



# STADTQUARTIER 2050

Partner:



überlingen



Assoziierte Partner:



Deliverable D3.1.4

## Energiebedarfsprofile

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

<b>1</b>	<b>Allgemeine Einführung, Hintergrund und Ziel</b>	<b>3</b>
<b>2</b>	<b>Methodisches Vorgehen</b>	<b>4</b>
<b>3</b>	<b>Erstellung von Bewohnerprofilen</b>	<b>9</b>
<b>4</b>	<b>Erstellung von Energiebedarfsprofilen</b>	<b>13</b>
<b>5</b>	<b>Literaturverzeichnis</b>	<b>20</b>

# 1 Allgemeine Einführung, Hintergrund und Ziel

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 Leuchtmannwendungen 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 reale Strom- und Wärmeverbrauchsdaten von Bewohnern der Demonstrationsquartiere erhoben, Anwendungsmöglichkeiten der Daten geprüft, sowie der rechtliche Rahmen im Bereich Datenschutz aufgearbeitet werden. Der vorliegende Bericht widmet sich dem Deliverable 3.1.4. Das Ziel dieses Deliverables ist es, konkrete Energiebedarfsprofile u.a. auf Basis von nutzerspezifischen Merkmalen (z.B. Anzahl, Alter und/oder Beruf der Bewohner) für die Anwendung im Quartier zu erstellen. 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, wie sich im Zwischenbericht so wie in Deliverable 3.1.3 entnehmen lässt. Da noch kein ausreichender Datenpool als Voraussetzung für die nachfolgende Datenanalyse vorhanden ist, wurde der externe, irische Datenschutz Commission for Energy Regulation herangezogen; dieser wurde bereits im Rahmen der Zwischenberichtserstattung und in Deliverable 3.1.4 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. Auf Basis dessen wurden Energielastprofile erstellt.

Vorab sei zusammengefasst, dass die Untersuchung im Rahmen des vorliegenden Deliverables sechs verschiedenen Cluster von Haushalten ergab, die

sich teils deutlich in ihren Energiebedarfsprofilen und demografischen Eigenschaften unterscheiden.

## 2 Methodisches Vorgehen

Im ersten Schritt wurden die vorbereiteten Daten aus Deliverable 3.1.3 geclustert. Als Datenpunkt wurde der Vektor der zuvor generierten durchschnittlichen Tage jedes Haushalts verwendet. Als Clustering-Algorithmus wurde K-Means gewählt, welcher mit dem k-Median-Algorithmus vergleichbar ist, der von Töppel (2019) zum Clustern dieses Datensatzes verwendet wurde. Für die Anzahl der Cluster  $k$  schlägt Töppel (2019) eine Anzahl von sechs bis neun Cluster für diesen Datensatz vor. Um die optimale Anzahl von Clustern herauszufinden, wurde die Elbow-Method angewandt. In Abbildung 1 ist die resultierende Grafik abgebildet. Im Bereich von sechs bis neun Clustern schlägt die Elbow-Method ein  $k$  von sechs vor. Da das Ziel darin besteht, Cluster mit unterschiedlichen demografischen Eigenschaften zu finden, ist ein  $k$  von sechs sinnvoll; denn dies entspricht der Anzahl der verschiedenen Haushaltskategorien, wie sie in der "Wohnmatrix" von InWIS (2019) vorgeschlagen wird, und auch der Anzahl der von Töppel (2019) erstellten Cluster.

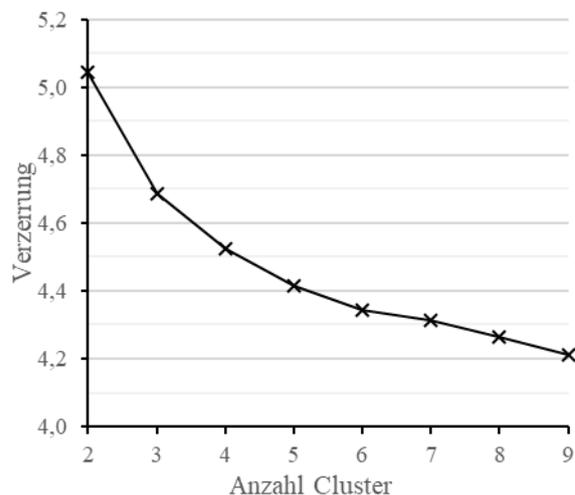


Abbildung 1: Resultierendes Diagramm der Elbow-Method

Die resultierenden Clustergrößen nach Anwendung eines k-Means-Algorithmus mit einem  $k$  von sechs sind in Abbildung 2 dargestellt. Cluster 2 ist am

größten und umfasst 31 % der Haushalte, während Cluster 1 am kleinsten ist und nur 4 % umfasst. Der durchschnittliche Stromverbrauch pro Cluster und Zeitslot ist in Abbildung 3 zu sehen. Abgesehen von den Clustern 5 und 6 hat jedes Cluster einen sich stark unterscheidenden durchschnittlichen Stromverbrauch. Cluster 1 ist das Cluster mit dem höchsten Stromverbrauch und umfasst die geringste Anzahl an Haushalten. Die drei größten Cluster 2,3 und 4, sind die Cluster, mit dem niedrigsten Stromverbrauch. Dies führt zu der Schlussfolgerung, dass sehr hohe Stromverbrauchsprofile im Rahmen der vorliegenden Ergebnisse eher selten sind.

