ENVIRONMENTAL RESEARCH
ENERGY

PAPER • OPEN ACCESS

Analyzing how the timing and magnitude of electricity consumption drive variations in household electricity-associated emissions on a high-VRE grid

To cite this article: Stepp Mayes *et al* 2024 *Environ. Res.: Energy* **1** 045010

IOP Publishing Environ. Res.: Energy 1 (2024) 045010 https://doi.org/10.1088/2753-3751/ad8bc6

ENVIRONMENTAL RESEARCH
ENERGY

CrossMark

OPEN ACCESS

RECEIVED
10 July 2024

REVISED
26 September 2024

ACCEPTED FOR PUBLICATION
28 October 2024

PUBLISHED
12 November 2024

PAPER

Analyzing how the timing and magnitude of electricity consumption drive variations in household electricity-associated emissions on a high-VRE grid

Stepp Mayes, McKenna Peplinski and Kelly T Sanders*

Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, CA 90089, United States of America

* Author to whom any correspondence should be addressed.

E-mail: ktsanders@usc.edu

Keywords: household electricity, electricity emissions, emissions factors, air conditioning

Supplementary material for this article is available online

Abstract

Electrifying the residential sector is critical for national climate change adaptation and mitigation strategies, but increases in electricity demand could drive-up emissions from the power sector. However, the emissions associated with electricity consumption can vary depending on the timing of the demand, especially on grids with high penetrations of variable renewable energy. In this study, we analyze smart meter data from 2019 for over 100 000 homes in Southern California and use hourly average emissions factors from the California Independent System Operator, a high-solar grid, to analyze household CO_2 emissions across spatial, temporal, and demographic variables. We calculate two metrics, the annual household electricity-associated emissions (annual-HEE), and the household average emissions factor (HAEF). These metrics help to identify appropriate strategies to reduce electricity-associated emissions (i.e. reducing demand vs leveraging demand-side flexibility) which requires consideration of the magnitude and timing of demand. We also isolate the portion of emissions caused by AC, a flexible load, to illustrate how a load with significant variation between customers results in a large range of emissions outcomes. We then evaluate the distribution of annual-HEE and HAEF across households and census tracts and use a multi-variable regression analysis to identify the characteristics of users and patterns of consumption that cause disproportionate annual-HEE. We find that in 2019 the top 20% of households, ranked by annual-HEE, were responsible for more emissions than the bottom 60%. We also find the most emissions-intense households have an HAEF that is 1.7 times higher than the least emissions-intense households, and that this spread increases for the AC load. In this analysis, we focus on Southern California, a demographically and climatically diverse region, but as smart meter records become more accessible, the methods and frameworks can be applied to other regions and grids to better understand the emissions associated with residential electricity consumption.

1. Introduction

In the United States, roughly one-third of all CO_2 emissions are produced through electricity generation [1] and approximately 39% of electricity consumption occurs in the residential sector [2]. Therefore, the emissions associated with residential electricity consumption are a significant portion of national greenhouse gas emissions and a target of mitigation strategies designed to offset emissions. At the same time, households are likely to experience new electricity demand needs as part of both climate change mitigation (e.g. heat pumps and EVs) and adaptation (e.g. increased need for AC in response to extreme heat). Household-level electricity consumption depends on many factors, including weather and climate patterns, building envelope characteristics, HVAC technologies, behavioral patterns, and more [3–5]. However, the location of a household and the timing of its demand also impact the emissions associated with its electricity consumption because regional grids have different grid mixes with generation sources that vary throughout the day [6, 7]. The timing of demand is particularly consequential for emissions on grids with high penetrations of variable renewable energy (VRE) (wind and solar power), where the average emissions factor (AEF) (which captures the emissions per unit of electricity demanded) can vary significantly throughout the day and year. Therefore, a combination of high-temporal resolution electricity consumption data and emissions factors is necessary to estimate and analyze the emissions associated with household-level electricity consumption.

A household's electricity-associated CO_2 emissions (referred to as HEE in this study) are interesting for many reasons. First, understanding the spatial and temporal distribution of total emissions, as well as the varying emissions-intensity of a household's electricity consumption, can inform programs designed to reduce HEE, a necessary aspect of climate change mitigation. For grids with high penetrations of fossil fuels, strategies for reducing emissions typically focus on reducing overall demand, such as improving energy efficiency or behavioral interventions. As grids adopt higher penetrations of VRE resources, demand-side management programs that encourage customers to shift their demand to hours with low carbon resource generation can have positive outcomes for both grid reliability and emissions reductions [8]. On a grid where demand and emissions can be highly decoupled, we are interested in identifying what factors cause households with high HEE to be high emitting. These households may have high demand, a more emissions-intense pattern of electricity consumption, or a combination of both, and evaluating how these aspects are contributing to emissions can help to inform which emissions-reducing strategies (i.e. load-shifting vs demand reduction) would be better suited to reduce HEE.

The distribution of HEE is also interesting from a climate-equity perspective. Previous research has shown that higher socioeconomic status households consume a disproportionate amount of electricity [9–12], but incorporating hourly emissions factors allows us to quantify the disparity in emissions, and determine if higher-income homes also have more emissions-intense consumption. Many analyses in this area have been limited by a lack of access to large-scale, household level data, without which it is difficult to study how different populations contribute to emissions. Ensuring climate change mitigation is done equitably requires a thorough understanding of the conditions that cause households to be high emitting.

