



Partner:



überlingen



Assoziierte Partner:



Deliverable D3.1.3

Anforderungsspezifikation zur Datenanalyse



Erstellt im Verbundvorhaben STADTQUARTIER 2050 im Rahmen der Förderinitiative „Solares Bauen/ Energieeffiziente Stadt“ aus dem 6. Energieforschungsprogramm

Autoren:

Dr. Valerie Graf-Drasch, Fraunhofer FIT

Augsburg, 20.07.2020

Version 1.0

Inhalt

1	Allgemeine Einführung und Hintergrund	3
2	Ziel	4
3	Daten, einhergehende Anforderungen, Methodik	5
4	Literaturverzeichnis	9

1 Allgemeine Einführung und Hintergrund

Seit März 2018 fördern die Bundesministerien für Bildung und Forschung sowie Wirtschaft und Energie das Projekt „STADTQUARTIER 2050 - Herausforderungen gemeinsam lösen: Beispiel-gebende Sanierung und Nachverdichtung von Stadtquartieren zu klimaneutralen Wohnsiedlungen mit Leuchturmanwendungen in Stuttgart und Überlingen“. Das Projektkonsortium besteht unter anderem aus Partnern der Wissenschaft, Wirtschaft, und Kommunen. Das Projekt STADTQUARTIER gliedert sich in sechs Arbeitspakete. Der Fokus dieses Berichtes liegt allgemein in Arbeitspaket 3 „Technologische Fragestellungen“ und konkret in Arbeitspaket 3.1.

Im Arbeitspaket 3.1 mit dem Titel „Ermittlung feingranularer Strom- und Wärmebedarfsdaten“ sollen Verbrauchsdaten analysiert, Anwendungsmöglichkeiten der Daten geprüft, sowie der rechtliche Rahmen im Bereich Datenschutz aufgearbeitet werden. Der vorliegende Bericht widmet sich dem Deliverable 3.1.3 und somit der Anforderungsspezifikation der Datenanalyse. Dieses Dokument diskutiert hierbei konkrete Anforderungen an den Datenpool sowie statistische Methodik zu dessen Auswertung, um eine Grundlage für die Evaluation der Daten zu schaffen. Die derzeitige Version dieses Berichts entspricht dem Wissens- und Planungsstand sowie dem allgemeinen Projektfortschritt aufgrund der Bewilligungssituation zum 20. Juli 2020. Aufgrund der ausstehenden Bewilligungen sind die Möglichkeiten zur projektinternen Datenakquise derzeit stark eingeschränkt. Da noch kein ausreichender Datenpool als Voraussetzung für die nachfolgende Datenanalyse vorhanden ist, wurde der externe, irische Datensatz Commission for Energy Regulation herangezogen; dieser wurde bereits im Rahmen der Zwischenberichtserstattung als mögliche Datengrundlage identifiziert. Um den Datensatz für den genannten Zweck verwenden zu können, wurde er bestmöglich an die zu untersuchenden Quartiere angepasst, indem Haushalte, welche mit Strom heizen, sowie jegliche Nicht-Wohngebäude entfernt wurden.

2 Ziel

Das übergeordnete Ziel des Arbeitspakets 3.1 ist es, feingranulare Energiebedarfsdaten zu ermitteln und zu analysieren. Auf Basis von Energiebedarfsdaten sollen Cluster von Bewohnertypen gebildet werden, die ein definiertes Energiebedarfsprofil vorweisen. Handlungsempfehlungen, die daraus abgeleitet werden, dienen außerdem dem Arbeitspaket 5.3 für die Entwicklung des Grid Optimizers.

Das konkrete Ziel des Deliverables 3.1.3 ist es, Anforderungen für die Analyse des Datenmaterials zu formulieren. Das schließt die Formulierung von statistisch geeigneten Modellen, um Zusammenhänge in den Daten bzgl. Strom- oder Wärmeverbrauch zu erkennen, mit ein. Darüber hinaus beinhaltet die Anforderungsspezifikation auch die Identifikation geeigneter Nutzerspezifischer Merkmale (z.B. Anzahl, Alter, und/oder Beruf der Bewohner) zugunsten der Datenanalyse.