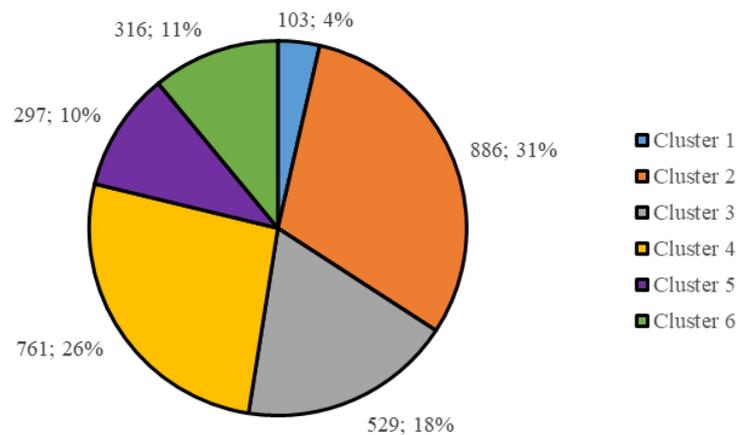


Abbildung 2: Clustergrößen

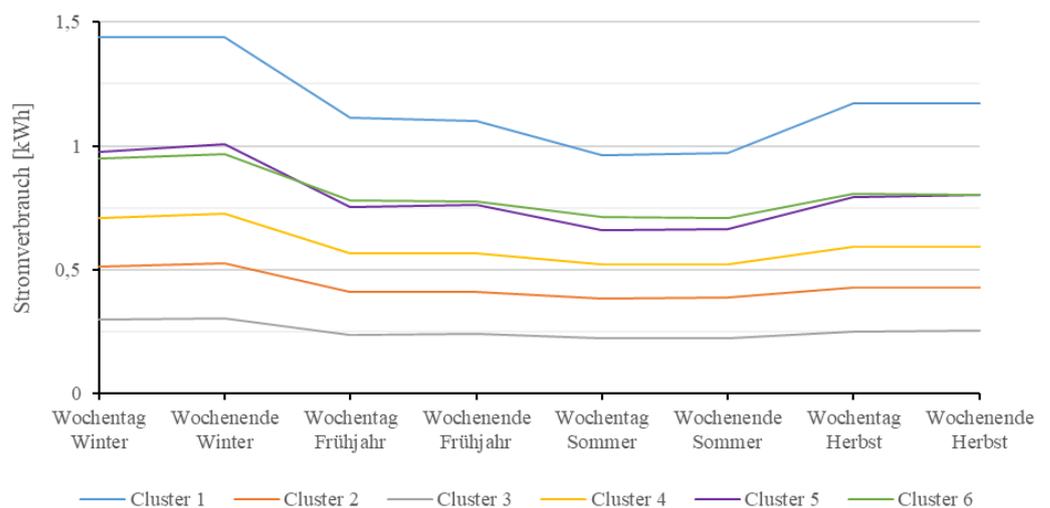


Abbildung 3: Stündlicher Durchschnittsstromverbrauch nach Cluster und Zeitslot

Zur Prognose des Energiebedarfs der simulierten Haushalte wird der SARIMA-Algorithmus verwendet, der für die Vorhersage von periodisch auftretenden Verläufen geeignet ist (Box et al., 1970, p. 327 ff.). Diese Entscheidung wurde getroffen, da das Energiebedarfsprofil eines Haushalts in der Regel Lastspitzen aufweist, die durch die Verwendung des Durchschnitts einer hohen Anzahl von Tagen geglättet würden. Durch die Verwendung eines SARIMA-Modells bleiben die Eigenschaften der Verbrauchsprofile erhalten, so dass bei der Prognose die Varianz realistischer Profile bestehen bleibt. Für jedes Cluster und jeden Zeitslot wurden die Daten aller Haushalte hintereinander gehängt. Diese Zeitreihe wurde von halbstündlichen zu stündlichen Daten aggregiert. Da die vollständigen Zeitreihen zu lang waren, um in angemessener Zeit verarbeitet werden zu können, wurde für jede Kombination aus Haushalt, Cluster und Zeitslot ein Tag nach dem Zufallsprinzip für die Berechnung von  $\gamma$  und die Ermittlung der optimalen Parameter des SARIMA-Modell ausgewählt. Die Erstellung der endgültigen SARIMA-Modelle mit den optimalen Parametern wurden mit jeweils 50 zufällig ausgewählten Tagen aus jedem Cluster und Zeitslot durchgeführt. Der Länge einer sich wiederholenden Periode wurde auf 24 gesetzt, da angenommen wurde, dass jeder Tag den gleichen Gesetzmäßigkeiten folgt. Da acht Zeitslots unterschieden wurden (jede Jahreszeit sowie Wochenende und Wochentag) und sechs Cluster existieren, wurden 48 SARIMA-Modelle erstellt. Für jedes Modell wurden 7.000 Tage für die weitere Verwendung in der Simulation vorhergesagt.