Third, by analyzing residential building electricity emissions, we can identify specific loads and patterns of consumption that are emissions intense. Building energy modelers often divide household electricity consumption into baseload electricity consumption (which captures loads that have limited dependence on time and temperature), HVAC loads, and variable loads [13]. Demand that is non-baseload can have specific temporal patterns, meaning it may occur in hours with particularly high, or low, levels of emissions intensity. For example, AC use can be concentrated in particular months of the year and hours of the day, but it is also a load that can be flexible or shed for short time periods [14]. Understanding the emissions associated with a time-dependent load like AC could inform if that load is a good target for an emissions-reduction strategy, such as improving efficiency or temporal load shifting. This flexibility may become increasingly important as AC ownership and use are expected to increase as part of climate change adaptation.

Southern California provides an interesting context in which to examine HEE. First, the region contains a diverse group of sociodemographic populations, a mix of urban, suburban, and rural areas, and many micro-climate zones. Furthermore, the grid in our study region, which is overseen by the California Independent System Operator (CAISO), is characterized by high penetrations of solar power and moderate penetrations of wind and hydro power. As a result, the mix of resources on the grid is highly dependent on time of day and year, and thus the emissions-intensity of electricity generation (the emissions per unit of energy generated) is highly time dependent [15]. The decoupling of grid emissions and generation creates a non-linear relationship between emissions and *demand*, and therefore analysis of HEE must consider both the magnitude and timing of household electricity consumption. The methods used in this study to analyze HEE can be applied on all grids, but it should be noted that many of the results and implications are CAISO-specific. Other electric grids will have distinct generation trends that impact whether a specific electricity consuming behavior, or pattern of behavior, occurs during a period of emissions-intense generation or not.

In this study, we use a smart meter dataset of hourly electricity consumption for approximately 107 000 homes in Southern California that spans over 2000 census tracts and seven building climate zones. We match the hourly electricity records with grid-level hourly emissions factors to analyze the associated emissions. We also introduce a novel metric, the household AEF (HAEF) that captures the emissions intensity of a household's electricity consumption during a period of time. Finally, we utilize an established methodology to identify electricity consumption for air-conditioning [16, 17] and analyze the emissions associated with AC use. We address three primary research questions in this work:

1. What are the CO_2 emissions associated with household electricity consumption in Southern California, and how do these vary across census tracts, climate zones, and sociodemographic status?
2. How does the emissions-intensity of electricity consumption vary from household to household?
3. Are AC-related emissions a significant portion of total residential building emissions in Southern California, and is electricity consumed for AC more, or less, emissions intense than electricity consumed for other activities?

This analysis reveals inequalities in emissions from the residential sector and can inform policy makers and researchers on how to think about electricity consumption and emissions on dynamic, high-renewable grids. It is important that we implement DSM strategies to ensure people's electricity and cooling needs are met, while minimizing the associated emissions. The results of this analysis can be used to develop a robust understanding of the emissions patterns of customers and inform these strategies. While some of the insights of this study may be unique to CAISO, which has extremely high solar penetration, this framework can be applied to all electric grids to inform region-specific demand-side programs.

2. Literature review

Studies that quantify emissions from the residential sector often rely on indirect estimates of emissions for entire regions due to challenges with obtaining large, highly-resolved datasets on household-level energy consumption [18–20]. This lack of direct measurements often limits the spatial and temporal resolution of the analysis but can be useful for capturing trends across regions or countries. A review by Geng *et al* observed that many studies used a version of input–output analysis, to estimate residential emissions at the country, regional, or state level [18]. Input–output models convert spending in a particular sector into emissions and are therefore unable to capture how individual electricity consumption patterns can affect emissions through their timing and location. For example, Baynes *et al* compared two methods of estimating regional residential sector emissions, one of which was an input–output model based on sectors in Australia that estimated emissions from aggregate expenditures [21].

For studies that have focused on emissions at finer scales, such as at the household or census-tract level, researchers have used a large variety of building and economic models in place of direct measures of electricity consumption [12, 22–25]. Lyons *et al* used an input–output model at the household level for 6884 residences in Ireland to estimate total energy consumption and the resulting emissions based on household spending in twenty sectors. They found that richer households generally emit more than those of lower income [12]. Mattinen *et al* used a bottom–up modeling approach in Finland to get spatial estimates of household CO_2 emissions, but developed estimates of energy consumption from household characteristics rather than direct measurements [22]. Goldstein *et al* addressed the complex relationships between, energy efficiency, CO_2 emissions, and racial energy inequity, but used regression models to estimate electricity consumption based on building characteristics [25]. Danylo *et al* take a downscaling approach with regional demand data in Ukraine and Poland, using proxy variables to estimate the spatial distribution of household-level emissions [23]. While these models and proxy variables can provide some insight into emissions trends and patterns, evaluating details such as the relationship between the timing of demand and the associated emissions requires accurate, user-specific data.

Another set of studies use survey data or utility bills, rather than modeling, as an estimate for electricity demand and then analyze emissions at the household or census-tract level, but these studies are still often limited by a lack of access to large datasets. Further, these studies calculate emissions using electricity consumption at coarse temporal resolutions and a constant emissions intensity term, which does not encompass how emissions factors vary diurnally or seasonally. Studies that use electricity billing typically captures electricity consumption at the monthly [26, 27] or yearly level [28]. This approach then requires complete knowledge of local pricing and rate schemes to give an accurate estimate of electricity consumption, and is still limited by the monthly resolution, which makes it impossible to analyze more refined temporal trends in emissions. In addition to the coarse temporal resolution for emissions calculations, many of these studies are limited by data access and/or sample-size. For example, Wilson relied on self-reported bills [27] and Vandeweghe had access to only census-tract averaged bills rather than individual bills [28]. In the case of [26, 27], the studies were also limited by having records for only a few thousand homes. Nassen *et al* used actual electricity meter records and estimated both direct and indirect residential building CO_2 emissions, but they were only able to collect electricity consumption data for 215 houses, with the remaining households' electricity consumption was estimated via regression [29].

