news/2020/01/08/WINNER-Zusammenfassung.html. 20.01.2021.

[22] J. Q. Wang, Y. Du und J. Wang. "LSTM based long-term energy consumption prediction with periodicity". In: *Energy* 197 (2020), S. 117197. ISSN: 0360-5442. DOI: https://doi.org/10.1016/j.energy.2020.117197. URL: https://www.sciencedirect.com/science/article/pii/S0360544220303042.

[23] *Wetter und Klima - Deutscher Wetterdienst*. https://www.dwd.de/. 24.01.2021.

[24] *What is "units" in LSTM layer of Keras?* https://zhuanlan.zhihu.com/p/58854907. 28.04.2022.