

Dokumentation Vorhersage-Tool

Um das Potential des im Rahmen des Projekt Maggies installierten Energiesystems vollends ausnutzen zu können ist es nötig den Energie-Verbrauch in der nahen Zukunft so genau wie möglich zu kennen. Nur wenn dieser bekannt ist, ist es möglich die Fahrweise der einzelnen System-Bestandteile in einer Art und Weise zu optimieren, dass die Energieverluste minimal sind. Der im Projekt entwickelte Optimierer ermöglicht auch weitere Ziele, wie zum Beispiel maximale Autarkie etc. . Die Freiheitsgrade, welche dem Optimierer zur Verfügung stehen entstehen hauptsächlich durch die Möglichkeit Wärme-Energie in den Puffer ein- und auszuspeichern, wodurch eine Entkopplung der Erzeugung vom Verbrauch entsteht. Außerdem stehen zur Erzeugung von Wärme und Strom jeweils mehrere Möglichkeiten zur Verfügung.

Wenn die vom Optimierer angenommenen Verbräuche zu stark von den tatsächlichen Verbräuchen abweichen besteht das Risiko, dass die Fahrpläne nicht eingehalten werden könne (z.B. weil der Puffer-Speicher zu sehr auskühlt). Dies gefährdet das Gesamtziel die Verluste zu minimieren und die Effizienz des Systems zu maximieren. Deshalb ist es notwendig, dass möglichst exakte Vorhersagen der Strom- und Wärmeverbräuche zur Verfügung stehen.

Im Rahmen der Literatur-Recherche hat sich gezeigt, dass es bisher keine Vorhersage-Techniken für Zeitreihen gibt die zum einen mit relativ hoch aufgelösten Daten arbeiten können und zum anderen einen Vorhersagehorizont von mindestens 36 Stunden aufweisen. Es muss jedoch zwingend der Verbrauch der nächsten 36h vorhergesagt werden, da im Strommarkt die Ein- und Ausspeisungen in und aus dem Verteilnetz am Vortag 15-Minuten-scharf dem Verteilnetzbetreiber gemeldet werden müssen. Für das U-Gebäude ist dies zwar noch nicht der Fall, da die Nennleistung der Erzeugung zu gering ist. Perspektivisch soll das in Rahmen des Projekts entwickelte Konzept jedoch auf ganze Stadtviertel angewendet werden können, womit der Grenzwert der Nennleistung deutlich überschritten wird. Es muss folglich bereits am Vortag 15-Minuten genau geplant werden wann und wie die Erzeuger des installierten Energiesystems laufen und wie die Energie im System verteilt wird. Dazu ist das Vorhersage-Tool notwendig. Bisherige Vorhersage-Algorithmen, die mit Daten des Energiemarkts arbeiten verwenden entweder Daten eines sehr hohen Aggregationslevels, z.B. der summierte Stromverbrauch eines ganzen Bundeslandes, und sind dadurch nicht anwendbar oder der Vorhersage-Horizont ist viel kürzer. Meist werden nur die Daten einige wenige Zeitschritte in die Zukunft vorhergesagt. Diese Algorithmen zu verwenden ist nicht zielführend.

Bei der Entwicklung des Vorhersage-Tools wurde sich auf Stromverbrauchs-Daten fokussiert. Dies hat zwei Gründe. Zum einen ist, wie bereits erwähnt, die Ein- und Ausspeisung von Strom ins öffentliche Verteilnetz ein kritischer Punkt, um das neuartige Energiesystem optimal und kostengünstig betreiben zu können. Wird dem Verteilnetzbetreiber kein Fahrplan gemeldet oder wird von diesem abgewichen werden Strafzahlungen fällig. Zum anderen ist ein möglichst genaues Wissen über den Stromverbrauch nötig, da es in dem geplanten Energiesystem keine Möglichkeit gibt Strom unmittelbar zu speichern. Dies bedeutet, der benötigte Strom muss instantan erzeugt werden. Lediglich über das BHKW kann zusätzlicher Strom erzeugt werden, wodurch jedoch gleichzeitig Wärme erzeugt wird, die in den Pufferspeicher eingespeist wird, was nur bis zu einem gewissen Maße möglich ist, ohne die Integrität des Gesamt-Fahrplans des Systems zu gefährden. Überflüssiger Strom kann theoretisch ebenfalls mit der Wärmepumpe in Wärme umgewandelt werden. Dies verursacht dasselbe Problem wie bereits erwähnt und ist zusätzlich höchst ineffektiv. Zum anderen vollziehen sich Änderungen im Konsum-Verhalten von Wärme auf einer längeren Zeitskala. Teilweise liegt dies begründet in der Tatsache, dass das gesamte Gebäude als träger