Im nächsten Schritt müssen die demografischen Kriterien der zu simulierenden Haushalte eingegeben werden. Die benötigten Parameter sind in Tabelle 1 aufgeführt. Für jeden simulierten Haushalt  $h$  wird die Wahrscheinlichkeit, zu einem beliebigen Cluster  $k$  zu gehören, mit Hilfe einer Variation von naiven Bayes-Klassifikatoren berechnet. Dies bedeutet, dass für jedes  $k$  gezählt wird, wie oft der genaue Satz demografischer Merkmale im Vektor  $x_h$  von  $h$  in den Umfragedaten gefunden wurde, und dann durch die Größe von  $k$  geteilt wird. Der resultierende Wahrscheinlichkeitsvektor  $P_{x_h}$  wird dann durch die Summe seiner selbst geteilt. Dies normiert die Summe der Cluster-Wahrscheinlichkeiten auf 100 %, die dann jeweils die Wahrscheinlichkeit der Zugehörigkeit von  $h$  zu den Clustern darstellen. Wenn die Summe von  $P_{x_h}$  gleich null ist, was bedeutet, dass  $x_h$  in der Umfrage überhaupt nicht vorkam, wird  $P_{x_h}$  auf die Gesamtwahrscheinlichkeit der Cluster gesetzt. Für jede Jahreszeit wurden die

Wahrscheinlichkeiten dann mit der Anzahl Tage der Jahreszeit multipliziert. Wenn aufgrund von Rundungen die Gesamtzahl der Tage nicht der Länge der Jahreszeit entsprach, wurde die Differenz zu dem Cluster mit der höchsten Wahrscheinlichkeit addiert oder von diesem abgezogen. Dieser Vorgang wird für jeden Haushaltstyp nur einmal durchgeführt, da sich diese Wahrscheinlichkeiten für eine bestimmte Gruppe von Haushalten nie ändern.

<i>Variable</i>	<i>Beschreibung</i>
$h$	Haushalt
$k$	Cluster
$x_h$	Merkmalsvektor des Haushalts $h$
$P_{x_h}$	Wahrscheinlichkeitsvektor von $x_h$ , der Angibt, mit welcher Wahrscheinlichkeit Haushalt $h$ in allen Clustern $k$ ist

Tabelle 2: Variablen zur Bestimmung der Clusterzugehörigkeit

Für jedes Cluster wurde die berechnete Anzahl von Tagen aus dem bereits generierten Satz von 7.000 generierten Tagen zufällig entnommen. Die Reihenfolge der Cluster, aus denen die Tage entnommen wurden, wurde für jeden Simulationslauf unterschiedlich randomisiert. Die ausgewählten Tage wurden dann in eine Matrix mit einer Spalte für jede Stunde des Tages und einer Zeile für jeden Tag des Jahres transformiert.

Für jeden Haushalt erhält man so ein Energiebedarfsprofil über ein Jahr. Um nun eine noch bessere Aussagekräftigkeit zu erlangen, kann man für jeden Haushalt mehrere Jahre generieren, um mithilfe einer Monte-Carlo Simulation herauszufinden, welche Auswirkungen das Verhalten der Menschen im Durchschnitt hat.

Der Ablauf der gesamten Datenaufbereitung bis zur Vorhersage des Energiebedarfs ist in Abbildung 4 zusammengefasst.

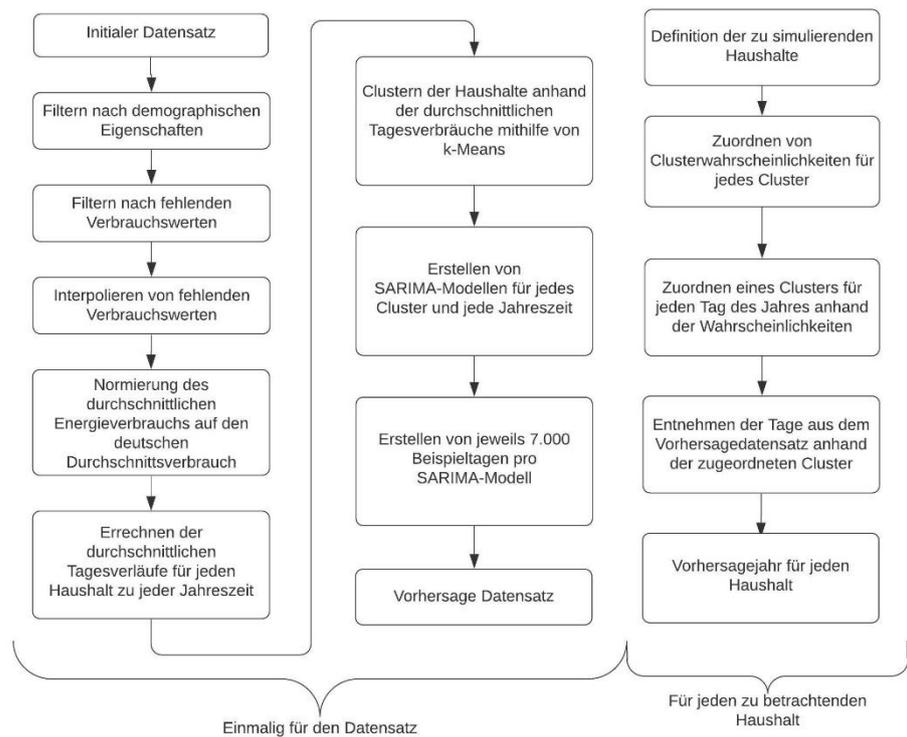


Abbildung 4: Ablauf der Datenaufbereitung und -vorhersage

### 3 Erstellung von Bewohnerprofilen

In diesem Kapitel werden die entstandenen Cluster hinsichtlich ihrer demografischen Eigenschaften beschrieben. Die durchschnittlichen Ausprägungen jedes Clusters werden dabei mit den durchschnittlichen Ausprägungen des gesamten gefilterten Datensatzes verglichen. Eine Zusammenfassung der demografischen Merkmale nach Cluster finden sich in Tabelle 2.

Cluster 1 beinhaltet viele große Familien mittleren Alters aus den unteren und mittleren Schichten mit mehr als einer Person unter 15 Jahren. Für gewöhnlich sind diese Eigentümer der Wohnung, in der sie leben, und haben eine durchschnittliche Einstellung zum Umweltschutz. Die Eltern sind für gewöhnlich arbeitend. Diesem Cluster wird der Titel „Großfamilien“ gegeben.