Typically, analyses of household electricity emissions rely on emissions factors with low temporal resolution, such as a single annual emissions factor for a study region [26–28, 30–32]. Researchers often utilize these coarse emissions factors because the corresponding electricity dataset is also not highly resolved.

In other instances, annual emissions factors are used to simplify the calculations and analysis, with the assumption that it describes the emissions intensity of the electricity demand well enough to uncover general trends. For example, Papachristos used a bottom–up approach to estimate the impact that installing smart meters in households would have on electricity related CO_2 emissions [30]. He first determined the impact of smart meters on individual household appliances and technologies and then multiplied the change in electricity by an annual emissions factor. While using an annual emissions factor may have been a reasonable approximation for the specific study area and goal of that analysis, household technologies can follow very specific patterns of use that depend on the time of day and time of year, and hourly emissions factors may change the estimate of electricity-associated emissions significantly, especially on high-renewable grids.

The importance of using emissions factors with high temporal resolution has been established in other studies that focus on specific aspects of residential emissions. Karimu *et al* and Nilsson *et al* both studied the relationship between household electricity emissions and real-time-pricing using season-hour and hourly emissions factors respectively [33, 34]. Nilsson *et al* found that real-time pricing could actually lead to increased emissions in Sweden, a country that relies on emissions intense imports from neighboring countries and has large diurnal variations in hydropower availability, because there is almost no correlation between the price and emissions. On the other hand, Karimu *et al* found that real-time pricing had limited impact on demand or emissions, but it is important to note that the relationship between emissions and generation cost will be highly grid-dependent, and thus real-time pricing may increase or decrease emissions depending on the regional grid. Miller *et al* studied residential emissions using a variety of hypothetical building end-use profiles that were created by the National Renewable Energy Laboratory to represent the range of residential electricity consumption patterns. They calculated CO_2 emissions using emissions factors at different resolutions (e.g. annual vs month–hour) and determined the level of bias created by choice of emissions factor [35]. Temporally-resolved emissions factors reveal that on many grids, the relationship between electricity demand, cost, and emissions is also non-linear on the demand side; a case study of a few households in Finland by Kopsakangas-Savolainen *et al* applied hourly emissions factors to electricity consumption and found that some households had higher electricity-associated emissions, but lower electricity bills, than others over a fixed time period [36]. Thus, it is important that both the electricity dataset and emissions factors have high spatial and temporal detail to quantify household electricity emissions and intensity with a large degree of accuracy across a study region.

In summary, while many studies of residential sector emissions exist, there is a lack of studies on the emissions from household electricity consumption that examine a large number of homes and span spatial and climatic extents, as well as a variety of sociodemographic groups. Furthermore, most previous studies have not used time-specific methods of estimating emissions that capture the varying mix of electricity-generating resources, making it impossible to examine the relative emissions intensity of different households or of specific electricity consuming activities.

3. Methods

3.1. Datasets

In this analysis, we used smart meter measurements from households across the greater Los Angeles area to study variations and patterns in HEE. The smart meter dataset, which was provided by the investor-owned utility Southern California Edison, contains hourly electricity records from roughly 135 000 homes for the years 2018 and 2019. (Note: after filtering, the final dataset consists of $\sim$107 000 homes). The records also included the customers' matching street level addresses, which allows us to study household-level electricity demand with detailed coverage across a large region. Due to privacy concerns and to meet the security requirements of SCE, the data were stored on USC's secure research computing environment.

Hourly emissions factors for CAISO's grid were paired with the smart meter hourly electricity records to evaluate HEE. To calculate the hourly emissions factors, we retrieved reports of the 2019 historical electricity demand (with electricity exports treated as demand) at 5 min intervals from the CAISO Today's Outlook webpage and aggregated them to the hourly level [15]. The CO_2 emissions associated with the generation for each hour (i.e. the emissions from all active powerplants in CAISO during an hour) were pulled from an online data repository (S3researchUSC/MEF-Regression (github.com)). The emissions values are demand-based, meaning they account for the emissions associated with electricity trades into and out of the CAISO region [8]. Then, the grid-level demand and emissions datasets were matched at the hourly level, and hourly AEFs were calculated by dividing the total hourly emissions by the total hourly demand. Thus, hourly AEFs are equal to the amount of CO_2 produced per unit of electricity consumed in each hour.

CAISO's grid mix has a high penetration of variable renewable resources, namely solar, which causes the AEFs to vary significantly throughout the day and across seasons, depending on which resources are online. For this analysis, we assumed that all customers in our dataset received the same grid mix from CAISO in each hour, and therefore, the AEFs associated with different customers' demand during the same hour were equivalent. Because we only have emissions factors for the year 2019, we omitted the year 2018 from the emissions analysis (Note: the 2018 electricity data was still used to determine AC behavior in the AC ownership and AC state algorithms referenced in section 3.3). Thus, the emissions analysis spans one year of household electricity consumption (2019), though some customers are missing smart meter records and have less than a year of data.

We merged in additional publicly available datasets to study the relationship between HEE and factors such as climate, building characteristics, and sociodemographic features. The study region of Southern California is an ideal location for this analysis as it encompasses multiple microclimates, a wide range of building stock, and diverse communities. Hourly temperature data were retrieved from weather stations throughout the study region and used to calculate the number of CDDs experienced by each household during the study period [37–39]. Building characteristics, including the total square footage and year of constructed, were retrieved from county assessor property databases and matched to the dataset using street level addresses [40–42]. The addresses are also used to match each household to the census tract and climate zone it is located in. Finally, socioeconomic data reported at the census tract level were pulled from CalEnviroScreen 3.0 [43] and the US Census Bureau [44]. CalEnviroScreen is a mapping tool that produces a score for each of California's ~8000 census tracts based on environmental, health, and socioeconomic information to identify the most vulnerable communities. The socioeconomic variables describe the census tracts' average income and occupancy, education and unemployment levels, age and racial makeup. To visualize the results of this study, we created maps using shapefiles of California's census tracts [45] and building climate zones, as designated by the California Energy Commission [46].

