

Characterizing residential sector load curves from smart meter datasets

Andrew S. Jin^{a,b}, Kelly T. Sanders^{a,*}

^a Department of Civil and Environmental Engineering, University of Southern California, United States

^b Environmental Laboratory, United States Army Engineer Research and Development Center, Vicksburg, MS, United States

HIGHLIGHTS

- Hourly smart-meter data of 160,000 homes were used to construct diurnal load profiles.
- Disparities in hourly loads between households were measured using the Gini index.
- The top 20% of households consumed about 40% of the daily electricity load.
- The bottom half of consumers used <25% of the daily electricity load.
- Customers in mild climates show less seasonal variability than those in hot climates.

ARTICLE INFO

Keywords:

Smart meter
Load profiles
Inequality
Demand response
Residential electricity
Big data

ABSTRACT

Understanding how and when residential electricity is used throughout the day is integral to the successful implementation of potential residential demand management strategies. Our analysis characterizes the daily hourly load profiles of approximately 160,000 residential electricity customers across the Southern California Edison (SCE) service area during the period spanning 2015 to 2016 and 2018 to 2019 across weekends, weekdays, seasons, and climate zones. We find that total daily electricity usage was highest in the hottest months of the year compared to milder months, particularly for households located in the hottest climate zones. The most energy-consumptive hours occurred during the mid-afternoon during the hottest months, in contrast to early evening high consumption in cooler months. We find that customers with average daily consumption at or above the 80th percentile cumulatively consume over 40% of electricity during the hottest months of the year residential load, while the bottom half of customers cumulatively consume <25% of the total residential load. The disparities in electricity usage across SCE households are higher in the mid-day, especially in milder months across all regions, and in mild climate zones compared to hotter climate zones since loads are not as dependent on high HVAC loads.

1. Introduction

In recent years, the global energy landscape has begun a transformative shift towards less carbon-intensive forms of electricity generation in efforts to mitigate the negative consequences of climate change [1]. Renewable electricity is projected to be the largest source of global electricity generation by early 2025 and is the only electricity generation source whose share is expected to grow [2]. Renewables are expected to displace fossil fuels in the United States electric power sector due to declining renewable technology costs for renewable power through 2050 [3]. Unlike conventional thermal generators, wind and solar PV generators have variable power outputs that create operational

challenges in balancing supply to meet demand [4]. To deal with these challenges, utilities around the United States and internationally have implemented new strategies to better forecast and manage energy supplies and demands [5].

In California, this challenge of balancing supply and demand is already growing more difficult because of a changing climate and changes in demand patterns due to increased electrification. Solar generation within the service territory overseen by California Independent System Operator (CAISO, which covers 80% of California's bulk power transmission [6]), represented 17% of total system electricity generated in 2022 [7]. In CAISO, high penetrations of solar power have created a so-called "duck curve" that is characterized by a deep daytime net-load

* Corresponding author.

E-mail address: ktsanders@usc.edu (K.T. Sanders).

<https://doi.org/10.1016/j.apenergy.2024.123316>

Received 9 February 2024; Received in revised form 14 April 2024; Accepted 22 April 2024

Available online 3 May 2024

0306-2619/© 2024 Published by Elsevier Ltd.

(where net load is the total load less variable renewable electricity) dip followed by a steep increase in net load that occurs as solar resources go offline in the early evening and CAISO simultaneously enters its evening peak [8]. The duck curve can lead to overgeneration and curtailment of renewables during the midday “belly” of the curve, while the “neck” of the curve requires conventional generators to rapidly ramp operations to address peak demand in the early evening, leading to higher costs and more emissions [9]. As California continues to grow its solar power portfolio, CAISO’s midday net load continues to dip lower year over year, exacerbating issues with overgeneration and ramping challenges [10]. While CAISO’s battery storage capacity, which increased tenfold from 2020 to 2023, provided key net peak capacity and energy during the summer of 2022, its deployment is still relatively limited and not yet sufficient to fully ameliorate the challenges of managing CAISO’s duck curve. Thus, grid managers must still contend with a deepened duck curve in the early afternoon as coincident with increases in the residential electricity demand.

Because grid operators must match supply constraints with electricity demand, systems like CAISO have moved from a paradigm of “matching available supply with dynamic demand” to “matching dynamic supply with dynamic demand” through demand-side management strategies, like demand response [11]. Demand response strategies are designed to influence customer use of electricity in order to encourage customers to use less power during peak times, or to shift energy use to off-peak hours [12]. Demand response strategies, as defined by the United States Federal Energy Regulatory Commission, are “changes in electric usage by demand-side resources from their normal consumption patterns in response to changes in the price of electricity over time, or to incentive payments designed to induce lower electricity at times of high wholesale market prices or when system reliability is jeopardized” [13]. Demand response programs can include voluntary or incentive-based demand response programs, such as direct load control, emergency demand response programs, or interruptible/curtailable rates [14].

While most demand-side management strategies have historically been targeted at larger electricity customers in non-residential sectors, state policies in California have accelerated the implementation of new time-of-use (TOU) rates in recent years [15]. In 2003, a residential TOU-pilot program was designed and implemented with the help of the three major investor-owned utilities in California: Pacific Gas and Electric Company (PG&E), Southern California Edison, and San Diego Gas and Electric Company [16]. Since then, a 2016 rulemaking effort has been put into place an effort to analyze the electricity load of CAISO and develop analyses to develop new TOU designs [17], with a pilot study commissioned in 2018 proposing the implementation of default TOU pricing for all residential electricity customers [18]. Time-of-use rates have been implemented as the default pricing scheme for customers in PG&E [19] and SDG&E [20] since 2019 and SCE since 2021 [21]. However, while there has been a concerted effort to develop such rates and understand the impact of such pricing rates on decreasing the load at peak hours, understanding the load shape of the residential sector remains a key gap in the literature.

Currently, the end-use load profile models developed to understand California’s demand in end-use sectors give a poor understanding of how residential electricity demand fluctuates throughout the day. Until 2017, the residential load profile used by the California Energy Commission was informed by the CEC’s Hourly Electric Load Model (HELM) [22], which was based on a small number of metered homes in the late 1980s. The most recent California end-use load shape report uses aggregated smart meter data from various houses to develop average daily residential load profiles [23]. However, the report only details an annual load profile and does not offer detailed insights into differences between customers within the residential sector (e.g., by housing type, by climate zone, by season, etc.). Hence, there are key gaps in our understanding of the differences in how customers within the same or adjacent geographic regions contribute to the aggregate residential load

curve.

Understanding customer load curves is critical to several residential electricity policies. While demand response strategies incentivize behavioral changes in individuals, the primary goal of demand response strategies is to reduce net demand during peak times [24]. While most conventional electricity rate structures are priced on the quantity of electricity consumed, these tariffs have to cover both the costs of generation and the larger fixed costs of delivering electricity such as distribution and transmission infrastructure [25]. Thus, utilities need to equitably balance pricing the fixed portion of costs associated with electricity consumption with incentivizing energy efficiency and reducing electricity use during peak times [26]. Understanding how different customers contribute to the overall load curve can help utilities design more effective and equitable residential demand response programs. Further, modulating this challenge has been the increase in rooftop solar in California, which has grown from 1200 MW in 2014 to 10,500 MW in 2024 [27]. While customers selected in our study occurred in 2015, when <5% of SCE customers had rooftop solar installations [28,29] solar adoption has been found to be growing fastest for the largest electricity consumers in wealthy neighborhoods, who gain the most benefits from new installations [30]. Thus, developing methods to construct load profiles through smart metering infrastructure can help to provide insights into how to better structure incentives equitably.

This study develops a longitudinal data-based approach that constructs aggregate residential load curves and evaluates the contributions of utility customers based on their annual net electricity consumption across different day types (i.e., weekday vs weekend), seasons, and climate zones. It uses a real-world hourly smart meter data consisting of over two billion hourly readings from a statistically representative sample of 160,000 electricity users within Southern California Edison’s service area to answer three major research questions regarding residential load curves:

- What are the diurnal and seasonal characteristics in the aggregate residential load curve in Southern California Edison’s service area?
- How do customers within different electricity consumption decile bins contribute to the overall load curve throughout seasons, and do higher consumption deciles have an outsized contribution to residential electricity load during specific hours of the day?
- How do customers in different climate zones contribute to the overall load curve (i.e., in terms of shape and magnitude)?