Vor der detaillierten inhaltlichen und technischen Diskussion der Anforderungsspezifikation, gilt es zu erwähnen, dass zur Vorhersage von Energiebedarfsprofilen bereits bekannte Konzepte, wie die sogenannten Standardlastprofile, bestehen. Diese werden beispielsweise vom Bundesverband der Energie- und Wasserwirtschaft (BDEW) bereitgestellt und sollen realistische Energiebedarfsprofile von Haushalten oder auch gewerblichen Kunden widerspiegeln. Allerdings sind diese Profile zu generisch gehalten, um auf eine breite Masse an Haushalten anwendbar zu sein. Um dies zu erreichen, wird das Energiebedarfsprofil anhand des spezifischen Jahresstromverbrauchs skaliert. Es findet demnach keine Unterscheidung zwischen verschiedenen Haushaltstypen auf Basis respektiver demografischen Eigenschaften wie beispielsweise Bewohnerzahl oder dem Berufsstand statt. In vorangegangenen wissenschaftlichen Arbeiten, wie beispielsweise von Töppel (2019), wurde herausgearbeitet, dass sich die Energiebedarfsprofile unterschiedlicher Haushalte neben dem Gesamtverbrauch auch im zeitlichen Verlauf unterscheiden. Dies motiviert eine detailliertere Herangehensweise und Bestimmung des Energiebedarfsprofils für jeden Haushalt.

Um eine genauere Bestimmung des Energiebedarfsprofils für jeden Haushalt zu ermöglichen, wird im Folgenden erklärt, welche Anforderungen erfüllt sein

müssen, und welche Methodik sich zu diesem Zwecke eignet. Zudem sollen aufbauend darauf im weiteren Verlauf der Bearbeitung des Arbeitspakets und einhergehender Deliverables konkrete Energiebedarfsprofile für die Anwendung im Quartier erstellt bzw. das Deliverable 3.1.4 bearbeitet werden.

3 Daten, einhergehende Anforderungen, Methodik

Als Datengrundlage für die Erstellung von Energiebedarfsprofilen werden historische Energiebedarfe, beispielsweise in Form von Stromverbrauchsdaten oder Wärmeverbrauchsdaten benötigt, auf deren Basis Aussagen über das künftige Verhalten der Bewohner getroffen werden können. Um diese Aussagen so exakt wie möglich zu formulieren, sollten die vorliegenden historischen Energiebedarfe zeitlich so hochauflösend wie möglich sein, die durch die Haushalte verursachte Last also beispielsweise sekundlich oder minütlich aufgezeichnet werden. Damit ließen sich auch kurzzeitige Lastspitzen sehr genau untersuchen und so sehr akkurate Vorhersagen der Energiebedarfsprofile erstellen. Eine Erhebung der Daten mit dieser Auflösung ist allerdings sehr aufwendig, da viel Speicherplatz für die Aufzeichnung und viel Rechenleistung für die Weiterverarbeitung benötigt wird. Sollten die Daten daher nur auf beispielsweise halbstündlicher oder stündlicher Ebene erhoben werden, ist dennoch das Ableiten von Energiebedarfsprofilen möglich. Kurzzeitige Lastspitzen werden hierbei zwar geglättet, jedoch ist es trotzdem möglich, sich damit einen guten bis sehr guten Überblick zu verschaffen und entsprechende Aktionen abzuleiten. Von den innerhalb des Projekts STADTQUARTIER 2050 betrachteten Quartiere stehen zum Zeitpunkt der Bearbeitung des vorliegenden Berichtes keine Energiebedarfsdaten zur Verfügung. Sollten diese zu einem späteren Zeitpunkt verfügbar werden, könnte man dadurch die Genauigkeit der Energiebedarfsprofile erhöhen.