Wärmespeicher fungiert. Außerdem gibt es keine direkte Kopplung zwischen der erzeugten Wärme und einem externen Netz. Es kann außerdem angenommen werden, dass das die grundlegende Modell-Struktur des Vorhersage-Tools für Strom auch für Wärme übernommen werden kann. Dazu muss dieses nur mit den entsprechenden Wärme-Verbräuchen trainiert werden.

Im Folgenden sind nun die Entwicklungsschritte des Vorhersage-Tools und die gewonnenen Erkenntnisse chronologisch aufgelistet:

- Literatur-Recherche zu existierenden Vorhersage-Tools für Energie-Zeitreihen. Die Ergebnisse wurden bereits oben besprochen
- Literatur-Recherche zu Möglichkeiten der Zeitreihenvorhersage; dazu Einarbeitung in:
 - o Software-Plattform Tensorflow und Keras
 - o Programmierung Rekurrente Neuronale Netze
 - o Programmierung LSTM-Netze
 - o Programmierung Convolutional Networks (CNN)
- Testen einfacher, verschiedener neuronalen Netze mit selbst erstellten Dummy-Daten
- Analyse des Verhaltens der Netze
- Tiefergehende Tests von LSTM-Netzen:
 - o Variation der LSTM-Strukturen
 - o Hyper-Parameter-Tuning
- Vorhersage-Ergebnisse der LSTM-Netze nicht zufriedenstellend, obwohl diese der Stand der Technik für NLP (natural language processing) sind und sehr häufig zur Erstellung von Vorhersagen verwendet werden, wenn auch nicht schwerpunktmäßig für Energieverbräuche.
 - o LSTM-Netze scheitern am Erkennen von Saisonalitäten (z.B. Wochentage)
- Intensiveres Einarbeiten in CNNs
- Erfolgreiche Tests von CNN-Modellen an synthetischen Daten --> Einbindung von Verbrauchsdaten der Margaretenau ist angedacht
- Zeitnah sind keine Daten aus der Margaretenau zu erwarten, da im U-Gebäude keine Smart-Meter verbaut sind und werden; erst nach erfolgreicher Installation des neuen Energiesystems kann mit Verbrauchsdaten gerechnet werden; geplante Inbetriebnahme des Systems ca. nach halber oder zwei-Drittel der Projektlaufzeit
- Suche nach öffentlich verfügbaren Datensätzen, die der zu erwartenden Charakteristik der Verbrauchs-Daten aus der Margaretenau entsprechen
- Verwendung und Vorverarbeitung der Daten aus dem CER Smart Meter Trial in Irland ¹ ; Der Smart Meter Trial enthält die Verbrauchsdaten über 1,5 Jahre von über 5000 Verbrauchern, zwei Drittel davon sind Privat-Haushalte. Aus diesen Daten wurden drei Datensätze erzeugt, mit denen von nun an anstelle der Verbrauchsdaten aus der Margaretenau gearbeitet wird:
 - o IRE15: kumulierter Stromverbrauch von 15 zufälligen Haushalten --> entspricht dem U-Gebäude
 - o IRE40: kumulierter Stromverbrauch von 40 zufälligen Haushalten --> entspricht den größeren Häusern in der Margaretenau
 - o IRE350: kumulierter Stromverbrauch von 350 zufälligen Haushalten --> entspricht ungefähr dem Verbrauch des gesamten Viertels
- Programmierung eines ersten Modells für die neuen Daten