In this study, we evaluate a large, multi-year smart meter dataset to create average electricity consumption profiles for households in Southern California. We classify utility customers into one of ten electricity consumption decile bins, which represent the percentile ranking of each household based on its typical daily electricity consumption. These deciles are determined by computing the average daily electricity consumption for each customer based on all their records in our multi-year dataset, which could span up to four years depending on the household. This study develops average load profile shapes for customers in Southern California that can be used to assess differences in the shape of the electricity load by season, day of the week, climate zone, and consumption-based decile bins. We then use a GINI metric to characterize disparities in the contribution of the overall load curve by customers in different decile bins [31]. Thus, this study not only characterizes the load shape of the residential sector electricity load profile in Southern California but also gives insight into the disparate contribution that customers who use different amounts of electricity or live in different geographic locations have on the diurnal patterns of the residential load profile. These insights can help enable (a) the tracking of how the residential electricity consumption shifts over time and (b) more equitable demand response targeted policies.

In the following sections, we review the state of the literature in the load profiling community and identify existing knowledge gaps. Then,

we present our methods to construct load profiles, characterize disparity, and describe the results of our study by analyzing the load profiles in Southern California by season, decile bin, and climate zone. Finally, we discuss the implications of the load profile shapes and how they can be used to inform improved demand response strategies.

2. Literature review

2.1. Load profiles

Load Profiles (also referred to as load shapes, demand shapes, load curves, demand curves, etc.) are a way of characterizing the variation of demand and electrical load over a specific period. In the academic literature, these have been largely referred to the diurnal hourly or sub-hourly graph of electricity use of individual customers (i.e. individual meters), groups of individuals, or different customer sectors (i.e. residential, commercial, industrial). The construction of load profiles has been used for a number of different tasks ranging from supply management requirements, such as electricity forecasting [32–35], and system design [36–38], to demand-side analyses, like tariff design [39–41]. Load profiles can generally be divided into two general models, one that characterizes a large group of potential customers to establish standard load profiles, an aggregate electricity profile that represents the overall behavior of the residential sector, and one that characterizes the behaviors of subgroups of customers that have similar behaviors [42].

The residential sector presents several unique challenges that make it harder to characterize load profiles. The decentralized nature of residential customers makes it difficult to characterize their consumption. While residential customers represent 39% of end-use electricity consumption in the United States, they represent approximately 87% of utility customers [43]. Electricity customers have a high diversity in building characteristics and occupant demographics that lead to dramatic differences in the load profiles of different households [44]. Even with nearly identical houses with the same appliances and similar demographic occupants, the load profiles of two different homes can be very different [45]. By contrast, other sectors, such as commercial, industrial, agriculture, and transportation, generally attract more significant financial motivation to analyze and understand their electricity usage patterns and have more centralized platforms from their respective industries to compare patterns of behavior [46]. Historically, load profiles of the residential sector can either be surveyed by utilities themselves (typically limited to a small number of voluntary survey participants), or the load profile of a customer group can be measured at the transformer or branch feeder level [47].

Several reviews have reviewed methods for developing load profiles in the residential sector [42,48–50]. Because of data limitations, two main archetypes have been used to interpolate the load profile, bottom-up and top-down strategies. Top-down approaches rely on aggregate data from utility operators, such as the all-sector load curve, and apply statistical methods and models to make inferences about what the residential sector may look like [51,52]. One approach to top-down modeling is Conditional Demand Analysis, which uses actual smart meter and appliance data to train models that disaggregate residential electricity use into functions of specific appliances, without the need for direct appliance measurement, and scales it up to the overall electricity load profile. This approach is exemplified in studies by Parti and Parti [53] and Aigener et al. [54]. Bottom-up approaches rely on detailed data from individual customers or appliances and use simulation models or other engineering methods to aggregate a residential load profile [55]. For example, Capasso et al. implemented a Monte-Carlo simulation based on various factors such as weather, time of day, and household appliance curves to create an aggregate total load [56]. While top-down approaches have high uncertainty regarding the actual behavior and preferences of individual customers, bottom-up approaches may not capture the true nature of customers because of their limited sample

sizes. Thus, generating data-driven load profiles, which have been enabled by the proliferation of smart meters, can provide a real-world-based analysis of electricity customers.

2.2. Smart meters data analysis in the residential sector

Smart meters, or Automated Metering Infrastructure (AMI), that provide hourly or sub-hourly information about electricity usage, are being installed in households in many regions. The growth of AMI has largely been driven by the operational benefits to utilities as automated metering can reduce the number of site visits to read meters, identify disruptions to service quickly, and more accurate and timely billing [57]. Because these operational benefits have been the driver of the implementation of smart meters, the use of smart meters as informational tools to understand how customers consume electricity has been underexplored [58].

Past reviews of smart meter analyses have found the large majority use relatively small numbers of smart-meter readings for their analysis, with many using small open-source residential datasets on the order of thousands of unique meters [59,60]. The use of massive datasets with >100,000 users that have been applied to academic literature has been limited to only a small number of distinct datasets [61–68]. Of these aforementioned papers, two papers were published directly by the utilities regarding internal operational research to improve the forecasting of electricity usage [62] or to improve automated customer feedback in billing [61]. By contrast, most of the papers that analyze these large datasets seek to understand how customers use electricity through the construction of load profiling and load analysis techniques [63–67].

One key topic of research is to understand how individual electricity customer's electricity load is affected by the use of different installed appliances. For example, Chen et al. used a dataset of over 180,000 customers in Southern California to analyze the existence of air conditioning within households by analyzing the temperature sensitivity of electricity usage [68]. Other analyses use smart meter electricity datasets tagged with the presence of different appliance technologies to analyze the impact of those technologies on the electricity profiles of houses. Anderson et al. studied a dataset of 667,373 houses in Denmark, with smart meter data matched to data about each household regarding the income and age of occupants, heating technology, and the presence of electric vehicles to understand differences in the level and timing of consumption throughout the day [63].

Load analyses have largely focused on characterizing load profiles in the residential sector by first segmenting customers by their similar electricity load profiles through clustering techniques. Generally, these methods characterize load profiles for groups of customers. Kwac, Flora, and Rajagopal (2014) first proposed and applied a method to segment households based on their hourly electricity consumption patterns using a dictionary of load shapes to characterize similar load profiles in 218,090 homes in Northern California [69]. A similar clustering method simplifies the load shapes by developing average load profiles for each user by month and for 2.5 million Illinois customers [70]. More sophisticated versions of this clustering methodology have been applied as well. For example, Ushakova and Mikhaylov propose a Gaussian-Mixture model to encode individuals' energy consumption over time based on 400,000 homes in the United Kingdom [67].

While the aforementioned group of papers characterized homes by similar behaviors, fewer papers have aggregated multiple customers based on temporal factors, such as seasonal and day-of-week differences, or subsets of the population before describing their load. A study analyzing the impacts of COVID-19 restrictions on residential electricity usage used 230,000 smart meters to analyze week-long hourly resolution smart meter load profiles in Santiago, Chile, finding that median household loads were formed for individual neighborhoods to compare residential demand [65]. Gunkel et al. (2023) [71] used the aforementioned Danish smart meter dataset in [63] to analyze the contributions

of individual customers within a customer class to the aggregate profiles of that class.

Current load profiling methods that utilize massive volume smart meter datasets have key limitations that we address in this study. No high-volume smart meter dataset analysis has tracked customers across a span of five years. Our study addresses a key gap in understanding how a household's load shape might vary across climate zones. Furthermore, no paper has analyzed the contribution of utility customers, binned according to their typical household electricity consumption, to the aggregate residential load profile of SCE. We then analyze the diurnal patterns of disparity in hourly electricity consumption of these binned customers. In the following section, we describe our methodology of creating data-driven load profiles and analyzing the contributions of and disparities between different users.

3. Methods

3.1. Data description

Our data was sourced from Southern California Edison, a large investor-owned utility serving the greater Los Angeles area. We obtained hourly electricity data from 200,000 residential smart meters, randomly selected so that the dataset is statistically representative of 5% of the SCE's 4.5 million residential households in their service area which spans the Greater Los Angeles [See Supplemental 1].

Each residential smart meter was associated with a unique service-account ID, representing data for one utility customer account. Data spanned from 2015 to 2016 and 2018–2019. Customers were initially requested and randomly selected from the SCE service area for the 2015–2016 period, and customer data was updated for the 2018–2020 period. Data from 2017 was not available due to privacy regulations required deletions by the utility before we requested updated data in 2021. Thus, some customers may have dropped out between years if the residents of that household moved. Data from 2020 was excluded due to the COVID-19-related changes in residential electricity usage. All electricity data were stored on USC's Center for High-Performance Computing with a highly secure HPC Secure Data Account, to remain in line with the security and confidentiality requirements of SCE. In our analysis, we only consider the electricity load profiles of sub-metered residential units (e.g. apartment buildings where each apartment has a meter) and filter out master-metered properties (e.g. apartment buildings or trailer parks with one master meter).

3.2. Individual load profile definition

In this paper, we begin by analyzing the individual load profiles of all utility customers (each representing a household with a smart meter) within our dataset. The dataset was organized into customer-day records, where each day is represented by a unique hourly electricity load profile. Each individual load profile record, l , is represented by a set where l_h represents the electricity used in a one-hour period, measured in kWh, at the hour, starting at h , of day, d , for an individual customer i [Eq. (1)]. Meter readings represent hourly net consumption, and no smart meter record can indicate negative electricity use in a given hour in our dataset (i.e. the meter will read zero, even in the case that solar electricity produced from a home is higher than the electricity consumed). Unlike the traditional electric industry-wide term “hour ending”, we will refer to hours by the starting hour of the electricity consumption. For example, hour $h=16$ will refer to the electricity consumed during the hour starting at 4:00 pm and ending at 5:00 pm. We then calculate the total daily electricity, e , measured in kWh, for an individual load profile.

$$l(i, d) = [l_0, l_1, \dots, l_h, \dots, l_{23}] \quad (1)$$

$$e(i, d) = \sum_h^{23} l_h \quad (2)$$

For each individual utility customer, i , the set D_i is the set of all complete sets of full-day records in a customer's dataset. Throughout this analysis, we will use the prime notation (D'_i) to describe the subset after filtering.

3.3. Filtering of non-residential and outlier residential electricity customer

In the dataset provided by SCE, each customer is identified by a service-account ID which is unique to each customer and is affiliated with a postal address. This address was parsed and matched with building assessor data in Los Angeles, San Bernadino, Orange, and Riverside Counties, and found that <1.5% of households were matched to non-residentially zoned parcels [Supplemental Methods – Section 2. D] The addresses were then attached to a geographic coordinate and tagged with a California Energy Commission Building Climate Zone.

The set D_i , is the set of all complete sets of full-day records, d , available for an individual customer, i , across Y_i , the set of all years, y , in which that customer has data. We include customer-years of data that meet the following conditions: First, the set of customer-day records for a customer in a given year must contain >200 customer-day load profiles. Second, the total electricity consumption in that given customer-year must be >20 kW-hours (Eq. (3)). If a customer does not have at least one qualifying year of data (i.e. Y'_i is an empty set), that customer is excluded from our analysis.

$$Y'_i = \left\{ y \in Y : |D_{i,y}| > 200 \wedge \sum_{d \in D_{i,y}} e(i, d) > 20 \text{ kWh} \right\} \quad (3)$$

We performed a manual assessment of the largest electricity customers in our dataset, whose average annual electricity usage was >100,000 kWh, to check if these customers represented one household. Records that were associated with addresses where more than one household was collectively measured by one meter (e.g., trailer parks or housing communities with only one SCE meter) were excluded [See Supplemental 2.D]. After all filtering steps were completed, our final dataset constituted 184 million customer-days of hourly smart meter readings data from 163,403 customers over 514,370 customer-years.

3.4. Customer electricity consumption decile bins

For each customer, we calculate the average daily electricity use value, $a(i, y)$, for each year y which satisfies our inclusion criteria Y'_i (Eq. (4)). Because meter readings in our dataset represent hourly net-consumption, $a(i, y)$, represents the average electricity consumed from the grid by a given customer. Averaging to the daily level reduces the influence of potentially missing data in each year and allows comparison between different customers. We then calculate the average annual-daily average electricity net consumption, $A(i)$, by averaging the set of eligible customer-years (Eq. (5)). Thus, $A(i)$, measured in kWh per day is the mean-annually-averaged daily net electricity consumption across the period of study for each customer.

To characterize the relative consumption of electricity customers to one another, we rank all electricity customers based on their average annual electricity consumption, $A(i)$. Annual electricity consumption is then used to order all individual customers with eligible data through a percentile function, $P(i)$ (Eq. (6)). Once percentile is calculated, each individual customer is assigned an electricity consumption tier, such that each decile bin, B_n , represents a decile of electricity usage, where decile 1 (0–10%) represents the lowest consumption customers consuming a daily average of 0.05 and 6.6 kWh and decile 10 (90–100%) represent the highest consumption customers consuming a daily average of >33.4 kWh (Eq. (7)) [See Supplemental 3].

$$a(i, y) = \frac{\sum_{d \in D_{i,y}} e(i, d)}{|D_{i,y}|} \quad (4)$$

$$A(i) = \frac{\sum_{y \in Y_i} a(i, y)}{|a_{i,y}|} \quad (5)$$

$$P(i) = \frac{\text{Rank}(A(i))}{|U|} \quad (6)$$

$$B(i) = [P(i) * 10] \quad (7)$$

3.5. Average daily load profiles

We define three season classifications primarily based on weather patterns in Southern California's Mediterranean climate: "Winter" defined as November–December–January–February (NDJF), "Spring" defined as March–April–May–June (MAMJ), and "Hottest Months" defined as July–August–September–October (JASO). JASO, combining summer and early fall, is the hottest period of the year, characterized by extended heat and dryness typical of Southern California summers, which often extend into the traditional autumn months. Because of diversity of climate zones contained within the SCE service area, these seasonal categories do not necessarily capture all of the region's microclimates but serve to capture its broad climate patterns.

To characterize net residential electricity consumption in SCE's service area, we create characteristic average daily load profiles that show the typical electricity load of a given set of customers on a given type of day. Throughout our analysis, we subset customers according to California Energy Commission Climate Zone (z) and electricity consumption decile bins (b). We subset customer-days by day of the week (i , e. weekend vs weekday) and season (i.e. JASO, MAMJ, JASO). Thus, average load profiles, $L(s, w, b, z)$, represent the mean electricity in each hour, L_h , from all customer-days for customers who lived in climate zone (z) with consumption in decile (b), on weekday/weekend (w) days in season (s) (Eqs. (8) & (9)).

$$L(s, w, b, z) = [L_0, L_1, \dots, L_h, \dots, L_{23}] \quad (8)$$

$$L_h = \frac{\sum_{i \in I'_{b,z}} \sum_{d \in D'_{s,w}} l_h(i, d)}{|D'_{s,w,b,z}|} \quad (9)$$

3.6. Gini coefficient calculation

To examine the distribution of electricity consumption among customers and quantify their relative contributions to the SCE aggregate residential electricity load, we derived average load profiles for each customer, i , denoted as $L(s, w, i)$, based on season (s) and day type (w). At each hour of the day, we use the typical customer profiles to construct a Lorentz curve [72] which shows the cumulative proportion of electricity consumed at hour, h , up to the i^{th} -ranked customer by $A(i)$ in ascending order $C_h(i)$:

$$C_h(i) = \frac{\sum_{j=1}^i l_h(s, w, c_j)}{\sum_{j=1}^n l_h(s, w, c_j)}$$

The Gini coefficient (GINI) is calculated using the differences between the Lorentz curve and a line of perfect equality [31] [See Supplemental 5]. We calculate the Gini coefficient as the:

$$\text{GINI}(h) = 1 - \sum_{i=1}^n (P_{u(i)} - P_{u(i-1)}) (C_h(i) + C_h(i-1))$$

In this case, the Gini coefficient is a value between 0 and 1, with 0 representing perfect equality (everyone consumes the same amount of electricity) and 1 representing perfect inequality (one customer consumes all electricity and no other customers consume any electricity).

4. Results

In this section, we describe the average residential load profiles by season and year. Because our data is statistically representative of the SCE service area, the shape and magnitude of the load curves are representative of the residential sector in SCE. (Fig. 1).

Weekend curves are similar in shape to weekday curves within the same season. The average daily load of weekends is slightly higher than that of weekdays within the same season. The majority of the differences between electricity use between weekdays and weekends occur during the 9 am through 5 pm workday, with 75% of the variability between the two curves in the hottest months and 62% in spring and winter months. During this 9 am to 5 pm period, there is higher weekend usage within the same season.

Like CAISO's all-sector electricity daily load profiles, the average daily load in the hottest months is higher than the average daily load in spring and winter months. In JASO, weekday average daily load is about 23.8 kWh across the period of study. By contrast, the MAMJ and NDJF weekday average daily load is 17.0 kWh, 28% lower than the JASO daily load. The seasonal load profiles are relatively similar in both shape and magnitude throughout the late evening and early morning periods (6 pm to 8 am) in all seasons. The biggest differences that drive the higher electricity in JASO occur during the mid-day, presumably due to cooling loads. In JASO, the residential load profile peaks at 17:00, two hours earlier than the 19:00 peak in both MAMJ and NDJF. This JASO peak hour electricity consumption is also significantly higher (1.47 kWh), with non-JASO peak hour consumption <1 kWh about 35% lower in electricity usage.

4.1. Average load profiles by electricity consumption decile bin

We analyzed diurnal electricity consumption profiles on weekdays in NDJF, JASO, and MAMJ by electricity consumption decile bins (Fig. 2). Each unique customer is assigned to the same decile bin in all seasons based on their average daily usage across the entire time period of study. Thus, some users may not be represented in all years. This figure shows that the differences in electricity usage across utility customers at each hour of the day can be large. For example, in JASO, customers in the highest electricity consumption decile bin consume 59.5 kWh of electricity per day, more than twice the average daily load of the average daily load of all users and >12 times the average daily load of the lowest electricity consumption decile bin.

The total daily load was highest in JASO for all decile bins of customers. Thus, while the impact on the diurnal shape of electricity consumption changes throughout seasons, the increase in the magnitude of total daily load remains a consistent trend for all decile bins. Most bins have about 40–50% lower weekday electricity consumption in non-JASO months compared to the JASO months. The exception is the smallest two deciles of customers where electricity differences hover around 15–20% lower weekday electricity consumption. This suggests that while there are seasonal differences in the magnitude of electricity consumption, the smallest customers (by electricity usage) exhibit less seasonal variability.

Generally, the average load profile of the middle 60% of customers is similar in shape to the average residential weekday load profile of all SCE customers. By contrast, the customers in the two highest consumption decile bins vary from lower decile bins. These large consumers have an average weekday load profile with a smooth singular peak in JASO and a double peak (i.e., one in mid-day and one in the early evening) in the Spring and Winter. For lower deciles, Spring and Winter load profiles generally consist of a smaller morning peak and a more pronounced evening peak. The timing of the peak hour shifts earlier in the day (16:00 and 17:00) for most decile bins in JASO, but the peak hour timing of the bottom three decile bins of customers stays consistent around hours 19:00 and 20:00 across the whole year. Customers in the bottom three decile bins have relatively flat profiles across seasons,

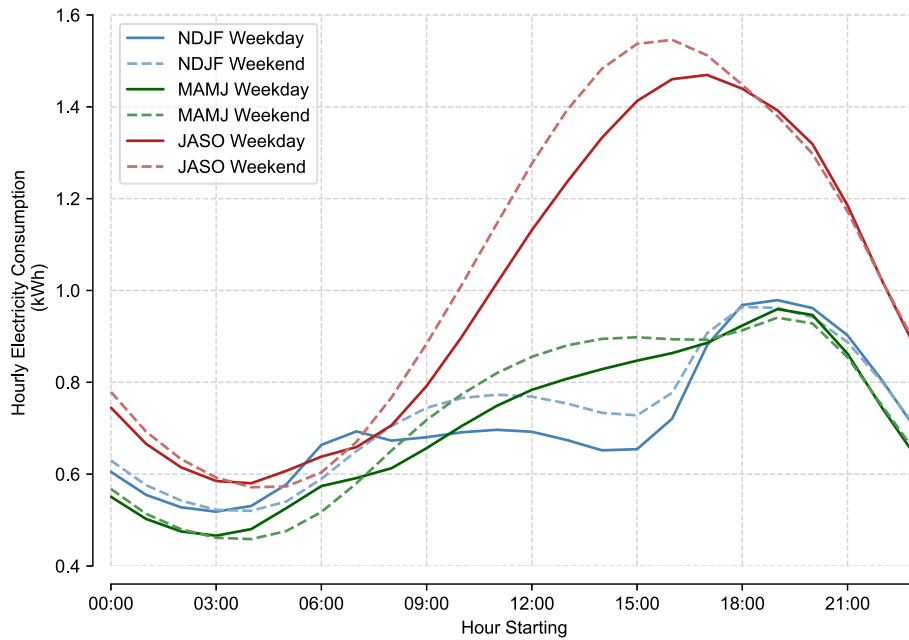


Fig. 1. The average daily load of a residential electricity customer in the SCE service area is highest during the hottest months (i.e., in JASO; July, August, September and October). Peak electricity usage occurs earlier and is much higher in hot months than other months. Daytime weekend usage is typically higher than weekday usage. Data represent the average hourly load profiles of approximately 160,000 SCE customers spanning the time period January 2015 through December 2019.

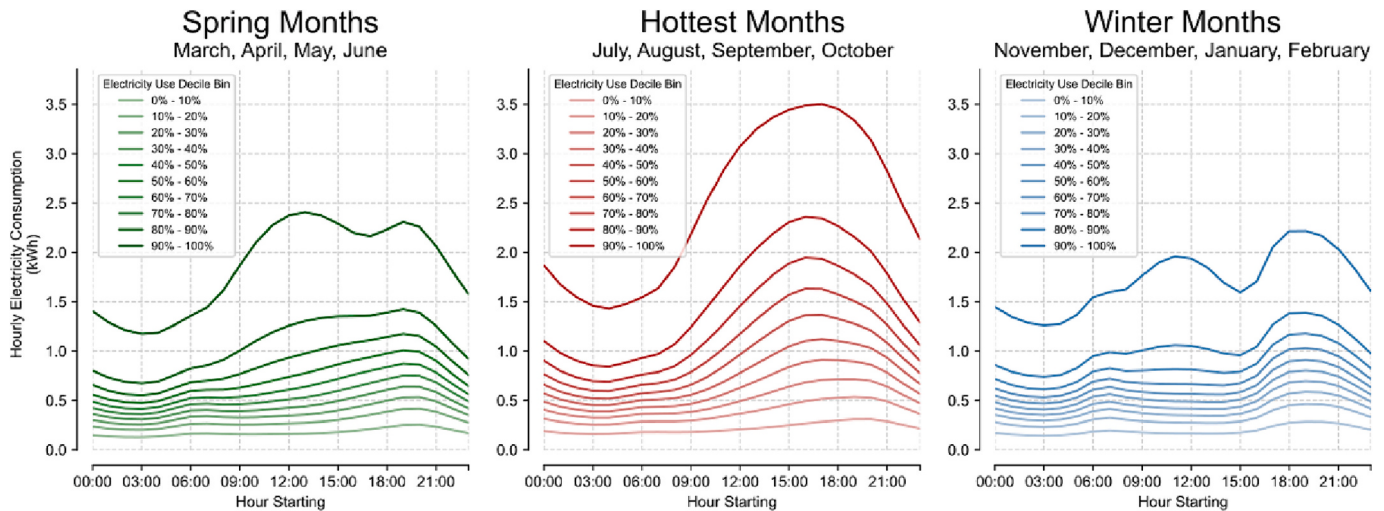


Fig. 2. The average load profiles of SCE customers segmented by electricity consumption decile bins show that the highest consuming customers represent a disproportionate fraction of daily load. Note: Each unique customer is assigned to the same decile bin in all seasons based on their average daily usage across the entire time period of study. Darker colors represent higher usage customers.

when compared with higher-usage customers. Like other customers, the magnitude of their average electricity usage is highest in JASO, but the difference in magnitude is smaller across seasons compared to higher usage customers.

4.2. Dispersion of electricity customers throughout the seasons

Next, we measure the diurnal patterns in disparities in electricity between the average electricity of customers. Cumulatively, the lower half of electricity customers ranked by annual electricity consumption consumed <25% of total JASO weekday usage. By contrast, the top 20 % of electricity users comprised >40% of all JASO weekday electricity usage.

The disparity of electricity usage is measured using the Gini

coefficient, which represents dispersion across a population on a scale from 0 to 1, changes by the time of day (Fig. 3). Higher values of GINI values represent higher dispersions across SCE customers (i.e. a smaller number of high-consumption customers consuming more than others). All three seasons generally show a similar pattern of low disparity between households in the late evening and early morning, and reach the highest dispersion levels at around midday (i.e., prior to afternoon peak hours). This result suggests that a relatively small proportion of high-consuming households account for a disproportionate amount of midday electricity usage across the year in SCE.

While the general shape of the diurnal disparity curve looks relatively similar across seasons, the magnitude of disparity shows large seasonal differences. The Gini coefficient is highest in magnitude during the mild spring months (green in Fig. 3), particularly on weekdays, most

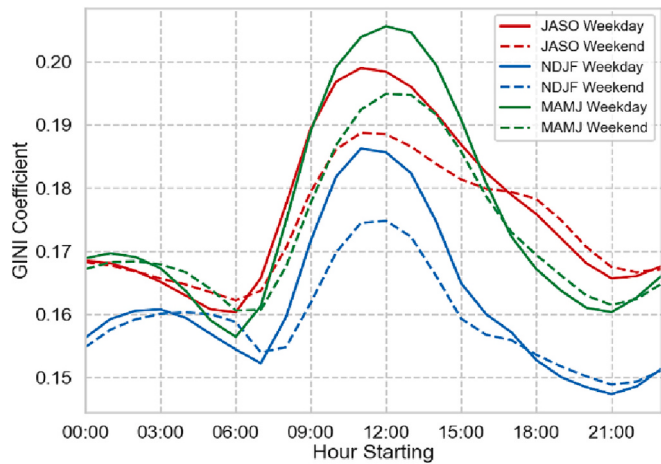


Fig. 3. The GINI Coefficient of electricity users suggests that the largest disparities in electricity usage across customers occur in the middle of the day, on weekdays, and in mild months. (Higher values of GINI values represent higher dispersions in electricity usage across SCE customers.)

likely due to large heterogeneities of HVAC usage and time spent at home compared to hotter and colder seasons. Cooling dominates midday loads across much of the customer base during the hottest months (red in Fig. 3), reducing midday dispersion across customers using AC compared to in mild weather months. The dispersion in the hottest months is likely bigger than the dispersion in the coldest months because space heating in California is dominated by natural gas units, which cannot be seen in these electricity data. Hence, the larger disparity in electricity consumption during hot months is being driven by large differences in residential cooling loads, which can be impacted by home size, home insulation and weatherization characteristics, AC unit technology, energy affordability concerns, etc. [73]. More generally, users across SCE demonstrate the widest range in loads in the midday during typical working days in every season. We conjecture that the disparity is higher on weekdays during typical working hours because there are bigger differences in time spent at home across populations of customers compared to weekends.

4.3. Climate zone differences in electricity use

SCE encompasses many different climate zones. Each climate zone represents a distinct geographic region, which encompasses not only differences in the climate that customers experience, but also differences

in housing stock, built infrastructure, and demographics. Cool climate zones include CZ 6 (coastal) and CZ 8 (inland coast). By contrast, the two climate zones in the valley, CZ 9 (Inland Valley) and CZ 10 (Interior Valley), are classified as moderate climate zones, with CZ 10 having summers that are classified as hot. CZ 14 (Low Desert) and CZ 15 (High Desert) are the hottest and driest climate zones in the region. The cool coastal climate zones maintain relatively similar load shapes year-round, with only minor increases in peak height and total daily electricity usage in the summer. (See Figure 4) By contrast, total daily electricity use during the hottest months significantly increases in moderate and hot climate zones, and peak-hour electricity shifts earlier in the afternoon during these hot periods. Decile breakdowns of residents within each climate zone can be found in Supplemental Information Fig. 6 and 7. In climate zones 10 and 14, the ratio between decile 10 (i.e., the highest-use customers) and decile 5 electricity loads is lower in the hottest months than compared to cooler months, suggesting that there is more diversity in electricity consuming behavior when loads are not being driven by cooling, which is consistent with the insights derived in the Gini coefficient analysis. In cooler climate zones, by contrast, the ratio between the decile 10 and decile 5 is higher in the hottest months, likely because there are large disparities in *if* and *how* households use AC during comparatively mild summers.

5. Discussion

This paper highlights that there is a big breakdown in the magnitude of electricity use across customers. We found that the top 20% of electricity customers use over 40% of the daily electricity load during weekdays in the hottest months while the bottom half of electricity use <25% of all net electricity consumption. During milder months (MAMJ), the largest consuming households show disproportionately high usage compared to the average customer, suggesting that these households have more discretionary loads. For Californian utilities such as SCE, most demand response efforts have targeted requests or incentives for load reductions during the period spanning 4 pm to 9 pm, when solar generation availability decreases and general residential electricity demand climbs towards peak net load. During our study period, the default rates for customers were standard tiered-rates based on monthly electricity usage with the ability to opt-in programs for time-of-use. Since 2021, customers have received time-of-use rates, with higher prices in the 4–9 pm period, by default and have the ability to opt-out to receive standard tiered rates [21]. Even so, during the hottest months of the year, the disproportionality of electricity use during the critical 4–9 pm period is greater in those hours compared to mild and winter months. Thus, during the hottest months of the year, large consumers use a

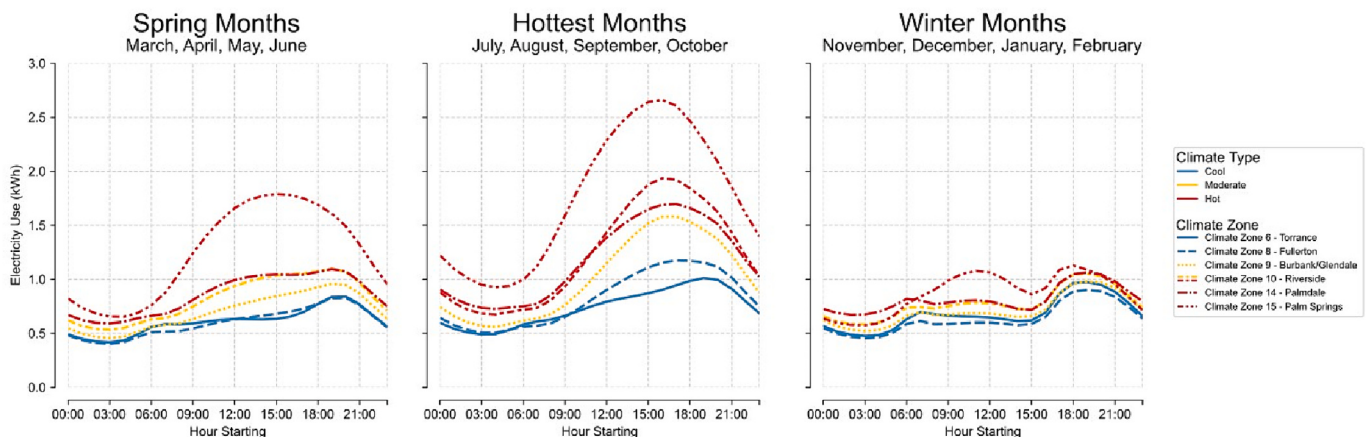


Fig. 4. Electricity load profiles differentiated by climate zone and season indicate that the hottest regions represent a disproportionate amount of electricity usage in SCE, particularly during the hottest months. Note: The hottest climate zones are denoted in red, moderate in yellow, and the coolest in blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

disproportionate amount of electricity use during the critical 4–9 pm period, suggesting that these customers are the most strategic for the utility to target for leveraging reductions during DR events. If utilities can get these customers to participate in DR activities, such as pre-cooling, that shift loads from these late afternoon hours (when the grid is dependent on expensive gas generators) to the midday (when the grid is served primarily by low-cost variable renewable energy generators), there will be valuable cost, emissions, and reliability benefits for the utility and customer base [74].

In general, higher electricity usage is associated with higher-income households and lower electricity with lower-income households, while use intensity (energy use that is normalized by household square footage) is generally higher in low-income households [75–78]. These findings approximately match our analysis, but as found in the aforementioned studies, these findings are not universally true. Within the top decile bin, about 16% of households were in census tracts with an average household income greater than \$150,000 [See Supplemental 8]. About 3.1% of households within the top decile bin of electricity use were in census tracts where the average household income was less than the living wage for an individual in California [79]. Thus, there likely exists some high electricity-consuming households with low household income who could be disproportionately impacted by new demand response initiatives, such as time-of-use rates.

However, designing rate structures requires not only balancing incentives for short-term load shifting, but also for social equity, public health, long-term technology adoption, and/or energy efficiency initiatives. Low-income households may have less efficient appliances and low insulation/weatherization interventions [80,81], higher energy loads because of higher occupancy [82], or require electricity-dependent medical equipment [83]. In the lowest decile bin, 13% of users lived in census tracts where the average household income was less than a living wage for an individual in California. These customers are likely to be the most disproportionately impacted by rate changes yet contribute little to the diurnal residential load profile, and potentially lack the flexibility to adjust their electricity-consuming behaviors during peak time. Recent work by Borenstein (2024) has found that a large proportion of the difference in consumption of those who would be in the top 20% of electricity use can be explained by the number of occupants in the house, gross consumption inclusive of distributed solar, differences in climate, electrified appliances, and the demographics of the household [84]. Many of these characteristics, Thus, while changes to pricing structures based on the amount of consumed electricity has the potential to benefit the grid, pricing structures also can be regressive in nature.

There are big trends in electricity use within California that could markedly shift the trends noted in this analysis over time. California has aggressive electrification efforts, including a goal to phase out gasoline powered vehicles by 2035. (The state currently leads the United States in electric vehicle registrations, with EV's constituting one-fifth of California's new car sales in 2023 [85,86]). EV chargers increase the electricity use of a household an average of between 3 and 10 kWh on the days vehicles are being charged, with higher power draws concentrated during specific hours [87,88]. Furthermore, the state plans to implement a ban on natural-gas appliances by 2030, with many municipalities already banning or discouraging the use of gas-fueled appliances, such as heaters, washers, dryers, and stoves. Currently the warmest months in SCE show disproportionate electricity consumption, but trends towards electrification will shift these trends diurnally and seasonally. For example, compared to the rest of the United States, Californian households are less likely to have appliances that use electricity with only 28% of space heaters (vs 40% nationwide) and 21% of water heaters (US: 47%) using electricity in 2023 [89]. Recent analysis of adoption of heat pumps in Texas showed that such changes can shift peak demand from the summer (driven by air-conditioning loads) to the winter (driven by heating loads) [90]. These appliances have the possibility of increasing and/or shifting the small morning peak currently seen in the Spring and

Winter load profiles.

As technologies such as rooftop solar and battery storage mature, designing rates and incentives for these technologies is becoming increasingly complicated. Small scale solar plays a growing role in the energy generation capacity in California, with a nameplate capacity of 11.7 GW at the end of 2023, and an estimated 19.2 TWh of annual generation. Rooftop solar installations in the residential sector generally sit behind the meter, and thus reduce the net electricity consumed from utilities in houses with PV panels installed. Poorly specified incentives and pricing structures for rooftop solar can put the benefits of such programs towards higher-consuming and higher-income customers while higher costs are absorbed by lower- to middle- income utility customers [30]. In California, Net Metering 3.0 which covers new installations of solar PV has also reduced the price paid for over-generated electricity (i.e., when solar generation exceeds electricity consumption for a given hour) for most solar PV customers, while increasing the potential benefits of Solar+Storage systems [91]. The use of smart meters has the potential to help the redesign of new pricing structures that can better balance the challenge of both equitable rates and incentivizing grid-efficient behavior. For example, in Denmark, smart meter data has been used to propose a new taxation policy that redistributes costs according to a household's consumption and PV usage to support both occupant-dense households and grid-side flexible prosumers [92].

Climate change will also affect household electricity use, particularly in regions that currently have relatively low AC penetration. AC penetration across the region is estimated at 69%, with the coastal areas of California having the lowest air conditioning penetration [68]. While our analysis shows that the cool, coastal climate zones (CZ6 and CZ8) have generally flat electricity load profiles, households (which tend to be wealthier than average) within these climate zones have the potential to grow significantly especially in the hottest months, due to relatively low rates of existing air conditioning penetration that will likely increase as temperature increases due to climate change [93]. (Even in these more mild climate zones, temperatures have recently hit record highs that have exceeded 100 degrees Fahrenheit [94]). Using data-driven methods, researchers and utilities will be able to track in near-real time electricity use behaviors.

Although smart meter data continues to grow more available, they are not yet ubiquitous and their distribution to third party researchers is often prohibited by utilities due to privacy concerns. However, as they become increasingly available, data-driven load profile studies like this one can offer critical insight into the actual distribution of electricity usage among customers, which cannot be generated from traditional bottom-up and top-down methods used to characterize residential load profiles. While physics-based models, like the Department of Energy's ResStock, utilize detailed conditional probability tables to characterize variables such as square footage, building construction, and heating and cooling appliances to generate regional load profiles across the United States for current and future scenarios [95], this benefit is constrained by the models' reliance on accurately representing the highly variable behaviors of individual users. Future research could integrate high-resolution smart meter data to validate the modeling of the residential load curve, especially for customers in different deciles of electricity use. Past research in customer behavior using customer segmentation and social science approaches have shown that customers may have dramatically different electricity use behaviors despite similar demographics, appliances, and building designs [45,69,96,97]. The results of this paper suggest that even when in the same service area, different customers have different load profiles because of differences such as by climate zones or electricity use bins. Therefore, data-driven models are essential in highlighting the divergences in behavioral patterns and the actual variances in electricity usage behaviors. Data-driven methods continue to grow in their capabilities (e.g., using smart meter data to deduce the presence and usage of appliances using non-intrusive load monitoring techniques [98,99]). As these methods improve, the insights from data-driven methods need to be integrated into bottom-up physics-

based models by providing more accurate representations of actual electricity consumption behaviors.

6. Conclusion

Developing insights into the patterns and magnitude of electricity usage in the residential sector is key for utility managers and policy makers to ensure the long-term reliability of the grid. To date, other studies developing load profiles of the residential sector have been limited by the size of available datasets. By using a statistically representative dataset of the Southern California Edison service area, we were able to characterize the load curves of statistically significant sub-populations of households across the SoCal Edison service area by characteristics such as total annual electricity usage, season, and climate zone. This work provides utility managers and policy makers in Southern California insights into how residential electricity is being used throughout the day and over different seasons.

This study provided a novel analysis of the contribution of groups of individual customers to the overall electricity curve and disparities in electricity use. We found that a small fraction of electricity users consume an outsized proportion of overall and peak electricity use. By contrast, the combined electricity use of the bottom half of household electricity consumers contributed to less than one-quarter of the electricity use. The disparities in electricity usage across SCE households are higher in the mid-day, especially in Spring since loads are not as dependent on high HVAC loads. Hotter months have higher disparities in electricity use during peak hours than cooler months because of wide differences in cooling-driven loads.

In the future, methods provided in this paper can give quick insights for utilities to understand how residential electricity may change across days, seasons, and climate zones, and how different groups across those spatio-temporal distinctions contribute to overall residential sector electricity usage. The region-specific trends identified in this study for Southern California are likely to change due to a warming climate, as well as climate mitigation policies promoting deep decarbonization and electrification [100,101]. However, researchers can build off the framework presented in this study to analyze diurnal changes in load profiles, both retrospectively to gain insight into how electricity load profiles may shift due to external circumstances (e.g. weather or local events) or to prospectively track trends and shifts in the residential load due to evolving factors affecting electricity consuming behaviors.

CRediT authorship contribution statement

Andrew S. Jin: Writing – original draft, Methodology, Formal analysis, Conceptualization. **Kelly T. Sanders:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgements

This work was funded in part by the National Science Foundation under grants CAREER Grant CBET-CAREER 1845931 the US Department of Defense Science, Mathematics, and Research for Transformation (SMART) Program, and an internal fellowship from USC Viterbi School of Engineering. Computation for the work described in this paper was

supported by the University of Southern California's Center for High-Performance Computing (hpcc.usc.edu). We also thank Southern California Edison for access to the smart meter data. The views herein do not necessarily reflect the views of any current or past employer.

Appendix A. Supplementary data

Supplementary information and figures to this article can be found online at <https://doi.org/10.1016/j.apenergy.2024.123316>.

References

- [1] International Energy Agency. World energy outlook 2023. 2024.
- [2] International Energy Agency. Renewables 2022. Renewable Energy Division in the Directorate of Energy Markets and Security; 2022.
- [3] US Energy Information Administration. Annual energy outlook 2023. 2024.
- [4] Kroposki B, Johnson B, Zhang Y, Gevorgian V, Denholm P, Hodge B-M, et al. Achieving a 100% renewable grid: operating electric power systems with extremely high levels of variable renewable Energy. *IEEE Power Energy Mag* 2017;15(2):61–73.
- [5] Bird L, Milligan M, Lew D. Integrating variable renewable energy: challenges and solutions. Golden, CO (United States): National Renewable Energy Lab. (NREL); Sep. 2013. NREL/TP-6A20-60451.
- [6] California Independent System Operator. Understanding the ISO [Online]. Available: <https://www.caiso.com/about/Pages/OurBusiness/Default.aspx>; 2024. Accessed: 25-Nov-2023.
- [7] California Independent System Operator. 2022 annual report on market issues and performance. Jul. 2023.
- [8] CAISO. What the duck curve tells us about managing a green grid [Online]. Available: https://www.caiso.com/Documents/FlexibleResourcesHelpRenewables_FastFacts.pdf; 2016. Accessed: 30-Aug-2023.
- [9] Denholm P, O'Connell M, Brinkman G, Jorgenson J. Overgeneration from solar energy in California. A field guide to the duck chart. Golden, CO (United States): National Renewable Energy Lab (NREL); Nov. 2015. NREL/TP-6A20-65023.
- [10] As solar capacity grows, duck curves are getting deeper in California [Online]. Available: <https://www.eia.gov/todayinenergy/detail.php?id=56880>; 2024 [Accessed: 07-Sep-2023].
- [11] Balasubramanian S, Balachandra P. Effectiveness of demand response in achieving supply-demand matching in a renewables dominated electricity system: a modelling approach. *Renew Sust Energ Rev* Sep. 2021;147:111245.
- [12] Gelazanskas L, Gamage KAA. Demand side management in smart grid: a review and proposals for future direction. *Sustain Cities Soc* Feb. 2014;11:22–30.
- [13] Reports on demand response and advanced metering | federal energy regulatory commission [Online]. Available: <https://www.ferc.gov/power-sales-and-markets/demand-response/reports-demand-response-and-advanced-metering>; 2024 [Accessed: 08-Sep-2023].
- [14] Albadi MH, El-Saadany EF. Demand response in electricity markets: an overview. In: 2007 IEEE power engineering society general meeting; 2007. p. 1–5.
- [15] Lee M. State OKs higher electric rates for most consumers. *San Diego Union-Tribune*; 2015.
- [16] Herter K, Wayland S. Residential response to critical-peak pricing of electricity: California evidence. *Energy* Apr. 2010;35(4):1561–7.
- [17] Decision adopting policy guidelines to assess time periods for future time-of-use rates and energy resource contract payment. vol. 2017; 2017. p. 19*.
- [18] George S, Bell E, Savage A, Messer B. California statewide opt-in time-of-use pricing pilot. Nexant Inc; 2024.
- [19] Brown TK. PG&E is changing how your electricity bill works — here's everything you need to know. *San Francisco Chronicle* 2021 [Online]. Available: <https://www.sfchronicle.com/local/environment/article/MicroClimates-PG-E-is-changing-how-your-16037469.php>. Accessed: 07-Sep-2023.
- [20] Nikolewski Rob. "SDG&E rolling out 'time of use' rates next month and what it means for you," *San Diego Union-Tribune*. 2019 [Online]. Available: <https://www.sandiegouniontribune.com/business/energy-green/sd-fi-sdge-timeofuse-20190221-story.html>. Accessed: 07-Sep-2023.
- [21] Lazarus David. Column: How you pay for electricity is changing. Here's what you need to know. *Los Angeles Times* 2021 [Online]. Available: <https://www.latimes.com/business/story/2021-09-24/column-edison-energy-rate-switch>. Accessed: 07-Sep-2023.
- [22] Linder KF, Dickson CT, Inglis MR. Hourly Electric-Load Model (HELM). Volume 1. Design, development, and demonstration. Final report. Washington, DC (USA): ICF, Inc; Jan. 1985. EPRI-EA-3698-Vol.1.
- [23] Baroiant S, Barnes J, Chapman D, Keates S, Phung J. California investor-owned utility electricity load shapes. 2024.
- [24] Siano P. Demand response and smart grids—a survey. *Renew Sust Energ Rev* Feb. 2014;30:461–78.
- [25] Burger SP, Knittel CR, Perez-Arriaga IJ, Schneider I, vom Scheidt F. The efficiency and distributional effects of alternative residential electricity rate designs. *Energy J* Jan. 2020;41(1):199–240.
- [26] Borenstein S, Bushnell JB. Do two electricity pricing wrongs make a right? Cost recovery, externalities, and efficiency. *National Bureau of Economic Research*; Jun-2018.

- [27] California Energy Commission. California distributed generation statistics [Online]. Available: <https://www.californiadgstats.ca.gov/>; 2024 [Accessed: 08-Apr-2024].
- [28] U.S. Energy Information Administration. Form EIA-861M [Online]. Available: <https://www.eia.gov/electricity/data/eia861m/index.php>; 2024. Accessed: 08-Apr-2024.
- [29] Lemay AC, Wagner S, Rand BP. Current status and future potential of rooftop solar adoption in the United States. *Energy Policy* Jun. 2023;177:113571.
- [30] Borenstein S. Private net benefits of residential solar PV: the role of electricity tariffs, tax incentives, and rebates. *J Assoc Environ Resour Econ* Sep. 2017;4(S1):S85–122.
- [31] Gastwirth JL. The estimation of the Lorenz curve and Gini index. *Rev Econ Stat* 1972;54(3):306–16.
- [32] Beccali M, Cellura M, Lo Brano V, Marvuglia A. Forecasting daily urban electric load profiles using artificial neural networks. *Energy Convers Manag* Nov. 2004;45(18):2879–900.
- [33] Hippert HS, Bunn DW, Souza RC. Large neural networks for electricity load forecasting: are they overfitted? *Int J Forecast* Jul. 2005;21(3):425–34.
- [34] Sandels C, Widén J, Nordström L. Forecasting household consumer electricity load profiles with a combined physical and behavioral approach. *Appl Energy* Oct. 2014;131:267–78.
- [35] Lindberg KB, Seljom P, Madsen H, Fischer D, Korpås M. Long-term electricity load forecasting: current and future trends. *Util Policy* Jun. 2019;58:102–19.
- [36] Phurailatpam C, Rajpurohit BS, Wang L. Planning and optimization of autonomous DC microgrids for rural and urban applications in India. *Renew Sust Energ Rev* Feb. 2018;82:194–204.
- [37] Iria J, Heleno M, Cardoso G. Optimal sizing and placement of energy storage systems and on-load tap changer transformers in distribution networks. *Appl Energy* Sep. 2019;250:1147–57.
- [38] Hajiaghapour-Moghami M, Azimi-Hosseini K, Hajipour E, Vakilian M. Residential load clustering contribution to accurate distribution transformer sizing. In: 2019 International Power System Conference (PSC); 2019. p. 313–9.
- [39] Chen CS, Hwang JC, Huang CW. Application of load survey systems to proper tariff design. *IEEE Trans Power Syst* Nov. 1997;12(4):1746–51.
- [40] Ramos S, Vale Z, Santana J, Rodrigues F. Use of data mining techniques to characterize mv consumers and to support the consumer-supplier relationship. 2024.
- [41] Mahmoudi-Kohan N, Moghaddam MP, Sheikh-El-Eslami MK. An annual framework for clustering-based pricing for an electricity retailer. *Electr Power Syst Res* Sep. 2010;80(9):1042–8.
- [42] Panapakidis IP, Alexiadis MC, Papagiannis GK. Load profiling in the deregulated electricity markets: A review of the applications. In: 2012 9th international conference on the european energy market; 2012. p. 1–8.
- [43] US Energy Information Administration. Electric power annual 2021. 2022.
- [44] Kavousian A, Rajagopal R, Fischer M. Determinants of residential electricity consumption: using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior. *Energy* Jun. 2013;55:184–94.
- [45] Socolow RH. The twin rivers program on energy conservation in housing: highlights and conclusions. *Energy Buildings* Apr. 1978;1(3):207–42.
- [46] Swan LG, Ugursal VI. Modeling of end-use energy consumption in the residential sector: a review of modeling techniques. *Renew Sust Energ Rev* Oct. 2009;13(8):1819–35.
- [47] Brandenburg L, Clarkson G, Grund Jr C, Leo J, Asbury J, Brandon-Brown F, et al. Load research manual. Volume 2. Fundamentals of implementing load research procedures. ANL/SPG-13(Vol.2). Nov. 1980. p. 6705684.
- [48] Proedrou E. A comprehensive review of residential electricity load profile models. *IEEE Access* 2021;9:12114–33.
- [49] Grandjean A, Adnot J, Binet G. A review and an analysis of the residential electric load curve models. *Renew Sust Energ Rev* Dec. 2012;16(9):6539–65.
- [50] Kang X, An J, Yan D. A systematic review of building electricity use profile models. *Energy Buildings* Feb. 2023;281:112753.
- [51] Parkash B, Lie TT, Li W, Tito SR. Hierarchical structure based energy consumption forecasting in top-down approach. In: Proceedings - 2022 7th Asia conference on power and electrical engineering, ACPEE 2022; 2022. p. 1732–7.
- [52] Fischer D, Surmann A, Lindberg KB. Impact of emerging technologies on the electricity load profile of residential areas. *Energy Buildings* Feb. 2020;208:109614.
- [53] Parti M, Parti C. The total and appliance-specific conditional demand for electricity in the household sector. *Bell J Econ* 1980;11(1):309–21.
- [54] Aigner DJ, Sorooshian C, Kerwin P. Conditional demand analysis for estimating residential end-use load profiles. *Energy* J Jul. 1984;5(3).
- [55] Kavgić M, Mavrogiani A, Mumovic D, Summerfield A, Stevanovic Z, Djurovic-Petrovic M. A review of bottom-up building stock models for energy consumption in the residential sector. *Build Environ* Jul. 2010;45(7):1683–97.
- [56] Capasso A, Grattieri W, Lamedica R, Prudenzi A. A bottom-up approach to residential load modeling. *IEEE Trans Power Syst* May 1994;9(2):957–64.
- [57] US Department of Energy. Advanced metering infrastructure and customer systems. 2016.
- [58] Every J, Li L, Dorrell D. Leveraging smart meter data for economic optimization of residential photovoltaics under existing tariff structures and incentive schemes. *Appl Energy* Sep. 2017;201:158–73.
- [59] Yildiz B, Bilbao JI, Dore J, Sproul AB. Recent advances in the analysis of residential electricity consumption and applications of smart meter data. *Appl Energy* Dec. 2017;208:402–27.
- [60] Wang Y, Chen Q, Hong T, Kang C. Review of smart meter data analytics: applications, methodologies, and challenges. *IEEE Trans Smart Grid* May 2019;10(3):3125–48.
- [61] Mukai T, Nishio K, Komatsu H, Kobayashi K, Sasaki M, Maeki W. Smart meter-based home energy report: design and operation of automatic generation system of hourly usage indication message. *Electron Commun Jpn* Mar. 2021;104(1):74–86.
- [62] Sobrino E, Santiago A, Gonzalez A, IEEE. Forecasting the electricity hourly consumption of residential consumers with smart meters using machine learning algorithms. In: Presented at the 2019 IEEE Milan Powertech; 2019.
- [63] Andersen F, Gunkel P, Jacobsen H, Kitzing L. Residential electricity consumption and household characteristics: An econometric analysis of Danish smart-meter data. *Energy Econ* Aug. 2021;100.
- [64] MacMackin N, Miller L, Cariveau R. Modeling and disaggregating hourly effects of weather on sectoral electricity demand. *Energy* Dec. 2019;188.
- [65] Sanchez-Lopez M, Moreno R, Alvarado D, Suazo-Martinez C, Negrete-Pincetic M, Olivares D, et al. The diverse impacts of COVID-19 on electricity demand: the case of Chile. *Int J Electr Power Energy Syst* Jun. 2022;138.
- [66] Novan K, Smith A. The incentive to overinvest in energy efficiency: evidence from hourly smart-meter data. *J Assoc Environ Resour Econ* Jul. 2018;5(3):577–605.
- [67] Ushakova A, Mikhaylov S. Big data to the rescue? Challenges in analysing granular household electricity consumption in the United Kingdom. *Energy Res Soc Sci* Jun. 2020;64.
- [68] Chen M, Sanders KT, Ban-Weiss GA. A new method utilizing smart meter data for identifying the existence of air conditioning in residential homes. *Environ Res Lett* 2019;14(9).
- [69] Kwac J, Flora J, Rajagopal R. Household energy consumption segmentation using hourly data. *IEEE Trans Smart Grid* Jan. 2014;5(1):420–30.
- [70] Zethmayr J, Makhija RS. Six unique load shapes: a segmentation analysis of Illinois residential electricity consumers. *Electr J Nov.* 2019;32(9):106643.
- [71] Gunkel PA, Klinge Jacobsen H, Bergaentzle C-M, Scheller F, Möller Andersen F. Variability in electricity consumption by category of consumer: the impact on electricity load profiles. *Int J Electr Power Energy Syst* May 2023;147:108852.
- [72] Gastwirth JL. A general definition of the Lorenz curve. *Econometrica* 1971;39(6):1037–9.
- [73] Hernández D, Bird S. Energy burden and the need for integrated low-income housing and energy policy. *Poverty Public Policy* 2010;2(4):5–25.
- [74] Mayes S, Zhang T, Sanders KT. Residential precooling on a high-solar grid: impacts on CO2 emissions, peak period demand, and electricity costs across California. *Environ Res Energy Oct.* 2023;1(1):015001.
- [75] Liddle B, Huntington H. How prices, income, and weather shape household electricity demand in high-income and middle-income countries. *Energy Econ* Mar. 2021;95:104995.
- [76] Chen C, Xu X, Adua L, Briggs M, Nelson H. Exploring the factors that influence energy use intensity across low-, middle-, and high-income households in the United States. *Energy Policy* Sep. 2022;168:113071.
- [77] Alberini A, Gans W, Velez-Lopez D. Residential consumption of gas and electricity in the U.S.: the role of prices and income. *Energy Econ* Sep. 2011;33(5):870–81.
- [78] U.S. Energy Information Administration. U.S. energy insecure households were billed more for energy than other households [Online]. Available: <https://www.eia.gov/todayinenergy/detail.php?id=56640>; 2024. Accessed: 29-Jan-2024.
- [79] Nadeau CA. MIT living wage calculator. 2023.
- [80] Xu X, Chen C. Energy efficiency and energy justice for U.S. low-income households: An analysis of multifaceted challenges and potential. *Energy Policy* May 2019;128:763–74.
- [81] Chen C, Xu X, Day JK. Thermal comfort or money saving? Exploring intentions to conserve energy among low-income households in the United States. *Energy Res Soc Sci* Apr. 2017;26:61–71.
- [82] Brown MA, Soni A, Lapsa MV, Southworth K, Cox M. High energy burden and low-income energy affordability: conclusions from a literature review. *Prog Energy* Oct. 2020;2(4):042003.
- [83] Casey JA, Mango M, Mullendore S, Kiang MV, Hernández D, Li BH, et al. Trends from 2008–2018 in electricity-dependent durable medical equipment rentals and sociodemographic disparities. *Epidemiol Camb Mass* May 2021;32(3):327–35.
- [84] Borenstein S. Energy hogs and Energy angels: what does residential electricity usage really tell us about profligate consumption?. In: National Bureau of Economic Research; Jan 2024.
- [85] Bloomberg DH. Electric cars now make up a fifth of California's auto sales. *Los Angeles Times* 2023 [Online]. Available: <https://www.latimes.com/business/story/2023-11-01/electric-cars-now-make-up-a-fifth-of-californias-auto-sales>. Accessed: 04-Jan-2024.
- [86] California leads the United States in electric vehicles and charging locations [Online]. Available: <https://www.eia.gov/todayinenergy/detail.php?id=61082>; 2024. Accessed: 04-Jan-2024.
- [87] Muratori M. Impact of uncoordinated plug-in electric vehicle charging on residential power demand. *Nat Energy* Mar. 2018;3(3):193–201.
- [88] Burlig F, Bushnell J, Rapson D, Wolfram C. Low energy: estimating electric vehicle electricity use. *AEA Pap Proc* May 2021;111:430–5.
- [89] Davis L. Three facts about electric heating in California. *Energy Inst Blog*. <https://energythaas.wordpress.com/2023/05/08/three-facts-about-electric-heating-in-california/>; 2023.
- [90] White PR, Rhodes JD, Wilson EJH, Webber ME. Quantifying the impact of residential space heating electrification on the Texas electric grid. *Appl Energy* Sep. 2021;298:117113.

- [91] Customer generation [Online]. Available: <https://www.cpuc.ca.gov/industries-and-topics/electrical-energy/demand-side-management/customer-generation>; 2024 [Accessed: 08-Apr-2024].
- [92] Gunkel PA, Kachirayil F, Bergaentzle C-M, McKenna R, Keles D, Jacobsen HK. Uniform taxation of electricity: incentives for flexibility and cost redistribution among household categories. *Energy Econ* Nov. 2023;127:107024.
- [93] Chen M, Ban-Weiss GA, Sanders KT. Utilizing smart-meter data to project impacts of urban warming on residential electricity use for vulnerable populations in Southern California. *Environ Res Lett* 2020;15(6).
- [94] Yee G. Temperature records broken across Los Angeles area at end of spring heat wave. *Los Angeles Times* 2022 [Online]. Available: <https://www.latimes.com/california/story/2022-04-08/record-high-temperatures-los-angeles-heat-wave>. Accessed: 04-Jan-2024.
- [95] Wilson EJH, Parker A, Fontanini A, Present E, Reyna JL, Adhikari R, et al. End-use load profiles for the U.S. building stock: methodology and results of model calibration, validation, and uncertainty quantification. Golden, CO (United States): National Renewable Energy Lab. (NREL); Mar. 2022. NREL/TP-5500-80889.
- [96] Firth S, Lomas K, Wright A, Wall R. Identifying trends in the use of domestic appliances from household electricity consumption measurements. *Energy Buildings* Jan. 2008;40(5):926–36.
- [97] Rajabi A, Eskandari M, Ghadi MJ, Li L, Zhang J, Siano P. A comparative study of clustering techniques for electrical load pattern segmentation. *Renew Sust Energy Rev* Mar. 2020;120:109628.
- [98] Schirmer PA, Mporas I. Non-intrusive load monitoring: a review. *IEEE Trans Smart Grid* Jan. 2023;14(1):769–84.
- [99] Abubakar I, Khalid SN, Mustafa MW, Shareef H, Mustapha M. An overview of Non-intrusive load monitoring methodologies. In: 2015 IEEE conference on energy conversion (CENCON); 2015. p. 54–9.
- [100] Gold R. Status report on electrification policy: where to next? *Curr Sustain Energy Rep* Jun. 2021;8(2):114–22.
- [101] Clean Energy. Jobs, and affordability act of 2022. 2022.