Prior to estimating HEE, we implemented an outlier analysis on the smart meter data to remove any households, or hours for a specific household, that may distort the emissions analysis. First, we removed households that likely have solar panels, based on an analysis developed by Chen *et al* [47]. This step was taken because households with rooftop solar have many midday hours where zero kwh of demand was recorded, which will alter the emissions values that are calculated using hourly consumption. We filtered out households with less than 20 kWh of annual electricity demand, as it is likely that they were unoccupied. Individual households with an annual electricity demand that is greater than the mean plus three times the standard deviation (calculated across all households passing previous outlier checks) were removed as well. For each household, we filtered hours with less than 0.01 kWh of electricity demand, as the almost negligible electricity demand indicates the house was likely unoccupied in those hours, and hours with electricity demand greater than the mean of their hourly demand plus 10 times the standard deviation, as these are likely erroneous measurements. Following these steps, we removed any household with (1) less than 80% of the annual electricity records or (2) one or months with less than 50% of the electricity records. This is necessary as certain periods of the year will feature a grid with lower or higher generation from low-carbon or carbon-free resources. As such, if a household was unoccupied for long periods, it is possible that the emissions calculations would be distorted. After implementing these filtering steps, the final dataset contains 107 096 homes.

3.2. Emissions calculations

For each household, the hourly electricity records were merged with the hourly AEF values by datetime to analyze HEE. We then define the variable Annual HEE_i as the emissions associated with the electricity demand of household i for the entire year of 2019.

First, the HEE were calculated by multiplying the electricity demand in each hour by the corresponding hourly grid AEF. Then, the total emissions of each household were found by summing their hourly emissions for an entire period. To adjust for missing smart meter readings in 2019, we multiply this total emissions value by the ratio of the number of hours in a year to the number of hours for which we have smart meter readings (effectively normalizing the total emissions by the number of hours present in each home's smart meter records), as shown in equation (2),

$$\text{Annual HEE}_i = \sum_{h=0}^{H_i} E_h \times \mathrm{AEF}_h \times \frac{8760}{H_i}. \tag{1}$$

In equation (1), Annual HEE_i is the HEE of household i over the course of 2019. E_h and $\mathrm{AEF}_h$are the hourly electricity demand and AEF in hour h. H_i is the number of smart meter electricity records of customer i in all of 2019.

In this analysis, we also considered how the timing of demand impacts the emissions intensity of the households in the datasets. Due to the time-varying fleet of grid resources that provide electricity to CAISO, the emissions intensity of the power supply varies significantly throughout the day and across seasons. Thus, two customers with the same total demand across a period of time will likely have different total emissions values due to different patterns of consumption. To study how the emissions intensity of household electricity demand varies across customers, we calculated a HAEF for each customer. The HAEF_i, measured in units of grams per kilowatt-hour, is equal to the total emissions of the household for a time period divided by the total demand of the household in that same time period, as described in equation (2). We calculate the HAEF_i for every household for the entire 2019 period, as well as separately for each season in 2019,

$$\mathrm{HAEF}_i = \frac{\sum_{h=0}^{H_i} E_h \times \mathrm{AEF}_h}{\sum_{h=0}^{H_i} E_h} \text{ for } H_i \in P. \tag{2}$$

In equation (2), the HAEF_i is the HAEF of household i defined for a given period P (annual or seasonal). The remaining variables were defined in equation (1),

The mean annual HEE_i and HAEF_i values for each census tract in the dataset were calculated to understand how emissions vary across the study region. Using the calculated emissions values, household and census tract level analyses were carried out to identify factors that contributed to higher emissions intensity and HAEFs. Regressions between the annual HEE_i and annual HAEF_i values and household level variables such as CDD, total square footage, and year of building construction were conducted. Similarly, regressions of the annual HEE_i and HAEF_i on census tract sociodemographic variables were performed. We use AEFs for our emissions calculations, rather than marginal emissions factors (MEFs), because we focus our analysis primarily on historical emissions and the comparison of actual load profiles, but acknowledge that MEFs may be more appropriate when discussing changes in consumption patterns.

3.3. AC emissions

We also investigated the emissions associated with electricity consumed for AC use and compared these results with overall HEE. First, we identified which households have AC using the AC ownership Algorithm developed in Peplinski *et al* [16]. In this methodology, households are fit to four models that represent different dependencies between hourly electricity use and demand. The model with the best fit determines if the household has AC or not. For each household that we determined has AC, we implemented the AC state Algorithm, a multiple linear regression model adapted from Dyson *et al* [17], to identify the hours in which customers turned their AC on (Refer to Peplinski *et al* for a detailed explanation of the methodology.) The model regresses electricity demand on temperature and dummy variables that represent hour of day and day of week interactions and can be used to predict electricity consumption if the AC is on or off. For each hour that was classified as on by the AC state Algorithm, we used the difference between the AC-on and AC-off electricity consumption predictions to estimate AC related electricity consumption. We used both 2018 and 2019 data as inputs to the AC ownership and AC state Algorithms to provide as much data as possible for the regression, however, we only utilized the 2019 results in our emissions analysis (recall that CAISO emissions data was not available for 2018).

Using the estimated AC demand, we then analyzed the emissions associated with AC use. Equations (1) and (2) were used to calculate the HEE from AC use specifically and the average emissions intensity of electricity consumed for AC for the year 2019. To do this, we substituted hourly AC electricity demand for E_h in each of the equations. By examining the emissions of AC use, we can gain insight into the emissions intensity associated with cooling demand versus the emissions intensity for total electricity demand.

4. Results and discussion

4.1. Distribution of HEE

We first analyzed the distribution of annual HEE for 2019 across the households and census tracts within our dataset. At the household level, we calculated an average annual HEE of 1310 kg of CO_2, with a 5th to 95th percentile spread of 373–2770 kg of CO_2. We then averaged the annual HEE to the census tract level, finding a narrower distribution with a 5th–95th percentile spread of 739–2001 kg of CO_2 in 2019. The distribution of household and census tract average annual HEE are shown in figure 1.

The distribution in figure 1 shows that a small number of households have particularly high annual HEE values. To gain a better understanding of the inequity of residential emissions, we use the annual HEE values to plot a Lorentz curve, a graphical representation of inequality, in figure 2.

The Lorenz curve depicts the percent of households in our dataset, on the x-axis, that is responsible for a percent of emissions, on the y-axis. Following the black curve with households ordered from lowest to highest annual HEE, we find that in 2019 the bottom 60% of households were responsible for just 37% of annual HEE, while the top 20% of households were responsible for ~40% of annual HEE. The inequality in HEE can also be described by the Gini index, which is measured using the difference between the line of

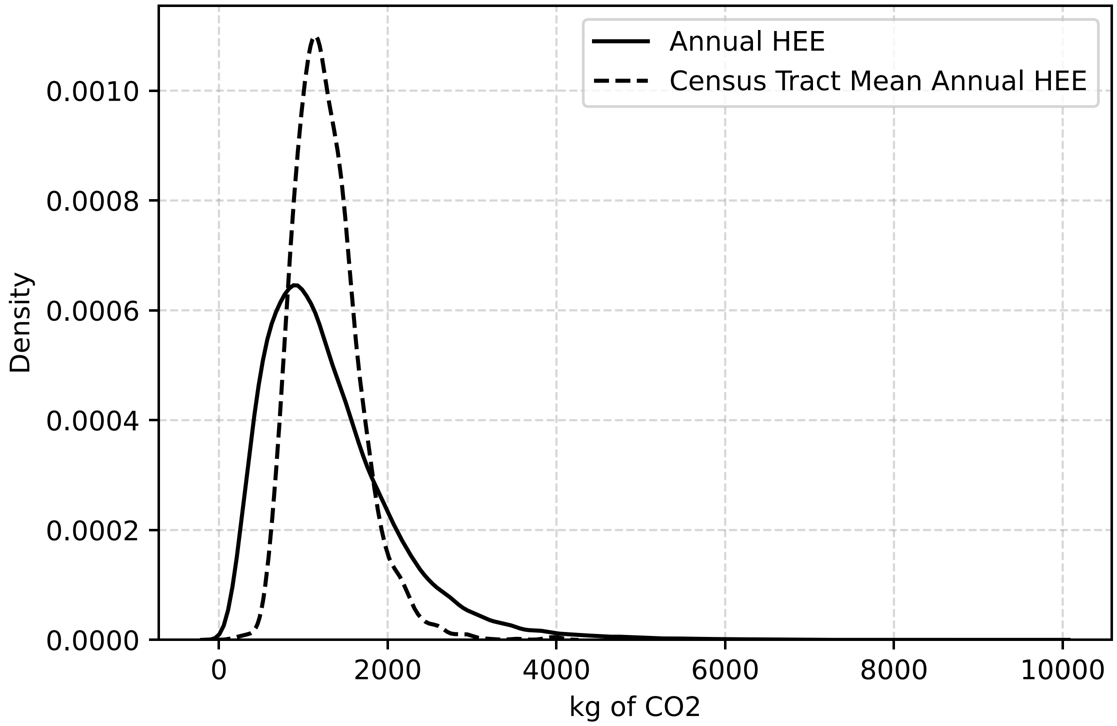

Figure 1. Distribution of annual HEE and census tract average annual HEE for the year 2019. We observe a large range in household emissions across the region. There is a much higher degree of variation in household level annual HEE than census tract level annual HEE.

perfect equality and the Lorenz curve. The Gini index of the annual HEE in our dataset is 0.33, which indicates an intermediate level of inequality (a Gini index of 0 represents perfect equality and 1 represents complete inequality) [48].

In figure 3 we show the mean annual HEE for each census tract in our dataset. Although averaging the annual HEE to the census tract level reduces the variation, there is still a large difference between the high- and low-emitting census tracts.

The desert and inland areas generally have higher annual HEE than the coastal and mountainous areas, possibly due to higher electricity demand due for cooling. However, the variation in annual HEE within the microclimates suggests that factors beyond location, such as sociodemographic traits or building characteristics, have a large impact on emissions. We explore the impact of these factors further in section 4.4.

4.2. HAEFs

HEE are highly dependent on the magnitude of a household's demand, but the timing of its electricity consumption is an important factor as well. This is especially true for households in our dataset, as California has high penetrations of renewable energy and thus, significant variations in the hourly AEF value. To illustrate that residential demand can occur at times with varying levels of emissions intensity, we first group the hourly demand data of all households by grid AEF percentile and then find the sum of demand in each grid AEF percentile across all households in our dataset. We show these results in figure 4.

Over the course of 2019, the lowest AEF hours feature high penetrations of VRE and minimal fossil-fuel based generation, while the highest AEF hours were dominated by natural gas generation and emissions-intense imports. CAISO's electricity imports typically come from regions with a higher fraction of fossil-fuel generation, including balancing authorities with coal generation in states like Arizona [6]. Residential demand in these high AEF hours therefore causes significantly more CO_2 emissions than in the low AEF hours. While there is not a strong correlation between demand and grid AEF, the households in our dataset do consume more electricity in high AEF hours than low AEF hours. Furthermore, each specific household has a unique distribution of demand that may be concentrated to high AEF or low AEF times. Therefore, in this study, we also examined how differences in the timing of customers' demand influenced their emissions.