Für Deutschland sind allgemein keine Stromverbrauchsdaten in ausreichender zeitlicher Auflösung und Länge frei verfügbar. Unter der Annahme, dass das Energiebedarfsverhalten irischer Haushalte den Haushalten im Fokus des Projekts STADTQUARTIER 2050 ähnelt, wird hier ersatzweise ein Datensatz aus Irland verwendet. Dieser wurde während der Datenaufbereitung so weit wie möglich an Deutschland angepasst. Dieser Datensatz wurde von der irischen Kommission für Energieregulierung (CER) während der Versuche zum

Kundenverhalten bei intelligenten Stromzählern (Electricity Smart Metering Customer Behaviour Trials) in den Jahren 2009 und 2010 erstellt und veröffentlicht. Er enthält etwa 1,5 Jahre halbstündlichen Stromverbrauch von 4.225 irischen Haushalten und 2.220 Unternehmen. Jeder Haushalt nahm zudem an einer Umfrage über seine demografischen Merkmale teil (Regulation Commission for Energy, 2012). Im Folgenden wird dieser Datensatz als „CER-Datensatz“ bezeichnet. Wärmeverbrauchsdaten lagen keine vor bzw. waren nicht verfügbar, sollten diese verfügbar gemacht werden, könnte mithilfe analoger Methodik, wie im Folgenden anhand der Stromverbrauchsdaten beschrieben, auch Energiebedarfsprofile in Form von Wärmeverbrauchsprofilen erstellt werden.

Der CER-Datensatz besteht aus einer Umfrage-Tabelle mit einer Auswahl an demografischen Daten für jeden Haushalt und den Stromverbrauchswerten in einer zweiten Tabelle. Die Stromverbrauchswerte sind sowohl mit ihrem Zeitstempel als auch mit dem entsprechenden Haushalt über eine numerische Zähler-ID verknüpft.

Für die Verwendung als Grundlage für die Energiebedarfsprofile wurden alle Nicht-Wohngebäude herausgefiltert. Haushalte, die angaben, dass sie mit Strom heizen, wurden ebenfalls ausgeschlossen, da diese einen verzerrten Stromverbrauch haben. Haushalte, die angaben, dass sie aufgrund ihrer finanziellen Situation Energie nicht frei nutzen könnten, wurden ebenfalls ausgeschlossen, da sie als Ausreißer interpretiert wurden. Schließlich wurden Haushalte ausgeschlossen, bei denen Antworten fehlten, die später für die Analyse der Cluster benötigt werden. Nach diesen Anpassungen verblieben 3.101 Haushalte in der Umfrage.

Zur Charakterisierung der Haushalte wurde die Erhebung in eine Matrix von Vektoren mit neun demografischen Eigenschaften umgewandelt, die auf den von Hierzinger et al. (2011) und Haufe & Dangschat (2017) verwendeten Eigenschaften basieren. Diese wurden als Einflussfaktoren auf den Energiekonsum von privaten Haushalten identifiziert. Tabelle 1 listet die Eigenschaften und ihre jeweils möglichen Ausprägungen auf.

Eigenschaften	Mögliche Ausprägung und deren Bedeutungen
Alter des Hauptbewohners	1 (bis 25), 2 (26-35), 3 (36-45), 4 (46-55), 5 (56-65), 6 (über 65)
Beschäftigung	1 (Arbeitend), 2 (Arbeitslos), 3 (In Rente)
Soziale Klasse (anhand NRS Social Grades)	1 (Eliten), 2 (gehobene Mittelschicht), 3 (Mittelschicht), 4 (Arbeiterschicht)
Anzahl der Bewohner	1 (1), 2 (2), 3 (mehr als 2)
Anzahl der Bewohner unter 15	0 (0), 1 (1), 2 (mehr als 1)
Besitz der Wohnung	0 (Eigentum), 1 (Gemietet)
Internetzugang	0 (nein), 1 (ja)
Anzahl der Bewohner, die tagsüber zuhause sind	0 (0), 1 (1), 2 (mehr als 1)
Wichtigkeit des Umweltschutzes für die Bewohner	0 (nicht wichtig), 1 (wichtig)

Tabelle 1: Eigenschaften von Haushalten

Anschließend wurden die Haushalte auf fehlende Stromverbrauchswerte überprüft. Etwa 1.000 Haushalte hatten fehlende Werte. Diese wurden linear interpoliert, solange nacheinander nicht mehr als 96 Werte (2 Tage) fehlten. Haushalte mit längeren Perioden fehlender Werte wurden aus dem Datensatz entfernt. Diese letzte Reduzierung führte zu 2.892 Haushalten, die schlussendlich analysiert wurden.