¹ <https://www.ucd.ie/issda/data/commissionforenergyregulationcer/>

- Durchführen eines Grid-Search der Hyper-Parameter um optimale grundlegende Parameter des Netzes zu definieren --> Anpassung der Netzarchitektur
- Implementierung eines Vorverarbeitungsprozesses:
 - o Einbindung der Temperaturzeitreihe, die von der Webseite des Irischen Wetterdienstes extrahiert wurde
 - o Regression der Eingabedaten und deren Varianz mit der Temperatur --> Verminderung des Temperatureinflusses auf die Verbrauchsdaten
 - o Normalisierung
- Fertigstellung eines ersten funktionsfähigen Vorhersage-Modells mittels CNN
- Verzögerungen beim Einbau des Heizungssystem führen dazu, dass weiterhin keine nutzbaren Monitoring-Daten für Strom- und Warmwasserverbrauch zur Verfügung stehen; zur Validierung des Modells sind jedoch mindestens mehrere Monate an Daten nötig, die aus verschiedenen Wärmeperioden des Jahres stammen müssen, um Bias zu vermeiden; Des Weiteren sollte das Modell mit Daten aus der Margaretenau trainiert werden, um optimal zu funktionieren. Dazu sind ein Jahr an Verbrauchsdaten nötig, die nach dem ursprünglichen Projekt-Plan zur Verfügung gestanden wären
- Planung der Validierung des Modells mit echten Verbrauchsdaten
- Anstreben einer virtuellen Simulations- und Testumgebung. Diese ersetzt jedoch nicht eine Validierung am Objekt
- Validierung des Modells mit vergleichbaren Daten aus dem Smart Meter Trial
- Weitere Verbesserungen des CNN-Modells anhand der Validierungsergebnisse
- Extraktion von Features aus den CER Daten
- Programmieren weiterer Vorhersage-Modelle, die mit einer geringeren Daten-Menge bereits zufriedenstellende Ergebnisse liefern (RandomForest, XGBoost, RidgeRegression, Exponential Smoothing, fully-connected NN-Modelle), da nicht zu erwarten ist, dass bis zum Ende des Projektes ausreichend Daten vorhanden sind um das entwickelte CNN-Modell ausreichend zu trainieren. Die neuen Modelle sind in der Lage mit einer geringeren Datenlage bereits Vorhersagen zu erstellen. Diese Vorhersagen sind dann selbstverständlich weniger akkurat. Zur Entwicklung und zum Training der neuen Modelle wurden ebenfalls die Daten aus dem CER Smart Meter Trial verwendet.
- Neue Vorhersage-Modelle sind entwickelt und mit CER-Daten validiert und prinzipiell einsatzbereit. Eine finale Validierung ist erst möglich wenn reale Monitoring-Daten zur Verfügung stehen.
- Es hat sich gezeigt, dass das XGBoost-Modell die vielversprechendsten Vorhersagen bei geringer Datenlage liefert

Der Programm-Code des CNN-Modells und des XGBoost-Modells sind auf der Maggie-Cloud unter "AP 3.2 --> Python Vorhersage-Tool" abgelegt.

Der Code ist vergleichbar zu einem Python-Paket aufgebaut. Dies ermöglicht eine Installation in Python. Eine detaillierte Dokumentation befindet sich im Code an den jeweiligen Stellen. Dies gewährleistet, dass das Vorhersage-Tool weiterverwendet werden kann. Zusätzlich befindet sich zu jedem Modell ein Beispiel an der entsprechenden Stelle.

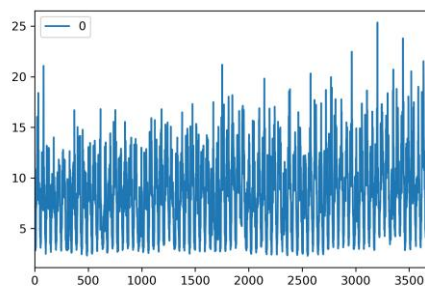
Im Folgenden sind noch die verwendeten Datensätze anhand jeweils dreier Plots dargestellt. Diese zeigen die unterschiedlichen Charakteristiken der einzelnen Datensätze. Es wird außerdem klar, warum eine Verbrauchs-Vorhersage für einzelne Häuser extrem anspruchsvoll ist: die Verbrauchsdaten sind hoch-volatil. Dies liegt daran, dass der Verbrauch eines einzelnen Haushalts einen sehr großen Einfluss auf den kumulierten Verbrauch hat. Besonders deutlich wird dies beim IRE15-Datensatz, aber auch noch bei den IRE40-Daten. Der IRE350-Datensatz weist einen glatteren Verbrauch auf, was einer geringeren Volatilität entspricht.