Abbildung 5: Demografische Eigenschaften Cluster 1 „Großfamilien“

Cluster 2 beinhaltet viele Ehepaare in der Altersgruppe über 65. Viele von ihnen sind Rentner. In den meisten dieser Haushalte sind beide Personen tagsüber zuhause, und es wohnen dort keine Kinder unter 15. Der Durchschnitt der Haushalte gehört der oberen Mittelschicht an. Überdurchschnittlich viele Haushalte haben keinen Internetanschluss. Diesem Cluster wird der Titel „Ehepaare höheren Alters“ gegeben.



Abbildung 6: Demografische Eigenschaften Cluster 2 „Ehepaare höheren Alters“

Cluster 3 beinhaltet viele Alleinstehende und Paare aus der Mittelschicht. Es sind viele Haushalte mit einem Alter des Hauptbewohners unter 35 oder über 55 vertreten. Personen unter 15 Jahren sind selten Teil der Haushalte. Überdurchschnittlich viele Haushalte haben keinen Internetanschluss. Diesem Cluster wird der Titel „Alleinstehende und Paare“ gegeben.



Abbildung 7: Demografische Eigenschaften Cluster 3 „Alleinstehende und Paare“

Cluster 4 beinhaltet viele jüngere Familien mit einem Alter des Hauptbewohners von 36 bis 45 Jahren. Dieses Cluster orientiert sich im Durchschnitt sehr stark an dem Gesamtdurchschnitt der Umfrage. Diesem Cluster wird der Titel „Familien mit Kindern Typ 1“ gegeben.

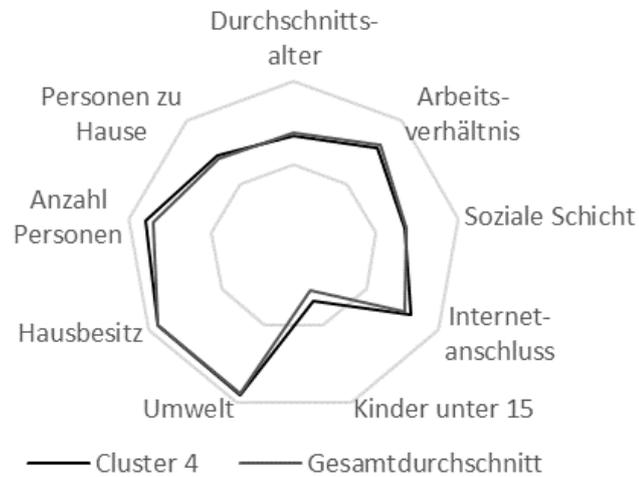


Abbildung 8: Demografische Eigenschaften Cluster 4 „Familien mit Kindern Typ 1“

Cluster 5 beinhaltet viele Familien mit einem Alter des Hauptbewohners von 46 bis 55 Jahren aus den unteren und mittleren Schichten. Am häufigsten kamen Haushalte vor, in denen mehr als eine Person unter 15 lebten. Viele Haushalte gaben an, dass tagsüber niemand zuhause ist und überdurchschnittlich viele haben Zugang zu Internet. Diesem Cluster wird der Titel „Familien mit Kindern Typ 2“ gegeben.

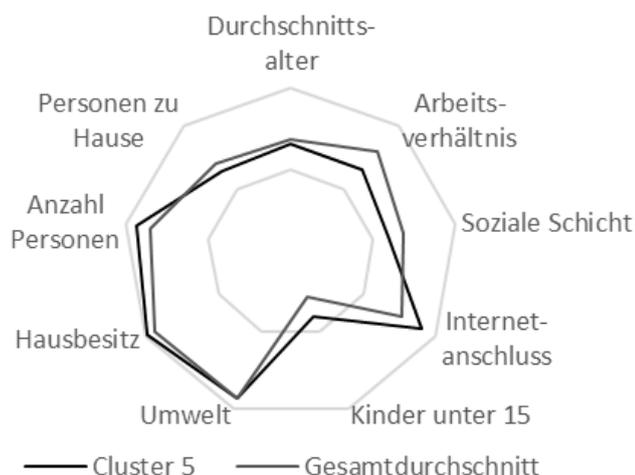


Abbildung 9: Demografische Eigenschaften Cluster 5 „Familien mit Kindern Typ 2“

Cluster 6 beinhaltet viele Familien aus der unteren Mittelschicht aller Altersgruppen mit oft mehr als einer Person unter 15 Jahren. Häufig wurde angegeben, dass mehr als eine Person tagsüber zuhause ist. Der Hauptbewohner ist für gewöhnlich arbeitend. Diesem Cluster wird der Titel „Einkommenschwache Familien“ gegeben.



Abbildung 10: Demografische Eigenschaften Cluster 6 „Einkommenschwache Familien“

Eigenschaft	1	2	3	4	5	6
Alter	Viele zwischen 36 und 55	Viele über 65	Viele unter 35 oder über 55	Viele zwischen 36 und 45	Viele zwischen 46 und 55	Gleichmäßige Verteilung
Beschäftigungsstatus	Größtenteils arbeitend	Viele in Rente	Viele in Rente	Gleichmäßige Verteilung	Größtenteils arbeitend	Größtenteils arbeitend
Soziale Schicht	Unterschicht und Mittelschicht	Gehobene Mittelschicht	Mittelschicht	Gleichmäßige Verteilung	Oberschicht oder gehobener Mittelschicht	Unterschicht
Anzahl von Bewohnern	Oft mehr als 2 Personen	Oft 2 Personen	1-2 Personen	Oft mehr als 2 Personen	Oft mehr als 2 Personen	Oft mehr als 2 Personen
Anzahl Kinder unter 15	Oft mehr als 1	Oft keine	Oft keine	Oft 1 oder mehr	Oft 1 oder mehr	Oft mehr als 1
Besitz der Wohnung	Viele Eigenheime	Gleichmäßige Verteilung	Viele Eigenheime	Gleichmäßige Verteilung	Viele zur Miete	Gleichmäßige Verteilung
Internetzugang	Viele mit Internetzugang	Über-durchschnittlich viele ohne Internet	Viele mit Internetzugang	Viele mit Internetzugang	Fast alle mit Internetzugang	Über-durchschnittlich viele ohne Internet
Anzahl Personen tagsüber zuhause	Gleichmäßige Verteilung	Oft mehr als 1 Person	Oft 0 oder 1 Person	Oft mehr als 1 Person	Oft 0 Personen	Oft mehr als 1 Person
Einstellung zu Umweltschutz	Durchschnittlich	Durchschnittlich	Viele für Umweltschutz	Durchschnittlich	Durchschnittlich	Durchschnittlich

Tabelle 3: Zusammenfassung der demografischen Eigenschaften nach Cluster

## 4 Erstellung von Energiebedarfsprofilen

Im Folgenden werden die durchschnittlichen, täglichen Energiebedarfsverläufe für die sechs Cluster aufgezeigt. Dabei wird jeweils ein Verlauf für einen Wochentag und einen Wochenendtag für Sommer und Winter verglichen.

Cluster 1 „Großfamilien“ zeigt den größten Gesamtenergiebedarf im Vergleich mit den anderen Clustern. Im Winter ist der Energiebedarf zu jeder Zeit höher als im Sommer. Sowohl an Wochentagen als auch an Wochenendtagen ist der Energiebedarf nachts und in den frühen Morgenstunden am geringsten, während der Verbrauch abends am höchsten ist. Im Winter fällt der Peak am Abend deutlich höher aus als im Sommer.

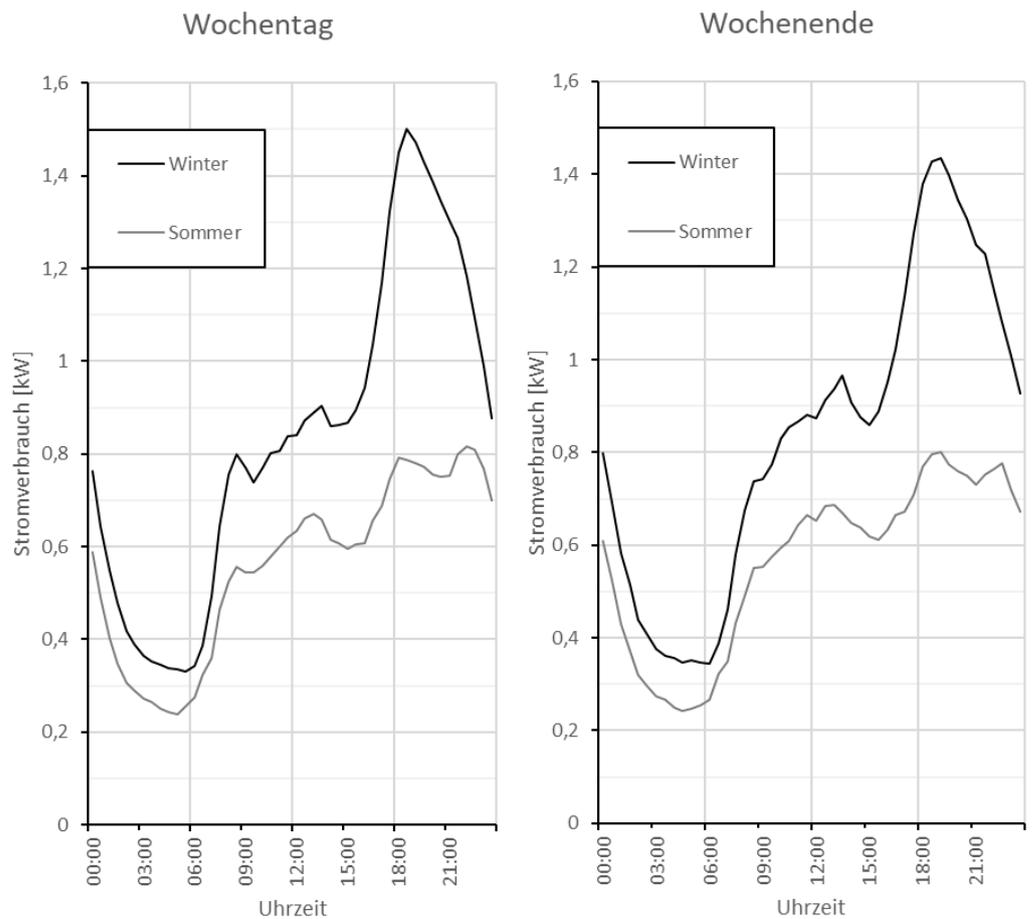


Abbildung 11: Energiebedarfsprofil Cluster 1

Cluster 2 „Ehepaare höheren Alters“ zeigt einen optisch sehr ähnlichen Verlauf des Energiebedarfs wie das Cluster 1, jedoch wird zu jeder Zeit deutlich weniger Strom benötigt. Im Sommer ist zudem der Unterschied zwischen dem Energiebedarf am Vormittag und am Abend nicht so ausgeprägt.

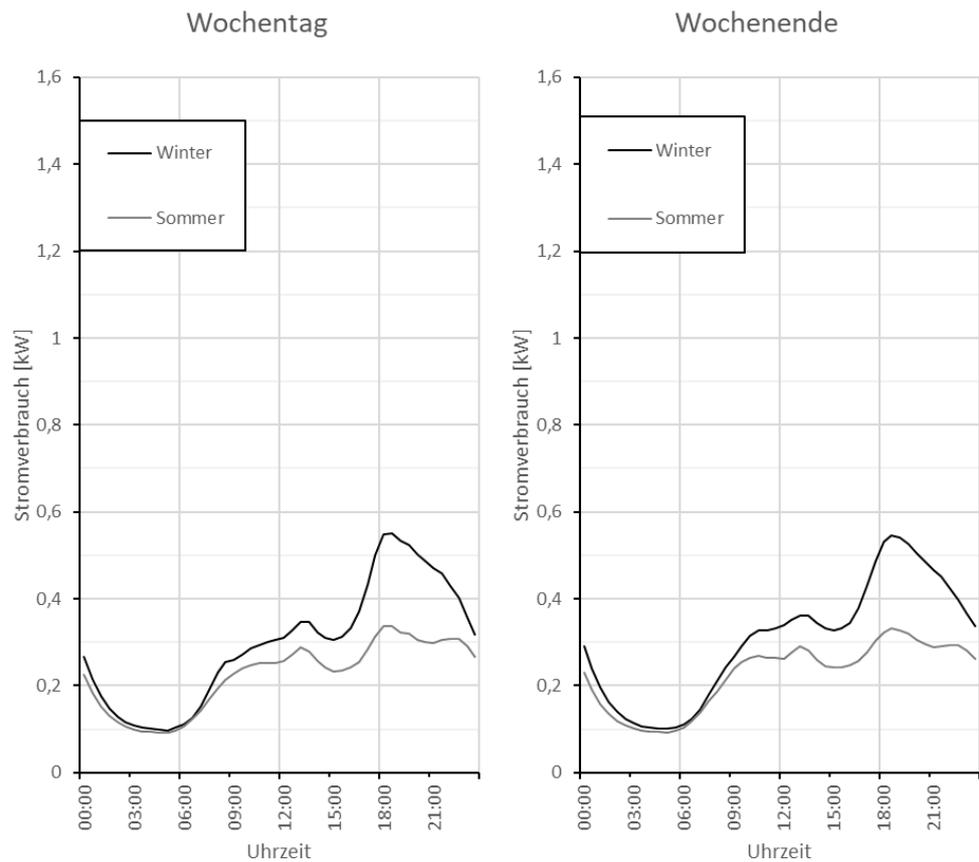


Abbildung 12: Energiebedarfsprofil Cluster 2

Cluster 3 „Alleinstehende und Paare“ hat den niedrigsten Energiebedarf aller Cluster zu jeder Zeit. Im Vergleich zu Cluster 2 steigt der Energiebedarf gerade am Wochenende erst später am Tag an. Ähnlich wie in Cluster 2 ist der Unterschied zwischen vormittags und nachmittags gerade im Sommer nicht so stark ausgeprägt wie in Cluster 1. Neben dem geringeren Gesamtenergiebedarf ist auch das Verhältnis zwischen Nacht und Tag nicht so groß wie in Cluster 1 und 2.

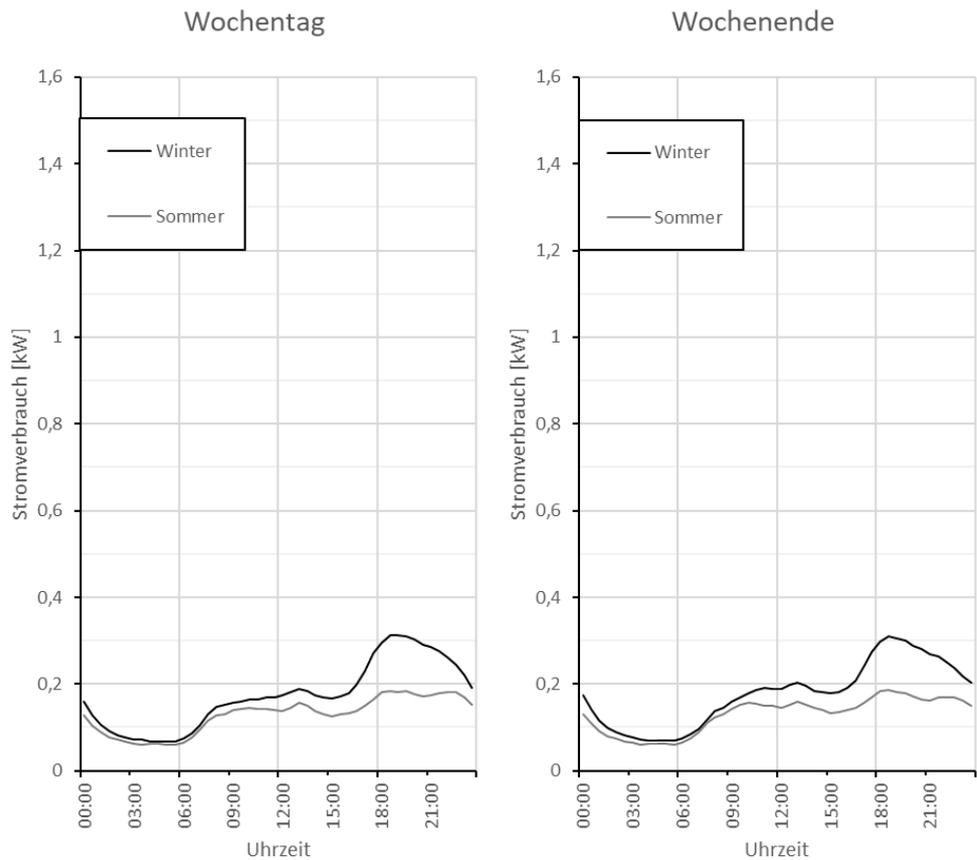


Abbildung 13: Energiebedarfsprofil Cluster 3

Cluster 4 „Familien mit Kindern Typ 1“ zeigt im Vergleich zu den Clustern 1 bis 3 gerade bei den Wochentagen einen späteren Anstieg des Energiebedarfs am Vormittag. Der Verbrauch liegt etwas höher als bei Cluster 2. Charakteristisch ist ebenfalls, dass der Energiebedarf am Abend gerade unter der Woche recht früh schon stark abfällt.

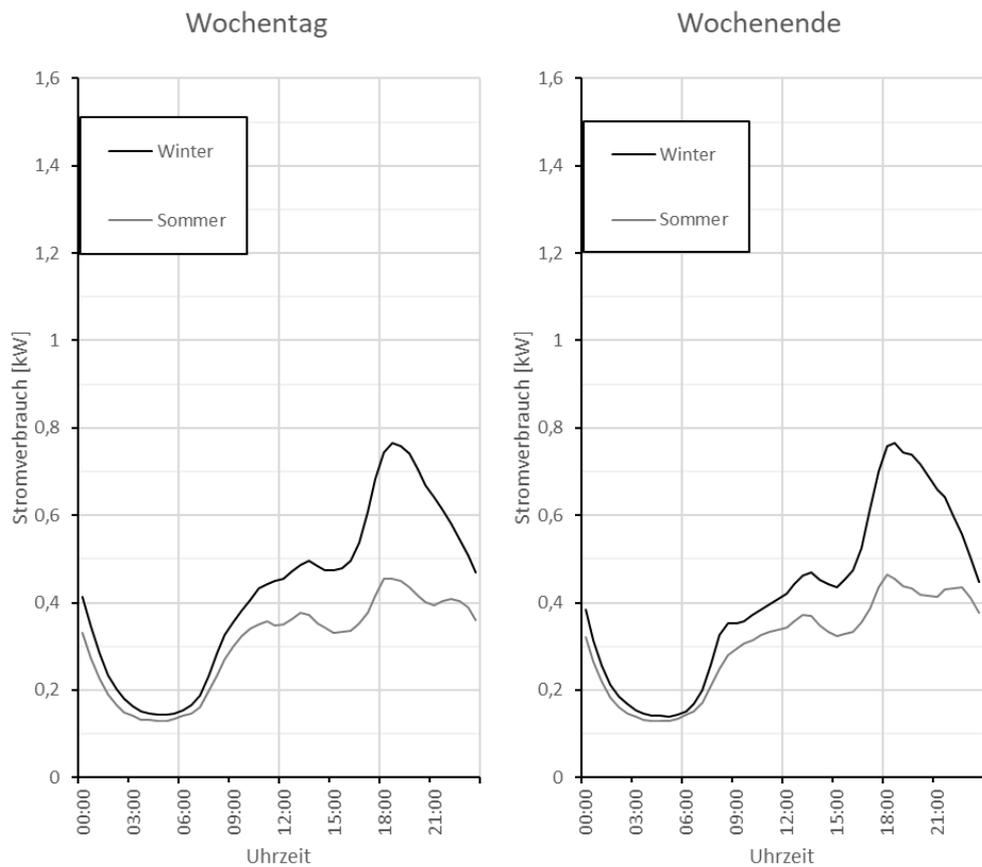


Abbildung 14: Energiebedarfsprofil Cluster 4

Cluster 5 „Familien mit Kindern Typ 2“ zeigt, ähnlich wie Cluster 4, einen späten Anstieg des Energiebedarfs in den Morgenstunden. Dieser ist vormittags jedoch sehr gering im Vergleich zu dem Verbrauch in den Abendstunden. Der höchste Energiebedarf am Abend ist etwas höher als in Cluster 4. Im Gegensatz zu allen anderen Clustern ist ein vergleichbarer Anstieg auch im Sommer zu verzeichnen, sodass der Energiebedarf morgens deutlich unter dem abendlichen Energiebedarf liegt.

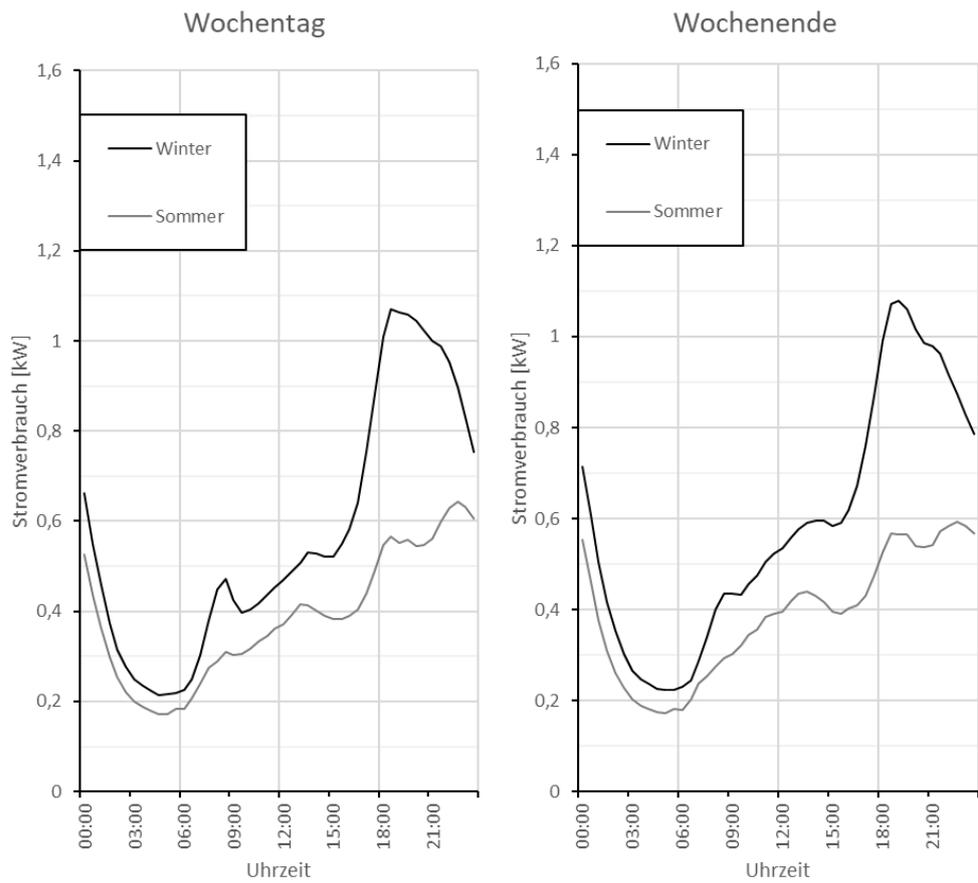


Abbildung 15: Energiebedarfsprofil Cluster 5

Cluster 6 „Einkommenschwache Familien“ zeigt einen ähnlichen Gesamtenergiebedarf wie Cluster 5, jedoch steigt der Energiebedarf in den Morgenstunden früher und deutlich stärker an. Der Energiebedarf im Sommer zeigt tagsüber eine geringere Steigung. Sowohl im Sommer als auch im Winter steigt der Energiebedarf morgens früh an und fällt abends wieder schnell ab.

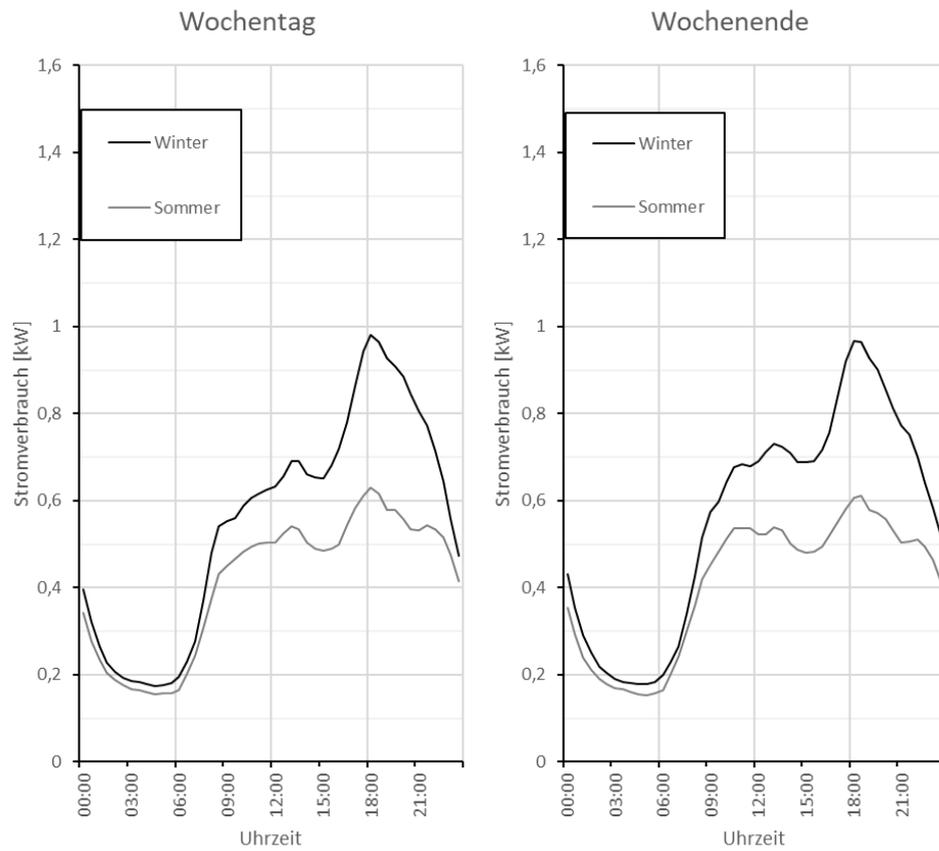


Abbildung 16: Energiebedarfsprofil Cluster 6

Aufbauend auf diesen Ergebnissen werden in D.3.1.5. Handlungsempfehlungen abgeleitet.

## 5 Literaturverzeichnis

- [BDEW 2019] Durchschnittlicher Haushaltsstromverbrauch 2008-2018 in kWh/a. <https://www.bdew.de/service/daten-und-grafiken/durchschnittlicher-haushaltsstromverbrauch/>
- [Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. 1970] Time Series Analysis: Forecasting and Control (Third Edit). Prentice-Hall International, Inc.
- [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.