Um den Aspekt zu adressieren, dass der CER-Datensatz 10 Jahre alt ist und aus einem anderen Land als Deutschland stammt, wurden die Daten auf den durchschnittlichen Stromverbrauch pro Haushalt und Jahr von 2.801 kWh/a normiert, um den durchschnittlichen Verbrauch eines deutschen Haushalts im Jahr 2018 abzubilden (BDEW, 2019). In Abbildung 1 sind zwei zufällig ausgewählte Stichprobentage aus dem Datensatz dargestellt. Zur Vorbereitung der Daten für das Clustern (Deliverable 3.1.4 „Energiebedarfsprofile“) wurden sie in die vier meteorologischen Jahreszeiten sowie in Wochentage und Wochenenden aufgeteilt. Irische Feiertage wurden als Wochenendtage betrachtet, da davon ausgegangen wird, dass das Stromverhaltensverhalten durch die

fehlende Arbeit an diesen Tagen ähnlich ist. Für jeden dieser acht Zeiträume und für jeden Haushalt wurde ein durchschnittlicher Tag berechnet.

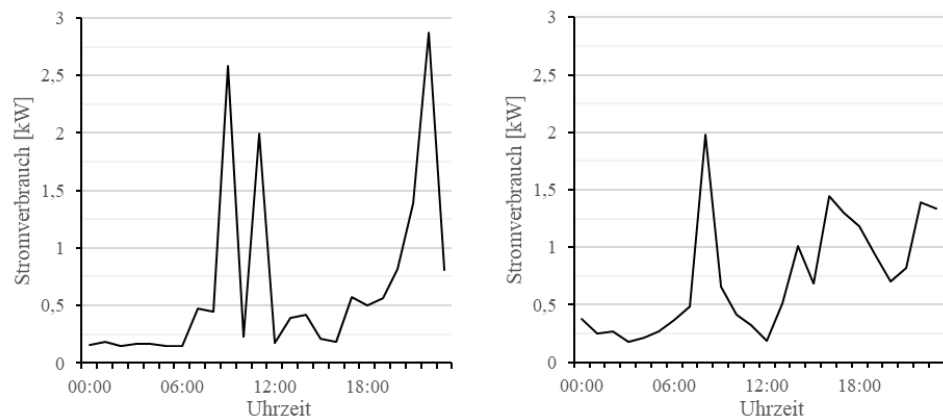


Abbildung 1: Zwei Stichprobentage aus dem gefilterten CER Datensatz

Diese 8 durchschnittlichen Tage pro Haushalt wurden dann zu einem 384-dimensionalen Vektor pro Haushalt verkettet und alle Vektoren zu einer einzigen Matrix kombiniert.

Das methodische Vorgehen zur Erstellung von Energielastprofilen im Rahmen des Deliverables 3.1.4 wird im Bericht zu Deliverable 3.1.4 im Detail in Kapitel 2 „Methodisches Vorgehen“ beschrieben.

4 Literaturverzeichnis

- [BDEW 2019] Durchschnittlicher Haushaltsstromverbrauch 2008-2018 in kWh/a. <https://www.bdew.de/service/daten-und-grafiken/durchschnittlicher-haushaltsstromverbrauch/>
- [Haufe, N., & Dangschat, J. S. 2017] Energie und soziale Ungleichheit. In Energie und soziale Ungleichheit (Issue January). <https://doi.org/10.1007/978-3-658-11723-8>
- [Hierzinger, R., Herry, M., Seisser, O., Steinacher, I., & Wolf-Eberl, S. 2011] Energy Styles Klimagerechtes Leben der Zukunft – Energy Styles als Ansatzpunkt für effiziente Policy Interventions. Neue Energien 2020, 107.
- [Regulation Commission for Energy. 2012] CER Smart Metering Project - Electricity Customer Behaviour Trial, 2009-2010 [dataset]. <http://www.ucd.ie/issda/data/commissionforenergyregulationcer/>
- [Töppel 2019] Ein Entscheidungsunterstützungssystem zur ökonomischen Bewertung von Mieterstrom auf Basis der Clusteranalyse. Internationale Konferenz Wirtschaftsinformatik (WI 2019), 4801, 1478–1492.