Die Verbrauchsdaten aus dem CER Smart Meter Trial können nicht ohne Weiteres weitergegeben werden, da bei Erhalt der Daten eine Vereinbarung mit der "Commission of Energy Regulation" getroffen wurde, die dies untersagt. Deswegen befinden sie sich nicht auf der Cloud. Es ist jedoch möglich eine Anfrage an die CER zu stellen um ebenfalls Zugriff auf die Daten zu erhalten.

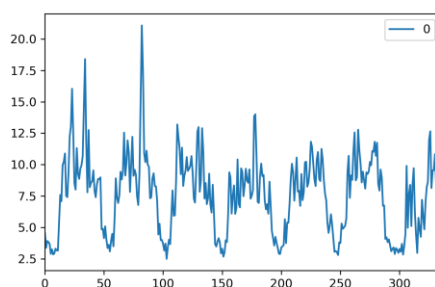
Zusätzlich befindet sich in den Unterlagen zum 4. Gesamttreffen in den Dateien 200218_09_3.2_Walbrunn_Lang.pptx, 200218_09_3.2_Lang_Plots1.html, 200218_09_3.2_Lang_Plots2.html eine Analyse verschiedener Vorhersage-Tools mit interaktiven Plots. Diese zeigen die Daten der IRE40-Zeitreihe sowie die entsprechenden Vorhersagen.

IRE15:

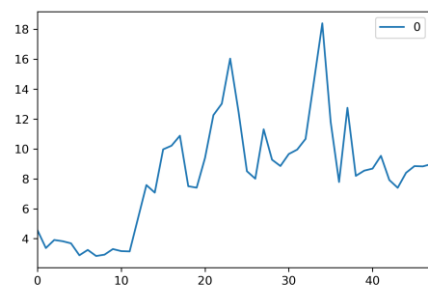
- Komplette Zeitreihe:



- Eine Woche:

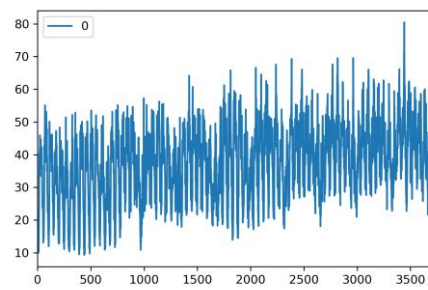


- Ein Tag:

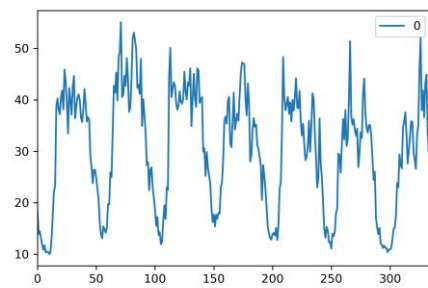


IRE40:

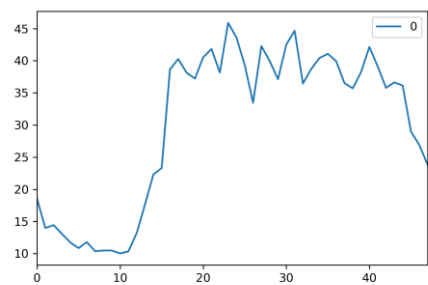
- Komplette Zeitreihe



- Eine Woche:

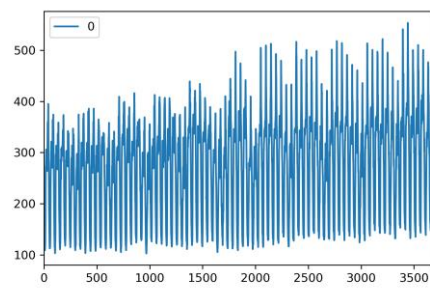


- Ein Tag:

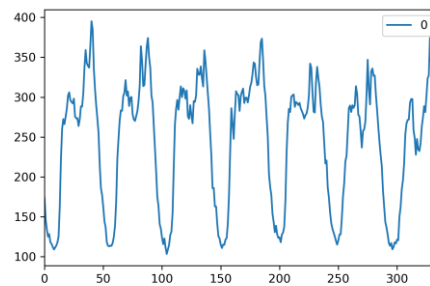


IRE350:

- Komplette Zeitreihe:



- Eine Woche:



- Ein Tag:

