



Kernkompetenzzentrum
Finanz- & Informationsmanagement



Projektgruppe
Wirtschaftsinformatik

Ein Entscheidungsunterstützungssystem zur ökonomischen Bewertung von Mieterstrom auf Basis der Clusteranalyse

von

Jannick Töppel

März 2019

Presented at: 14th International Conference on Wirtschaftsinformatik (WI),
Siegen, Germany, 2019

WI-945

Universität Augsburg, D-86135 Augsburg
Besucher: Universitätsstr. 12, 86159 Augsburg
Telefon: +49 821 598-4801 (Fax: -4899)

Universität Bayreuth, D-95440 Bayreuth
Besucher: Wittelsbacherring 10, 95444 Bayreuth
Telefon: +49 921 55-4710 (Fax: -844710)



Universität
Augsburg
University



UNIVERSITÄT
BAYREUTH



Ein Entscheidungsunterstützungssystem zur ökonomischen Bewertung von Mieterstrom auf Basis der Clusteranalyse

Jannick Töppel¹

¹ Universität Augsburg, Kernkompetenzzentrum Finanz- & Informationsmanagement,
Augsburg, Deutschland
jannick.toepfel@fim-rc.de

Abstract. Für den Erfolg der Energiewende spielt die dezentrale Stromerzeugung eine entscheidende Rolle. Aus diesem Grund wurde das Geschäftsmodell Mieterstrom entwickelt, welches sich über die Erzeugung und Lieferung von Strom in direktem räumlichen Zusammenhang definiert. Dabei soll durch den direkten Verkauf von Strom an Mieter ein höherer Gewinn erzielt werden, im Vergleich zur klassischen Stromeinspeisung ins Netz. Zur Förderung von Mieterstrom in Deutschland wurde 2017 von politischer Seite ein umfassendes Förderprogramm beschlossen und somit die Rahmenbedingungen deutlich verbessert. Basierend auf Smart Meter Daten wird in diesem Beitrag deshalb ein Entscheidungsunterstützungsmodell zur Investitionsbewertung von Mieterstrommodellen entwickelt und evaluiert. Dafür wird in einem ersten Schritt eine Clusteranalyse durchgeführt, um anschließend auf Basis der durchschnittlichen Stromverbrauchsprofile der Cluster sowie bedingter Wahrscheinlichkeiten die Rentabilität eines Mieterstrommodells vorherzusagen. Die in einer Fallstudie evaluierten Investitionen weisen eine Amortisationszeit von 8 bis 14 Jahren sowie eine CO₂-Einsparung von über 60% auf.

Keywords: Mieterstrom, Investitionsbewertung, Clusteranalyse, Smart Meter Daten.

1 Einleitung

Vor dem Hintergrund des voranschreitenden Klimawandels sowie knapper werdender Ressourcen, wurden in den vergangenen Jahren weltweit erste Anstrengungen unternommen, um diesen negativen Entwicklungen entgegenzuwirken [1]. In Deutschland spielt zur Erreichung der internationalen Klimaziele die sogenannte Energiewende eine zentrale Rolle, wovon sich der kurzfristige Ausstieg aus der Atomenergie und der mittelfristige Ausstieg aus der Kohleenergie verbirgt [2]. Eine wichtige Voraussetzung für das Erreichen dieser Ziele ist der Aufbau einer dezentralen Energieversorgung, für welche der Deutsche Bundesrat am 7. Juli 2017 mit dem Gesetz zur Förderung von Mieterstrom einen wichtigen Grundstein legte [3]. Mieterstrom stammt aus einer Photovoltaikanlage (PV-Anlage) oder einem Blockheizkraftwerk

(BHKW) eines Wohngebäudes und wird direkt den ansässigen Mietern zur Verfügung gestellt [4]. Überschüssiger Strom kann gegen die übliche EEG-Einspeisevergütung dem Stromnetz zugeführt werden. Diverse steuerliche Erleichterungen und Zuschüsse machen es für Vermieter nun deutlich attraktiver, den produzierten Strom möglichst vollständig an ihre Mieter zu verkaufen. Um den Mietern die Teilnahme am Mieterstrommodell interessant zu machen, wird in der Regel der Marktpreis für Strom leicht unterboten. Für den Vermieter wird dadurch der Eigenverbrauch, also der direkt im Gebäude verbrauchte Anteil des erzeugten Stroms, zu einer entscheidenden Performancekennzahl. Um die Stromverbräuche und Stromeinspeisungen präzise messen zu können, spielen Smart Meter eine zentrale Rolle in Mieterstrommodellen. Die erfassten, hochauflösenden Messergebnisse werden insbesondere für die komplexe Abrechnung benötigt, welche in der Regel über eine unterstützende IT Infrastruktur abgewickelt wird [5].

Die durch die Gesetzesänderung geschaffene Ausgangslage hat bereits erste Wohnbaugesellschaften zu Investitionen in das Mieterstrommodell bewegt [6]. Deren Rentabilität hängt, neben der Entwicklung des Strompreises, maßgeblich vom Eigenverbrauch im Tagesverlauf ab [7]. Ein Beispiel: Wenn die meisten Mieter tagsüber während der sonnenreichen Stunden nicht zuhause sind, muss der Investor den Großteil des produzierten Photovoltaikstroms ins Netz einspeisen und profitiert daher nicht von den höheren Preisen, welche er von den Mietern verlangen könnte. Ex ante sind, in Abwesenheit von Smart Metern, die Verbrauchsprofile in der Regel nicht bekannt oder nur in aggregierter Form (Jahresverbräuche) verfügbar und können so nicht bzw. kaum zur Investitionsbewertung herangezogen werden.

Auf Basis von Smart Meter Stromdaten wurden in den vergangenen Jahren eine Vielzahl von Data Analytics Lösungen entwickelt. So wurden Haushalte in Abhängigkeit von Verbrauchsprofilen in verschiedene Cluster eingeteilt [8] und anhand des Verbrauchsprofils auf Eigenschaften, wie z.B. das Gebäudealter, die Art der Warmwasserversorgung [9] sowie den Beschäftigungsstatus oder die Wohnform [10], geschlossen. Dabei liegt der Fokus der meisten Arbeiten jedoch nicht auf der Entscheidungsunterstützung durch Informationssysteme (IS) bei der Investitionsbewertung, sondern vielmehr auf der Unterstützung von politischen Entscheidungen oder der Entwicklung neuer Dienstleistungen für Endkunden, Netzbetreiber und Stromanbieter.

Aus den vorangegangenen Überlegungen resultiert die folgende Forschungsfrage: Wie kann ein Entscheidungsunterstützungssystem basierendes auf Smart Meter Daten das Verbrauchsprofil einer Immobilie vorhersagen, um Mieterstrommodelle ökonomisch zu bewerten?

Die datenbasierte Vorhersage von Verbrauchsprofilen für Mieterstrommodelle ist ein Beitrag zur Energie Informatik, welche als Teilgebiet der IS-Forschung Informationssysteme zur Steigerung der Energieeffizienz und der Nachhaltigkeit entwickelt [11]. Um die formulierte Forschungsfrage zu beantworten, wurde der Design Science Ansatz gewählt [12]. Dieser Ansatz zeichnet sich durch die Entwicklung und anschließende Evaluation eines Artefakts aus. Hevner et al. [12] heben hervor, dass ein

entwickeltes Artefakt (Modell) eine praktische Implikation zeigen sollte. Diese Anforderung wird in Kapitel 2 durch die Einbettung des Artefakts in ein Geschäftsmodell gewährleistet, welches auf Basis einer kurzen Einführung zu Mieterstrom und Datenanalysen von Smart Meter Daten skizziert wird. Anschließend wird in Kapitel 3 ein datenbasiertes Modell zur Vorhersage von Eigenverbräuchen entwickelt, welches zur ökonomischen Bewertung von Mieterstrommodellen verwendet werden kann. In Kapitel 4 wird das entwickelte Modell anhand realer Daten evaluiert und erprobt. Abschließend werden in Kapitel 5 die Ergebnisse und Limitationen zusammengefasst sowie auf weitere Forschungsfragen eingegangen.

2 Mieterstrom und Datenanalyse

Mit sinkenden Kosten für die lokale Stromerzeugung [13] und steigenden Endkundenpreisen für Strom, wird die dezentrale Stromerzeugung und der direkte Konsum vor Ort immer attraktiver. Der wirtschaftliche Rahmen, in welchem ein Investor (Vermieter) von diesen Entwicklungen profitieren kann, ist das oben eingeführte Mieterstrommodell. Behr und Großklos [4] definieren Mieterstrom anhand vier wesentlicher Elemente: (1) Die dezentrale Erzeugung von Strom, welche entweder regenerativ oder mit hoher Effizienz in Kraft-Wärme-Kopplung erfolgt, (2) der Standort der Energieerzeugungsanlage befindet sich in einem Gebäude oder auf dem Wohngelände, (3) zwischen der Energieerzeugungsanlage und den Mietern besteht ein räumlicher Zusammenhang sowie (4) für die Belieferung der Mieter werden keine allgemeinen Versorgungsnetze verwendet. Darüber hinaus regelt der deutsche Gesetzgeber in diversen Gesetzen und Verordnungen weitere Detailpunkte des Mieterstroms. Neben dem Stromverkauf an Mieter kann der Vermieter nach wie vor den erzeugten Strom auch ins Stromnetz einspeisen und erhält dafür eine garantierte Einspeisevergütung. Diese liegt jedoch in der Regel um ca. 30 bis 50% niedriger als die Rückflüsse aus dem Direktverkauf an die Mieter. Aus diesem Grund muss für die ökonomische Rentabilität eine entsprechende Strommenge von den Mietern abgenommen werden, um die Mehrkosten für Smart Meter und Abrechnungssysteme zu kompensieren. Diese Mehrkosten sowie weitere Aspekte der Implementierung wurden daher bereits in ersten Studien für den deutschen Markt zusammengefasst [5], [14]. In Abbildung 1 ist ein typisches Verbrauchsprofil einer Immobilie mit vier Haushalten abgebildet, welche über eine Photovoltaikanlage (10kWp) sowie über eine Brennstoffzelle (3kW) verfügen. Es ist zu erkennen, dass nachts die Brennstoffzelle die Grundlast des Gebäudes vollständig deckt und die Photovoltaikanlage von ca. 4 bis 20 Uhr Strom produziert, allerdings nicht genug um den Peak am Abend abzudecken. Die dargestellten Verläufe sind hochgradig stochastisch, da sie von diversen Faktoren abhängen. So ist die Stromproduktion einer Photovoltaikanlage stark von Wettereinflüssen abhängig [15] und das Verbrauchsprofil vom Nutzerverhalten oder der Anwesenheit der Mieter [16]. Für ein Mieterstrommodell ergibt sich folglich eine doppelte Unsicherheit, bestehend aus der produzierten Strommenge sowie dem Strombedarf der Mieter.

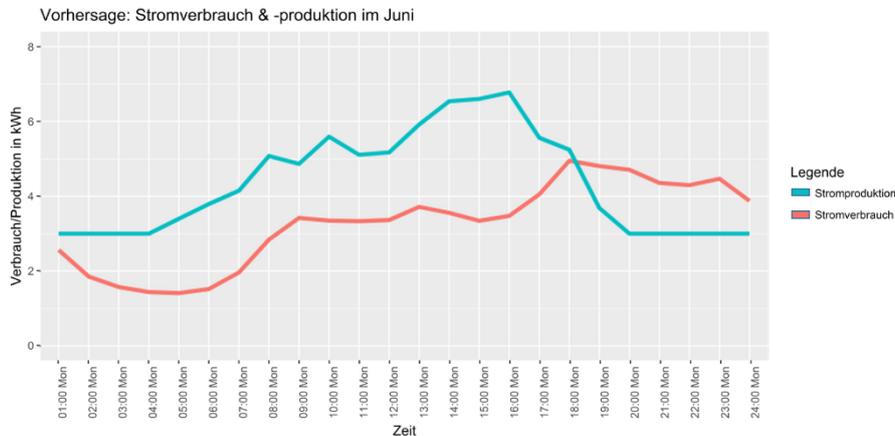


Abbildung 1. Beispielhafter Stromverbrauch und -produktion (3kW Brennstoffzelle und 10kWp Photovoltaikanlage) eines Mehrfamilienhauses (vier Einheiten)

Die für Mieterstrom eingesetzten Stromerzeugungstechniken Photovoltaik (PV) und BHKW weisen unterschiedliche Erzeugungsprofile auf. Die PV-Anlage besitzt sowohl einen Tages- als auch einen Jahresgang mit höchster Stromerzeugung zu der Mittagszeit und im Sommer. Dabei wird die Leistung einer PV-Anlage maßgeblich von der Umgebungstemperatur sowie der Globalstrahlung beeinflusst, und kann so zwischen einzelnen Tagen deutlich schwanken [15]. In der wissenschaftlichen Literatur ist daher in den vergangenen Jahren eine Vielzahl von Publikationen zur Vorhersage von Erzeugerprofilen bei PV-Anlagen entstanden. In der Regel werden dabei hybride Modelle verwendet, sodass auf Basis von Ingenieursgleichungen [17] lediglich die Eingangsvariablen der Gleichungen wie Temperatur- und Globalstrahlungsverläufe u.a. durch Neuronale Netze [18], Markov Modelle [19] oder Zeitreihen [20] stochastisch modelliert werden. Ein weiterer Teil der Literatur beschäftigt sich darüber hinaus mit Randparametern, wie zum Beispiel der Verschattung von PV-Anlagen [21] oder dem Anlagendesign [22]. Ein BHKW liefert dahingegen im durchgängigen Betrieb eine nahezu konstante Strommenge. Sollte die zusätzlich anfallende Wärmemenge im Sommer nicht benötigt werden, kann das Abschalten der Anlage rentabel sein. Die Berechnung des Eigenverbrauchs hat in den letzten Jahren durch den vermehrten Einsatz von Stromspeichern an Relevanz gewonnen [15], wobei zur Berechnung von Eigenverbräuchen in der Regel einzelne Objekte betrachtet oder Standardprofile verwendet werden.

Großklos [7] misst daher im Zuge der Investitionsbewertung von Mieterstrommodellen neben der Energieerzeugungsvorhersage auch dem Verbrauch und der tatsächlichen Stromabnahme eine wichtige Rolle bei. Da Stromverbrauchsprofile sowohl große Schwankungen im Tagesverlauf aufweisen als auch einem saisonalen Jahresverlauf unterliegen, sind für die Berechnung des Eigenverbrauchs möglichst hochauflösende Vorhersagen notwendig. Die verschiedenen Einflüsse auf Stromverbrauchsprofile untersuchen Sodenkamp et al. [9], indem sie mit Machine Learning Algorithmen und Verbrauchsdaten aus der Schweiz

versuchen, anhand des Stromverbrauchs auf Eigenschaften eines Haushalts zu schließen (z.B. Alter des Hauses, die Verwendung von Wärmepumpen oder elektrischen Kochgeräten). In dieselbe Richtung arbeiten auch Beckel et al. [23], indem sie für einen irischen Datensatz diverse Eigenschaften zu Haushaltsbewohnern sowie des Gebäudes mit einer Trefferquote von zum Teil 70 bis 80% bestimmen. Für denselben Datensatz untersuchen McLouglin et al. [16] den Einfluss auf den Stromverbrauch von Faktoren, wie die Anzahl der Schlafzimmer, Zusammensetzung des Haushaltes oder den Wohnungstyp mit Hilfe einer multiplen linearen Regression. In einer späteren Arbeit wenden die Autoren dann statistische Clustermethoden wie k-Means, k-Medoids und Self Organising Maps an, um anhand des Stromverbrauches die Haushalte in verschiedene Klassen aufzuteilen [8]. Insgesamt identifizieren sie 10 Cluster, die jeweils charakteristische Verbrauchsverläufe aufweisen. Zu der gleichen Anzahl an Clustern kommen auch Haben et al. [24], unter Verwendung eines sogenannten Finite Mixture Clustering. Einen etwas anderen Weg schlagen Wang et al. [25] ein, indem sie erst die Verbrauchsdaten jedes Haushalts in ein Markov-Modell überführen und anschließend die Haushalte nach deren Wahrscheinlichkeitsdichten clustern, um so bestimmte Verhaltensweisen zu identifizieren. Abschließend schätzen Al-Wakeel et al. [26] fehlende Stromverbrauchswerte anhand des k-Means Algorithmus, indem sie verschiedene Distanzmaße von den Ausgangsdaten zu den Clusterzentren verwenden.

Die vorgestellten statistischen Methoden werden in der Praxis bei der ökonomischen Bewertung von Mieterstrom kaum angewandt, sodass bis heute Standardprofile oder Referenzprojekte zur Bewertung von Mieterstrommodellen herangezogen werden [7]. Ziel dieser Arbeit ist es daher, ein datenbasiertes Modell zur Vorhersage des haushaltsspezifischen Stromverbrauchprofils auf Basis weniger Faktoren, wie zum Beispiel der Anzahl der Zimmer, dem Beschäftigungsverhältnis des Mieters und der Haushaltsgröße zu entwickeln. Dieses Stromverbrauchsprofil kann anschließend zur Berechnung des Eigenverbrauchs verwendet werden, der wiederum als Grundlage für die Investitionsbewertung von Mieterstrommodellen dient.

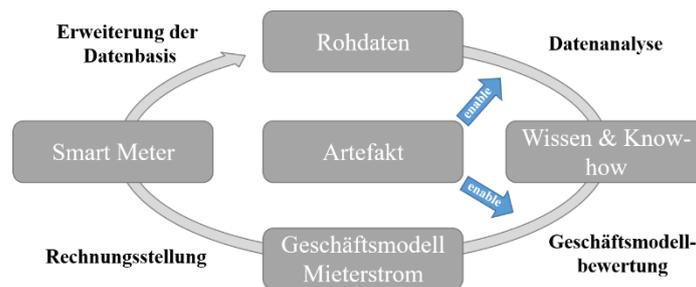


Abbildung 2. Datenkreislauf eines Geschäftsmodells zur Beratung bei Mieterstrom

Abbildung 2 zeigt das Artefakt dieses Beitrags entlang eines Datenkreislaufs für ein Geschäftsmodell zur Verwaltung und Beratung von Mieterstrom. Hierzu wird, ausgehend von Rohdaten zu Verbrauch und Haushaltseigenschaften, Know-how auf Grundlage von Datenanalysen aufgebaut. Dies ermöglicht im zweiten Schritt, in Form

eines Bewertungsalgorithmus, die ökonomische Bewertung des Mieterstrommodells. Nach Umsetzung des Mieterstroms werden dann insbesondere zu Abrechnungszwecken weitere Daten über Smart Meter gesammelt, welche den Datenpool erweitern und so den gesamten Kreislauf kontinuierlich verbessern.

3 Modellteil

Im Folgenden werden die notwendigen Gleichungen zur Modellierung eines Mieterstrommodells auf Basis von PV und BHKW eingeführt. Um die ökonomische Bewertung eines Mieterstrommodells durchführen zu können, ist eine möglichst exakte Schätzung des Eigenverbrauchs notwendig. Auf dieser Basis lässt sich für ein Mieterstrommodell die folgende Gleichung des Cashflows im Jahr $t \in \{1, \dots, T\}$ aufstellen, die Steuereffekte nicht berücksichtigt:

$$CF_t = E_t^G \cdot \varepsilon_t \cdot P_t^{MS} + E_t^G \cdot (1 - \varepsilon_t) \cdot P_t^{Netz} - K_t, \quad (1)$$

wobei $E_t^G \in \mathbb{R}^+$ die erzeugte Strommenge in kWh bezeichnet, $\varepsilon_t \in [0,1]$ den Eigenverbrauchsfaktor und P_t^{MS} den erzielten Preis pro kWh durch den Verkauf von Mieterstrom sowie P_t^{Netz} dementsprechend die Einspeisevergütung für die Stromabgabe ins Netz. Die Kosten K_t zum Zeitpunkt t setzen sich aus den Verwaltungskosten, den Abrechnungskosten (Smart Meter und IT-Infrastruktur) sowie ggf. anfallenden Kosten für Energieträger zusammen. Zur Vereinfachung werden die Erlöse aus der Wärmeerzeugung eines BHKW nicht berücksichtigt, können aber über verminderte Kosten modelliert werden. Hinter Gleichung (1) steckt die Annahme, dass zusätzlich benötigter Strom aus dem Stromnetz gewinnneutral an die Mieter weiterverkauft wird. Der Eigenverbrauchsfaktor ε_t lässt sich als Bruch von erzeugter Strommenge E_t^G und der von den Mietern abgenommenen Strommenge E_t^{MS} darstellen:

$$\varepsilon_t = \frac{E_t^{MS}}{E_t^G}. \quad (2)$$

Für eine gegebene Partition, welche einen Tag anhand von N Messpunkten in $N - 1$ gleich große Intervalle teilt, kann die abgenommene Mieterstrommenge E_t^{MS} (auf Jahresbasis) wie folgt geschätzt werden:

$$E_t^{MS} \approx \sum_{i=1}^{365} \sum_{j=1}^N \min(e_{ij}, v_{ij}), \quad (3)$$

mit e_{ij} als die erzeugte Strommenge in kWh am Tag i zwischen Messpunkt j und $j - 1$ sowie v_{ij} als der dazugehörige Stromverbrauch der Mieter. Dabei wird innerhalb der Summenzeichen der limitierende Faktor für den Mieterstrom berechnet, also die verfügbare Strommenge oder der Strombedarf der Mieter und dieser anschließend für alle Intervalle aufaddiert. Die gesamte erzeugte Strommenge E_t^G kann analog zu Gleichung (3) berechnet werden, indem lediglich die e_{ij} aufsummiert werden. Als nächster Schritt wird nun eine Methode zur Modellierung der erzeugten sowie

verbrauchten Strommengen e_{ij} und v_{ij} eingeführt. Für die Modellierung der erzeugten Strommenge aus einer PV-Anlage e_{ij}^{PV} wird die Osterwald Methode [27] verwendet, welche trotz ihrer Einfachheit eine hohe Prognosegüte aufweist [17]. Demnach kann die Leistung L_{ij}^{PV} einer PV-Anlage über die folgende Gleichung definiert werden:

$$L_{ij}^{PV} = L_{ij}^{max}(G_{ij}, T_{ij}^c) \cdot (1 - L_V) \cdot (1 - L_R) \cdot (1 - L_T) \cdot (1 - L_S). \quad (4)$$

$L_V \in [0,1]$ bezeichnet die Verluste durch Verschattung, $L_R \in [0,1]$ definiert die Verluste durch Reflexion des Lichtes, $L_T \in [0,1]$ definiert die Verluste in den Leitungen des PV-Moduls beim Stromtransport sowie $L_S \in [0,1]$ die Systemverluste u.a. bei der Stromumwandlung. Die maximale Leistung $L_{ij}^{max}(G_{ij}, T_{ij}^c)$ in Abhängigkeit der Globalstrahlung G_{ij} und der Zelltemperatur der PV-Anlage T_{ij}^c ist definiert als:

$$L_{ij}^{max}(G_{ij}, T_{ij}^c) = L^{peak} \cdot G_{ij} \cdot [1 + 0,0035 \cdot (T_{ij}^c - T^*)]/G^*, \quad (5)$$

mit L^{peak} als die Gesamtleistung der PV-Anlage unter genormten Bedingungen bei einer Temperatur von $T^* = 25^\circ C$ und $G^* = 1000 W/m^2$. Die Zelltemperatur T_{ij}^c einer PV-Anlage kann auf Basis der Außentemperatur T_{ij}^A sowie der Globalstrahlung G_{ij} wie folgt geschätzt werden [28]:

$$T_{ij}^c = T_{ij}^A + (T^{PV} - 20) + G_{ij}/800, \quad (6)$$

wobei T^{PV} die ausgelegte Zelltemperatur im Betriebsmodus ist. An dieser Stelle sei ergänzend vermerkt, dass die einfallende Globalstrahlung G_{ij} maßgeblich von der Ausrichtung der PV-Anlage und des Sonnenwinkels abhängt. Nach Reikard [20] sowie Fang und Lahdelma [29] können die Globalstrahlung G_{ij} und die Außentemperatur T_{ij}^A aufgrund ihrer statistischen Eigenschaft adäquat über Seasonal Autoregressive Integrated Moving Average (SARIMA) Modelle quantifiziert werden. Abschließend kann die Leistung L^{BHKW} eines BHKW vereinfachend als konstant angenommen werden, sodass sich die folgende Gleichung für die produzierte Strommenge e_{ij} ergibt:

$$e_{ij} = (L_{ij}^{PV} + L^{BHKW}) \cdot 24h/N. \quad (7)$$

Hinter Gleichung (7) steckt die Annahme, dass innerhalb eines Zeitschrittes die Leistungen der PV-Anlage wie auch des BHKWs konstant sind. Als zweiter Schritt wird nun die Modellierung des Stromverbrauchsprofils v_{ij} der am Mieterstrommodell beteiligten $K \in \mathbb{N}$ Haushalte eingeführt. Der Gesamtverbrauch v_{ij} lässt sich als die Summe der Einzelverbräuche der Haushalte darstellen:

$$v_{ij} = \sum_{k=1}^K v_{kij}, \quad (8)$$

wobei v_{kij} der Stromverbrauch des k -ten Haushaltes am i -ten Tag zwischen Messpunkt j und $j - 1$ bezeichnet. Über jeden Haushalt sei ferner ein Merkmalsvektor $x \in \mathbb{N}^D$ bekannt, welcher aus $D \in \mathbb{N}$ einzelnen Merkmalen besteht. Zusätzlich sei bekannt, dass für Haushaltsstromverbräuche $Q \in \mathbb{N}$ Standardstromverbrauchsprofile existieren,

anhand welcher sich Haushalte in verschiedene Cluster aufteilen lassen. Für jedes Cluster $q = 1, \dots, Q$ seien f_1, \dots, f_Q die Dichtefunktionen von $\mathbb{R}^D \rightarrow \mathbb{R}$, welche die Verteilung und Häufigkeit des Merkmalsvektors x in jedem Cluster charakterisieren. Zusätzlich seien die allgemeinen Wahrscheinlichkeiten π_1, \dots, π_Q der Cluster bekannt, welche die Häufigkeiten der einzelnen Cluster charakterisieren und deshalb auch $\sum_{q=1}^Q \pi_q = 1$ gelten muss. Nach dem Satz von Bayes gilt dann für jeden Haushalt mit Merkmalsvektor x :

$$f_x(q) := P(\text{Haushalt gehört zu Cluster } q | X = x) = \frac{\pi_q f_q(x)}{\sum_{l=1}^Q \pi_l f_l(x)}. \quad (9)$$

Mit Hilfe von Gleichung (9) kann somit die Wahrscheinlichkeit, dass ein Haushalt zu einem Cluster q gehört, anhand von verschiedenen Dichtefunktionen und Wahrscheinlichkeiten berechnet werden, die sich nach einer Clusteranalyse bestimmen lassen. Seien zum Beispiel $M \in \mathbb{N}$ Haushalte mit Merkmalsvektoren $x_1, \dots, x_M \in \mathbb{R}^D$, welche anhand ihres Stromverbrauchprofils in Q Cluster aufgeteilt wurden, gegeben, dann lassen sich die Schätzer $\hat{\pi}_q$ und \hat{f}_q definieren durch:

$$\hat{\pi}_q = \frac{\#\text{Anzahl Haushalte in Cluster } q}{M}, \quad (10)$$

sowie

$$\hat{f}_q(x) = \frac{\#\text{Anzahl Haushalte in Cluster } q \text{ mit Merkmalsvektor } x}{\#\text{Anzahl Haushalte in Cluster } q}. \quad (11)$$

Durch Einsetzen der beiden Schätzer in Gleichung (9) erhalten wir dann den kombinierten Schätzer:

$$\hat{f}_x(q) = \begin{cases} \frac{\hat{\pi}_q \hat{f}_q(x)}{\sum_{l=1}^Q \hat{\pi}_l \hat{f}_l(x)}, & \text{wenn } \sum_{l=1}^Q \hat{f}_l(x) > 0 \\ \hat{\pi}_q, & \text{sonst.} \end{cases} \quad (12)$$

Der zweite Teil der Gleichung (12) dient der Robustheit des Schätzers, da der Nenner den Wert 0 annehmen würde, falls ein Merkmalsvektor x nicht in den ursprünglichen Daten enthalten wäre. Nun seien die durchschnittlichen Stromverbräuche innerhalb eines Clusters \hat{v}_{ij}^q am Tag i zwischen Messpunkt j und $j - 1$ bekannt für $q = 1, \dots, Q$. Dazu sei ergänzend bemerkt, dass die Verbräuche eines Clusters zu einer bestimmten Partition aus den historischen Verbräuchen zu den einzelnen Monaten, an einem durchschnittlichen Wochentag, über den betroffenen Zeitraum gemittelt werden können und so Schätzwerte für v_{ij}^q ableitbar sind. Daraus lässt sich der erwartete Verbrauch \hat{v}_{kij} des k -ten Haushalts mit Merkmalsvektor x_k wie folgt berechnen:

$$\hat{v}_{kij} = \sum_{l=1}^Q \hat{f}_x(l) \cdot \hat{v}_{ij}^l. \quad (13)$$

Wird dieser Schätzer in die Ausgangsgleichung (8) eingesetzt, erhält man den entsprechenden Schätzer für den Stromverbrauch der Haushalte. Hinter dem eingeführten Merkmalsvektor x stehen numerisch codierte Merkmale, welche einen Haushalt charakterisieren und so auf den Verbrauch schließen lassen. Diese Merkmale, wie zum Beispiel die Anzahl der Haushaltsmitglieder, können sich über die Zeit ändern. Es wird jedoch zur Vereinfachung angenommen, dass dieser Vektor konstant ist.

Abschließend lässt sich nun der Kapitalwert (KW) des Mieterstrommodells mit Kalkulationszinssatz r und Investitionskosten I_0 zum Zeitpunkt $t = 0$ wie folgt berechnen:

$$KW = \sum_{t=1}^T \frac{CF_t}{(1+r)^t} - I_0. \quad (14)$$

Zudem ergibt sich die folgende Amortisationszeit (AZ):

$$AZ = \min_{t \in [1, T]} t + T \cdot I_{\sum_{i=1}^t CF_i < I_0} \quad (15)$$

wobei $I \in \{0,1\}$ definiert ist als Indikatorfunktion und somit AZ die Anzahl der Perioden definiert, bis die jährlichen Cashflows in Summe die anfänglichen Investitionskosten I_0 übersteigen, unter der Annahme das $AZ \leq T$.

4 Evaluation des Modells

In der folgenden Evaluation werden zwei Mieterstrominvestitionen auf Basis von Realweltdaten ökonomisch bewertet und versucht die Vorhersagegüte des Modells zu bestimmen [14]. Dabei wird ein Mieterstrommodell für vier Haushalte und eines für eine Quartierslösung (30 Haushalte) bewertet. Es wird angenommen, dass der Wärmebedarf der Beispielgebäude für die Heizung und Warmwasser nicht über Strom zur Verfügung gestellt wird. Zur Stromerzeugung werden bei den vier Haushalten eine $3kW$ Brennstoffzelle und eine $10kWp$ PV-Anlage installiert, welche annahmegemäß über 20 Jahre konstante Leistungen erbringen. In der Quartierslösung wird jeweils das 7,5-Fache dieser Leistung installiert. Damit liegt die Quartierslösung noch unter der vom Gesetzgeber vorgegeben Fördergrenze für Anlagen von insgesamt bis zu $100 kW$. Zusätzlich wird über den Betrachtungszeitraum ein konstantes Nutzerverhalten der Mieter angenommen. Die Kostenersparnis durch Skaleneffekte bei der Quartierslösung entsprechen 10% pro kW -Leistung, wohingegen die Vergütung der Netzeinspeisung vom Fördergeber für Großprojekte, nach deutscher Förderrichtlinie, um 2 Cent reduziert wird. Weitere Annahmen und Parameter sind in Tabelle 1 zusammengefasst.

Tabelle 1. Modellparameter und Annahmen der Evaluation

Variante	T	r	I_0	K	P_t^{Netz}	P_t^{MS}	L_V	L_R	L_T	L_S	T^{PV}
4-Mieter	20	2%	50.000€	3.000€	0,12€	0,25€	0,16	0,07	0,00	0,12	47°C
Quartier			337.500€	20.250€	0,10€						

4.1 Vorstellung der Daten und Datenaufbereitung

Als Stromverbrauchsdaten werden öffentliche Smart Meter Daten von 4225 Haushalten der irischen Kommission für Energieregulierung (CER) verwendet, welche über einen Zeitraum von Juli 2009 bis Dezember 2010 alle 30 Minuten erfasst wurden [30]. Neben den Stromverbräuchen wurden diverse weitere Haushaltsmerkmale erfasst. Die verwendeten Merkmale für das Modell sind in Tabelle 2 zusammengefasst. Da die betrachteten Immobilien annahmegemäß über keine Stromheizung verfügen, werden sämtliche Haushalte aus den Daten entfernt, welche nicht dieser Annahme entsprechen oder deren Daten nicht vollständig sind. Danach werden für die vier meteorologischen Jahreszeiten, für einen durchschnittlichen Werktag, in jeweils dreistündigen Zeitintervallen ab 0 Uhr, der durchschnittliche Stromverbrauch je Haushalt sowie die Änderungsraten berechnet (64 Kennzahlen). Für die Modellierung der Temperatur und der Globalstrahlung (hor.) werden Daten vom Deutschen Wetterdienst für die Stadt Würzburg über den Zeitraum 2005 bis 2017 in stündlicher Auflösung verwendet [31].

Tabelle 2. Verwendete Merkmale zur Charakterisierung der Haushalte. In Klammern sind die Ausprägungen der vier Beispielhaushalte der ersten Investition dargestellt

<i>Merkmale des Mieters</i>	<i>Ausprägungsmöglichkeiten</i>
Alter Hauptmieter	<35 ⁽⁴⁾ , 36-45 ⁽²⁾ , 46-55 ⁽³⁾ , 56-65, >65 ⁽¹⁾
Beschäftigungsstatus	Angestellt ⁽²⁻⁴⁾ , Selbstständig, Rente ⁽¹⁾ , Arbeitslos
Wohnform	Single ^(1,4) , Eltern mit Kindern ≤ 15 J. ⁽²⁾ , Sonstige ⁽³⁾
Anzahl Bewohner > 15 Jahren	0 ^(1,4) , 1-2 ^(2,3) , >2
Anzahl Kinder ≤ 15 Jahren	0 ^(1,3,4) , 1 ⁽²⁾ , >1
Anzahl Schlafzimmer	1-2 ^(1,3,4) , 3 ⁽²⁾ , >3
Internetanschluss	vorhanden ^(v) ⁽²⁻⁴⁾ , nicht vorhanden ^(n.v.) ⁽¹⁾
Geschirrspülmaschine	vorhanden ⁽²⁻⁴⁾ , nicht vorhanden ⁽¹⁾
Spielekonsole	vorhanden, nicht vorhanden ⁽¹⁻⁴⁾

4.2 Empirische Ergebnisse der Clusteranalyse

Zur Clusteranalyse wird der Partitioning Around Medoids (PAM) Algorithmus [32] und die Plattform R-CRAN sowie das Paket ‘cluster’ [33] genutzt. Die Auswahl der Clusteranzahl erfolgt über das Elbow-Kriterium, für welches ein Fehlermaß gegen die Anzahl der Cluster geplottet wird. Als Fehlermaß wird dabei die Summe der Entfernungsquadrate zwischen den Clustern zum Gruppenmittelpunkt des Clusters verwendet (Total within Sum of Squares). Mit zunehmender Anzahl der Cluster nimmt das Fehlermaß der Cluster ab, bis ab einem bestimmten Punkt eine deutliche Verflachung zu erkennen ist. Dieser Punkt markiert für die verwendeten Daten bei $Q = 6$ die optimale Anzahl der Cluster. Damit ist die gewählte Clusteranzahl etwas geringer als die vergleichbarer Arbeiten [8], was insbesondere darauf zurückzuführen ist, dass nur Haushalte betrachtet werden, die nicht mit Strom geheizt werden. In Abbildung 6 werden die durchschnittlichen Stromverbräuche an einem Werktag je Cluster dargestellt, worin sich die unterschiedlichen Verbrauchsprofile erkennen lassen. In Tabelle 3 sind die Ausprägungen dargestellt, welche im Vergleich zum Durchschnitt

besonders häufig in den Clustern auftreten. Daraus lassen sich grob ein Cluster für Singles (1), drei Cluster für Familien in den unterschiedlichsten Lebensphasen (2-4), ein Cluster für Paare (5) und eines für Rentner (6) ableiten. Zur Modellierung der SARIMA-Prozesse für den Verlauf der Globalstrahlung und der Temperaturkurven wird die Plattform R CRAN und das Paket ‘forecast’ verwendet [34]. Dabei werden für jeden Kalendermonat auf stündlicher Basis, für die Stunden zwischen Sonnenaufgang und -untergang mit Dauer $s_i \in \mathbb{N}$, für $i = 1, \dots, 12$ jeweils zwei $SARIMA(p, d, q) \times (P, D, Q)_{s_i}$ Modelle für $p, d, q, P, D, Q \leq 1$, gefittet und nach dem AIC-Kriterium ausgewählt, sodass sich 24 Zeitreihen ergeben. Um die Modellgüte der Zeitreihen als gegeben beurteilen zu können, werden im Anschluss die Modellparameter auf Signifikanz ($\alpha = 0.05$) untersucht, die Residuen mit Hilfe des Ljung-Box Tests auf Autokorrelation getestet sowie auf die Standardnormalverteilung untersucht.

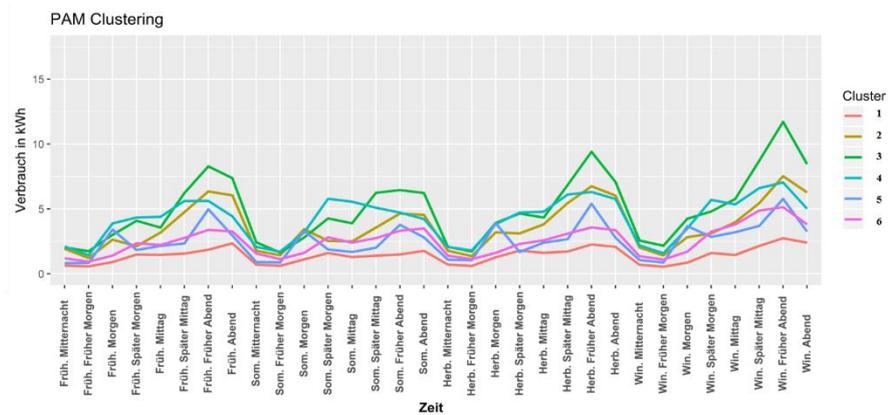


Abbildung 3. Stromverbräuche der Cluster nach Jahres- und Tageszeiten (Werktag) auf Basis von Smart Meter Daten der irischen Kommission für Energieregulierung (CER) [30]

Tabelle 3. Darstellung der Merkmale je Cluster die um 10 Prozentpunkte gestiegen sind zum Durchschnitt (~) auf Basis des Clustering von Stromverbrauchsdaten [30]

Merkmale /Cluster	1(rot)	2(braun)	3(grün)	4(h.blau)	5(blau)	6(pink)
Alter des Hauptmieters	~	36-55	46-55	~	~	>65
Beschäftigungsstatus	~	Angest.	Selbst.	~	Angest.	Rentner
Wohnform	Single	Familie	Familie	Sonstige	~	Sonstige
Anzahl Mitbewohner über 15 Jahren	0	1-2	>2	>2	1-2	~
Anzahl Kinder unter 15 Jahren	0	1,>1	>1	~	~	0
Anzahl Schlafzimmer	1-2, 3	>3	>3	>3	~	~
Internetanschluss	n.v.	v.	v.	v.	~	~
Geschirrspülmaschine	n.v.	v.	v.	v.	~	~
Spielekonsolen	n.v.	v.	v.	v.	~	n.v.
Anteil Haushalte (%)	27,3	15,4	9,1	9,6	11,8	26,8

4.3 Monte Carlo Simulation und empirische Ergebnisse

Auf Basis der in Abschnitt 4.2 kalibrierten Modelle werden nun für vier bzw. dreißig zufällig ausgewählte Haushalte aus dem Datensatz die in Tabelle 1 beschriebenen Mieterstrommodelle ökonomisch und ökologisch bewertet. Ziel ist es, die Erwartungswerte des Eigenverbrauchsfaktors ε_t sowie den der erzeugten Strommenge E_t^G auf Basis einer Monte Carlo Simulation mit 1.000 Durchläufen vorherzusagen. Auf dieser Basis können dann der Kapitalwert und die Amortisationszeit der Investitionen berechnet werden. Mit Erwartungswerten in Höhe von 33.400 kWh/a für die erzeugte Strommenge sowie von 65% für den Eigenverbrauchsfaktor, ergeben sich bei dem Mieterstrommodell mit vier Haushalten eine Amortisationszeit von 13 Jahren sowie ein erwarteter Kapitalwert von ca. 12.000€ über den Betrachtungszeitraum. Dem gegenüber steht die Quartierslösung mit einer erwarteten Amortisationszeit von 10 Jahren sowie einem erwarteten Kapitalwert von ca. 220.000€. Trotz der geringeren Einspeisevergütung amortisiert sich die Quartierslösung somit im Schnitt drei Jahre vor dem Mieterstrommodell mit nur vier Haushalten. Dies liegt unter anderem an dem höheren Eigenverbrauchsfaktor der Quartierslösung von insgesamt 76%, welcher in Tabelle 4 auf monatlicher Basis dargestellt ist. Der höhere Eigenverbrauchsfaktor lässt sich über die rechtsschiefe Verteilung der Stromverbräuche erklären, sodass mit einer steigenden Anzahl an Mietern, Haushalte mit hohen Verbräuchen statistisch häufiger auftreten und somit auch mehr Strom abgenommen wird. Insgesamt wird auf Grundlage des Anlagendesigns die 7,5-Fache Strommenge im Quartier produziert. Würden sich die Investitionskosten I_0 und die jährlichen Kosten K der Quartierslösung im Vergleich zum Mieterstrommodell mit vier Haushalte um jeweils 20% verringern pro kW -Leistung (davor 10%), wird eine Amortisationszeit von 8 Jahren erreicht. Diese Einsparungen können in etwa als obere Grenze für die betrachtete Quartierslösung gelten [35]. Wie in Tabelle 4 dargestellt, tritt bei der Vorhersage des monatlichen Eigenverbrauchsfaktors bei der Investition mit vier Mietern ein absoluter Fehler von 3,5 Prozentpunkten und bei der Quartierslösung von 2,5 Prozentpunkten auf.

Table 4: Empirische Ergebnisse des Entscheidungsunterstützungsmodells

	Jan.	Feb.	Mär.	Apr.	Mai	Jun.	Jul.	Aug.	Sep.	Okt.	Nov.	Dez.
<i>4-Mieter</i>												
$E[\varepsilon]$	0,79	0,72	0,65	0,59	0,56	0,55	0,57	0,58	0,62	0,67	0,74	0,84
ε_{real}	0,75	0,69	0,60	0,58	0,54	0,53	0,55	0,55	0,58	0,63	0,69	0,77
<i>Quartier</i>												
$E[\varepsilon]$	0,88	0,83	0,77	0,70	0,69	0,67	0,69	0,70	0,74	0,79	0,84	0,90
ε_{real}	0,86	0,83	0,77	0,65	0,63	0,67	0,67	0,67	0,70	0,78	0,82	0,87

Wird unter Betrachtung des gesamten Lebenszyklus der Mieterstrominvestition die CO₂-Einsparung berechnet, ergeben sich für die Quartierslösung eine jährliche Einsparung in Höhe von 95 Tonnen CO₂, was einer Reduktion von 64% im Vergleich zum deutschen Kraftwerksmix gleichkommt.¹ Dies entspricht in etwa dem CO₂-

¹ Es wurde ein CO₂-Ausstoß pro kWh Strom von 100 Gramm bei PV, 240 Gramm bei einer Brennstoffzelle und 550 Gramm im deutschen Kraftwerksmix angenommen [36].

Ausstoß eines Mittelklassewagens über ca. 600.000 km Fahrstrecke. Für das zweite Szenario ergibt sich eine jährliche Einsparung von 13 Tonnen CO₂. Der durchschnittliche Stromverbrauch pro irischem Haushalt von 3.700 kWh/a, auf dessen Basis die Ergebnisse berechnet wurden, liegt leicht über dem durchschnittlichen Verbrauch von 3.200 kWh/a in Deutschland [37]. Ein auf dieser Grundlage kalibriertes Modell wird den Strombedarf deutscher Haushalte überschätzen und dadurch tendenziell zu kurze Amortisationszeiten berechnen. Wird der durchschnittliche Mehrverbrauch der irischen Haushalte pauschal abgezogen, ergeben sich für Deutschland im Schnitt ein Jahr längere Amortisationszeiten.

5 Zusammenfassung, Limitationen und Ausblick

Für den Erfolg der Energiewende spielt die dezentrale Stromerzeugung eine entscheidende Rolle. Dabei bietet Mieterstrom für die auftretenden Herausforderungen einen vielversprechenden Lösungsansatz, welcher insbesondere durch die Nutzung des steigenden Datenvolumens durch Smart Meter zu neuen Geschäftsmodellen führen kann. In dieser Arbeit wurde deshalb zur Investitionsbewertung von Mieterstrommodellen ein Entscheidungsunterstützungssystem entwickelt und evaluiert. Die Ergebnisse der Wirtschaftlichkeitsbetrachtung zeigen, dass sich Mieterstrommodelle unter den getroffenen Annahmen je nach Größe nach 8 bis 14 Jahren amortisieren. Zusätzlich konnten bei der hybriden Nutzung von PV-Anlagen und Brennstoffzellen CO₂-Einsparungen von über 60% nachgewiesen werden. Zu den Defiziten der Annahmen: (1) Ein gleichbleibendes Nutzverhalten kann nur schwer über lange Zeiträume angenommen werden, da die natürliche Veränderung der Lebenssituation eines Haushaltes zu Veränderungen im Verbrauchsverhalten führen kann. Diese Annahme könnte jedoch mit einem zusätzlichen Teilmodell zur Fortschreibung der Haushaltsmerkmale entkräftet werden. (2) Es werden keine Ausfälle durch Vertragskündigungen berücksichtigt, wie sie zum Beispiel durch Mieterwechsel entstehen. An dieser Stelle könnten Langzeitstudien den notwendigen Aufschluss über die Häufigkeit dieser Ausfälle liefern. (3) Die berechneten Amortisationszeiten sollten auf Basis deutscher Verbrauchsdaten validiert werden, um so die vollständige Übertragbarkeit auf Deutschland zu gewährleisten. (4) Entstehende Preisrisiken durch die stochastische Preisentwicklung für Energieträger und die zukünftige Entwicklung der Betriebskosten des Mieterstrommodells werden aktuell noch nicht abgebildet. Dies kann jedoch auf Basis von bestehender Literatur ergänzt werden. Zukünftige Erweiterungen des Modells sollten neben den vier genannten Limitationen insbesondere die Integration von Speichertechnologien sowie das Zusammenspiel von Mieterstrom und Elektrofahrzeugen berücksichtigen. Durch diese Trends könnten sich die Profile von abgerufenen und bereitgestellten Strommengen signifikant ändern und so zu neuen Rahmenbedingungen für Mieterstrommodelle führen. Um ein beratungs- und serviceorientiertes Geschäftsmodell basierend auf dem entwickelten Ansatz zu realisieren, ist zusätzlich eine allumfassende IT-Unterstützung notwendig. Hier ist es weiterhin notwendig die technische Ausgestaltung solcher IT-Anwendungen im Detail zu betrachten. In diesem Zusammenhang spielt auch das

Thema Datenschutz eine wichtige Rolle, da die Privatsphäre der Mieter ebenfalls geschützt werden muss. Trotz der genannten Limitation zeigt der vorliegende Beitrag auf, wie ein Entscheidungsunterstützungssystem für Mieterstrom konzipiert sein müsste und kann somit zukünftig zur Identifikation und Umsetzung erfolgreicher Investitionen beitragen.

Literatur

1. Davis, S.J., Caldeira, K., Matthews, H.D.: Future CO₂ emissions and climate change from existing energy infrastructure. *Science (New York, N.Y.)* 329, 1330–1333 (2010)
2. Hake, J.-F., Fischer, W., Venghaus, S., Weckenbrock, C.: The German Energiewende – History and status quo. *Energy* 92, 532–546 (2015)
3. Pepermans, G., Driesen, J., Haeseldonckx, D., Belmans, R., D’haeseleer, W.: Distributed generation: definition, benefits and issues. *Energy Policy* 33, 787–798 (2005)
4. Behr, I., Großklos, M.: Mieterstrom – ein Beitrag zur Energiewende. In: Behr, I., Großklos, M. (eds.) *Praxishandbuch Mieterstrom*, pp. 3–14. Springer Fachmedien, Wiesbaden (2017)
5. Beucker, S., Hinterholzer, S., Schweinoch, M., Zernahle, O.: Messkonzepte für Mieterstrom. In: Behr, I., Großklos, M. (eds.) *Praxishandbuch Mieterstrom*, 36, pp. 81–98. Springer Fachmedien, Wiesbaden (2017)
6. Leprih, U., Neumann, W.: Zur Rolle des Mieterstroms in Energiesystemen der Zukunft. In: Behr, I., Großklos, M. (eds.) *Praxishandbuch Mieterstrom*, pp. 15–27. Springer Fachmedien, Wiesbaden (2017)
7. Großklos, M.: Erzeugungs- und Verbrauchsprofile für die Planung von Mieterstrom. In: Behr, I., Großklos, M. (eds.) *Praxishandbuch Mieterstrom*, pp. 113–123. Springer Fachmedien, Wiesbaden (2017)
8. McLoughlin, F., Duffy, A., Conlon, M.: A clustering approach to domestic electricity load profile characterisation using smart metering data. *Applied Energy* 141, 190–199 (2015)
9. Sodenkamp, M., Kozlovskiy, I., Hopf, K., Staake, T.: Smart Meter Data Analytics for Enhanced Energy Efficiency in the Residential Sector. In: *Wirtschaftsinformatik 2017 Proceedings*. AIS electronic library, St. Gallen. (2017)
10. Hopf, K., Sodenkamp, M., Kozlovskiy, I., Staake, T.: Feature extraction and filtering for household classification based on smart electricity meter data. *Comput Sci Res Dev* 31, 141–148 (2016)
11. Watson, R.T., Boudreau, M.-C., Chen, A.J.: Information Systems and Environmentally Sustainable Development: Energy Informatics and New Directions for the IS Community. *MIS Quarterly* 34, 23 (2010)
12. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. *MIS Quarterly* 28, 75 (2004)
13. Nestle, U., Kunz, C.: *Studienvergleich: Stromgestehungskosten verschiedener Erzeugungstechnologien*. Agentur für erneuerbare Energien, Berlin, (2014)
14. Paschka, D.: Wirtschaftlichkeit von Mieterstrom. In: Behr, I., Großklos, M. (eds.) *Praxishandbuch Mieterstrom*, pp. 125–139. Springer Fachmedien, Wiesbaden (2017)
15. Masa-Bote, D., Castillo-Cagigal, M., Matallanas, E., Caamaño-Martín, E., Gutiérrez, A., Monasterio-Huelín, F., Jiménez-Leube, J.: Improving photovoltaics grid integration through short time forecasting and self-consumption. *Applied Energy* 125, 103–113 (2014)
16. McLoughlin, F., Duffy, A., Conlon, M.: Characterising domestic electricity consumption patterns by dwelling and occupant socio-economic variables: An Irish case study. *Energy and Buildings* 48, 240–248 (2012)

17. Fuentes, M., Nofuentes, G., Aguilera, J., Talavera, D.L., Castro, M.: Application and validation of algebraic methods to predict the behaviour of crystalline silicon PV modules in Mediterranean climates. *Solar Energy* 81, 1396–1408 (2007)
18. Mellit, A., Pavan, A.M.: A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. *Solar Energy* 84, 807–821 (2010)
19. Mellit, A., Benghanem, M., Arab, A.H., Guessoum, A.: A simplified model for generating sequences of global solar radiation data for isolated sites: Using artificial neural network and a library of Markov transition matrices approach. *Solar Energy* 79, 469–482 (2005)
20. Reikard, G.: Predicting solar radiation at high resolutions: A comparison of time series forecasts. *Solar Energy* 83, 342–349 (2009)
21. Masa-Bote, D., Caamaño-Martín, E.: Methodology for estimating building integrated photovoltaics electricity production under shadowing conditions and case study. *Renewable and Sustainable Energy Reviews* 31, 492–500 (2014)
22. Yin, E., Li, Q., Xuan, Y.: Optimal design method for concentrating photovoltaic-thermoelectric hybrid system. *Applied Energy* 226, 320–329 (2018)
23. Beckel, C., Sadamori, L., Staake, T., Santini, S.: Revealing household characteristics from smart meter data. *Energy* 78, 397–410 (2014)
24. Haben, S., Singleton, C., Grindrod, P.: Analysis and Clustering of Residential Customers Energy Behavioral Demand Using Smart Data. *IEEE Trans. Smart Grid* 7, 136–144 (2016)
25. Wang, Y., Chen, Q., Kang, C., Xia, Q.: Clustering of Electricity Consumption Behavior Dynamics Toward Big Data Applications. *IEEE Trans. Smart Grid* 7, 2437–2447 (2016)
26. Al-Wakeel, A., Wu, J., Jenkins, N.: k -means based load estimation of domestic smart meter measurements. *Applied Energy* 194, 333–342 (2017)
27. Osterwald C.R.: Translation of device performance measurements to reference conditions. *Solar Cells*, 269–279. (1986)
28. International Electrotechnical Commission: Crystalline silicon terrestrial photovoltaic (PV) modules—design qualification and type approval. International Standard IEC 61215 (2005)
29. Fang, T., Lahdelma, R.: Evaluation of a multiple linear regression model and SARIMA model in forecasting heat demand for district heating. *Applied Energy* 179, 544–552 (2016)
30. Commission for Energy Regulation (CER): CER Smart Metering Project - Electricity Customer Behaviour Trial, 2009-2010 [dataset]. 1st Edition, www.ucd.ie/issda/CER-electricity (Accessed: 12.09.2018)
31. DWD Climate Data Center: Historische stündliche Stationsmessungen für Deutschland, <ftp://ftp-cdc.dwd.de/pub/CDC/>(Accessed: 12.09.2018)
32. Kaufman, L., Rousseeuw, P.J.: Finding Groups in Data. An Introduction to Cluster Analysis. John Wiley & Sons Inc, Hoboken (2009)
33. Maechler, M., Rousseeuw, P., Struyf, A.: cluster - Cluster Analysis Basics and Extensions. R package version 2.0.7-1 (2018)
34. Hyndman, R.J., Khandakar, Y.: Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 1-22 (2008)
35. BKI-Baukosteninformationszentrum Deutscher Architektenkammern, B.-B.D.: BKI Baupreise kompakt 2018 - Neu- und Altbau. Statistische Baupreise für Positionen mit Kurztexen. BKI, Stuttgart (2017)
36. Memmler, M., Lauf, T., Wolf, K., Schneider S. (eds.): Emissionsbilanz erneuerbarer Energieträger. Bestimmung der vermiedenen Emissionen im Jahr 2016. Im Auftrag des Umweltbundesamtes (2017)
37. Sustainable Energy Authority of Ireland: Energy in the residential sector. 2018 Report (2018)