We evaluated this phenomenon by calculating each customer's HAEF_i, which describes the average emissions intensity of a household's electricity consumption over a period of time. The annual HAEF_i values span a large range from lowest to highest emission-intensity household, though they show less spread across

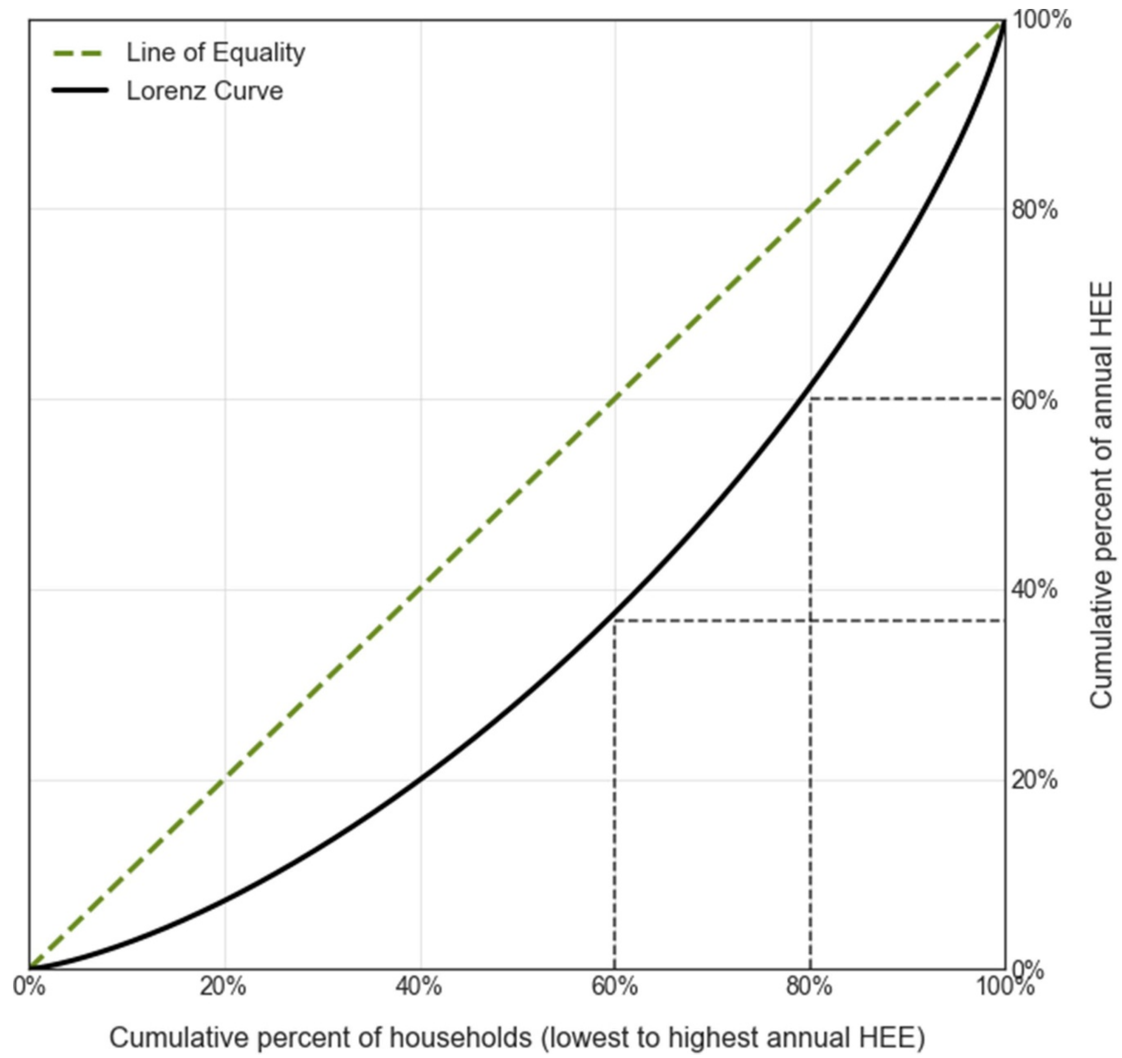

Figure 2. The Lorenz curve of annual HEE across all households in the dataset in 2019, a graphical representation of the distribution of emissions, is shown in black. The *x*-axis shows the percentage of customers (ordered from least to most emitting) and the *y*-axis shows the corresponding cumulative percentage of annual HEE that these customers are responsible for. The line of perfect equality is shown in green (e.g. *x*% of customers consume exactly *x*% of emissions), and the area between the curves represents the degree of inequality.

users than annual HEE. We find a mean annual HAEF of 189 g of CO_2 per kWh with a 5th–95th percentile spread of 173–201 g of CO_2 per kWh. We also calculate HAEF by season and include these distributions in figure 5. The seasons are divided as follows: Winter: December, January, February; Spring: March, April, May; Summer: June, July, August; Fall: September, October, November.

Although annual HAEFs are less variant than annual HEE, some users still have patterns of electricity use that are far more emissions intense. For example, the customer with the most emissions intense electricity consumption has an annual HAEF $\sim$1.7 times that of the customer with the least emissions intense electricity consumption (230 versus 130 g of CO_2 per kWh). While homes with high emissions due to high electricity consumption are likely good targets for programs focused on energy efficiency and demand reduction, the homes with high HAEFs may be targets for load-shifting or other DSM programs designed to create more sustainable patterns of consumption. However, it should also be noted that the impact of load-shifting on emissions is at times evaluated using MEFs, which may follow a different seasonal or diurnal pattern than AEFs. Therefore, a household could be also considered a good candidate for reducing emissions via DSM due to having a large portion of consumption during high-MEF hours (identifying these household is beyond the scope of this study). Because CAISO's grid has less renewables in the fall and winter months, we find generally higher fall and winter HAEFs. More importantly, we observe higher variance in HAEF in the spring and summer, meaning there is a large range across households in how emissions intense their electricity consumption is. This suggests that load flexibility could be more impactful in in the spring and summer in terms of reducing emissions. For example, the highest spring HAEF is roughly 2.2 times higher than the lowest spring HAEF, meaning that the highest HAEF household could reduce its spring-time

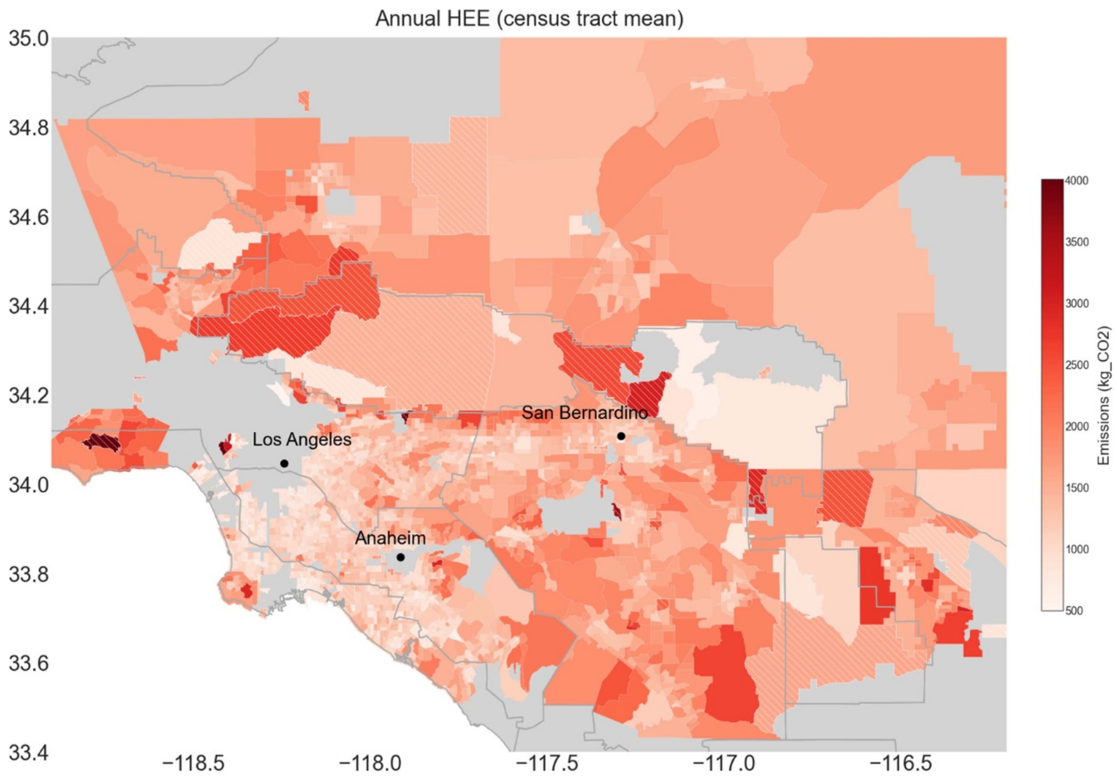

Figure 3. Choropleth map of the average annual HEE for each census tract in 2019, calculated by taking the mean of the annual HEE for each household in our dataset located within the census tract. We generally observe higher emissions in inland census tracts, although there are notable exceptions. Shading indicates census tracts for which fewer than 10 homes are present in our dataset.

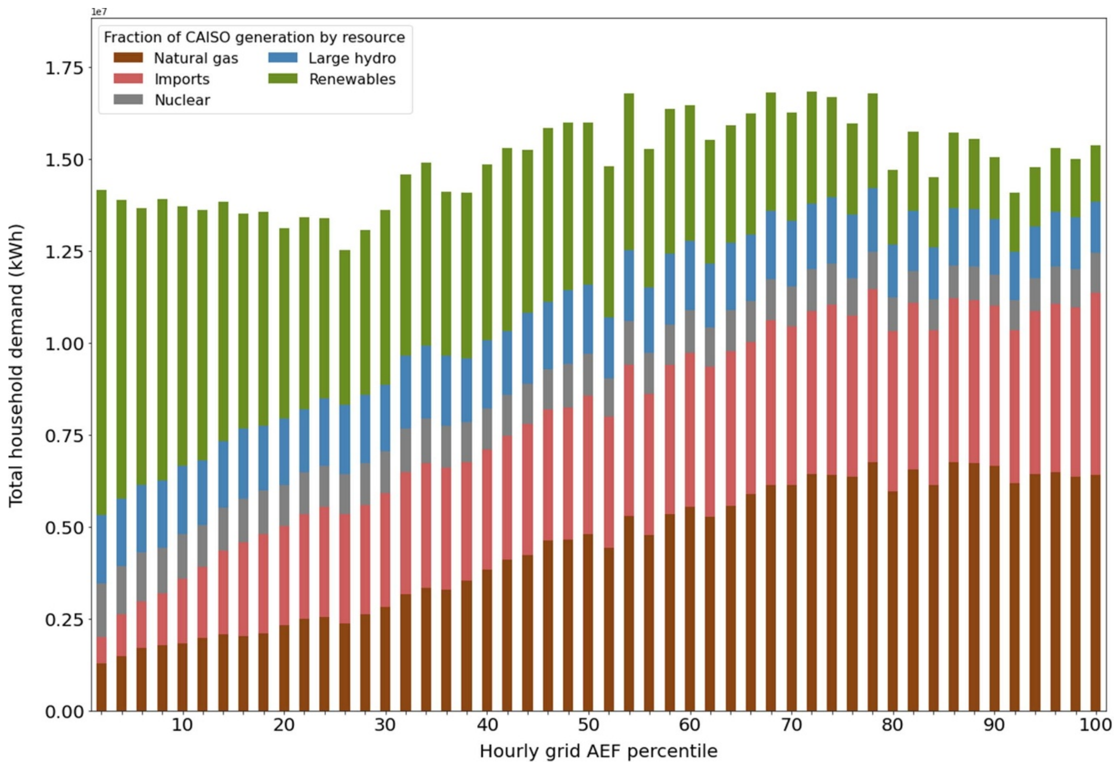

Figure 4. Sum of 2019 demand across all 107 096 households grouped by grid AEF percentile. En masse, households consume slightly more electricity during higher AEF hours. The grid mix varies drastically from the lowest AEF times (43.5 gCO_2 per kWh), when most of the generation is carbon-free, to the highest AEF times (295 gCO_2 per kWh), when the majority of generation is natural gas and imports.

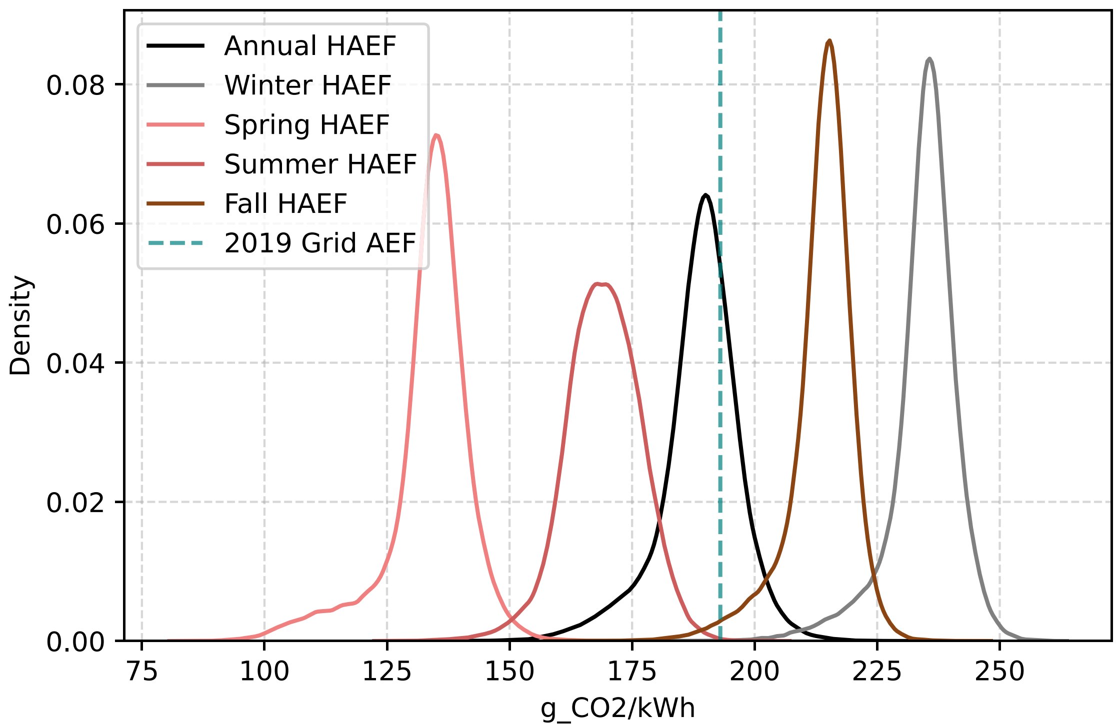

Figure 5. Distribution of annual HAEF and seasonal HAEFs shown with the annual grid AEF in 2019. There is a large spread in the emissions intensity of electricity consumption, with more of the variety in annual HAEF coming from consumption in spring and summer months.

electricity-associated emissions by 54% by consuming electricity with equal emissions intensity to the customer with the lowest HAEF, without reducing their demand by a single kWh.

The results of this analysis indicate that emissions reductions in the residential sector could be realized by encouraging different load patterns. In figure 6, we show the average normalized daily electricity load profile in each season for the top and bottom 10% of households by seasonal HAEF.

Figure 6 shows that high HAEF households typically demand more of their electricity in the evening, while low HAEF household consume more of their electricity in the middle of the day across all four seasons. The high HAEF households may be good candidates for a load-shifting program, with a focus on identifying evening loads that can be shifted earlier in the day. In summer, the highest HAEF households have a peak only a few hours earlier than the lowest HAEF households, implying that small temporal shift can be extremely impactful on emissions. Interestingly, the distributions in figure 5 shows that the mean HAEF is lower than CAISO's annual grid AEF, suggesting that SCE's residential sector electricity behavior is less emissions intense than other loads on CAISO's grid. Additional analysis is needed to determine if the residential sector as a whole consumes electricity with lower emissions-intensity than other sectors.

We also examine these results across the study region at the census tract level, again taking the mean of all households residing within a census tract, as shown in figure 7.

While individual users display a lot of variance in HAEF, the fluctuation is muted at the census tract level, suggesting that the electricity demand patterns that lead to emissions-intense consumption are not driven by locational or climatic factors. Therefore, programs that target emissions reductions through reducing HAEFs would likely need to identify specific customers, rather than areas, as targets. Further investigation between HAEF and climatic variables, building characteristics, and sociodemographic factors can be found in section 4.4.

4.3. AC emissions overview

AC demand represents a significant portion of household electricity consumption and, consequentially, a significant portion of HEE can be tied to AC use. Averaging across all households, about 12% of annual HEE come from AC, with that number rising to roughly 16% when averaging across only homes classified as having AC in this dataset. The fraction of annual HEE resulting from AC use ranges from 0% to 72% across all households in our dataset.

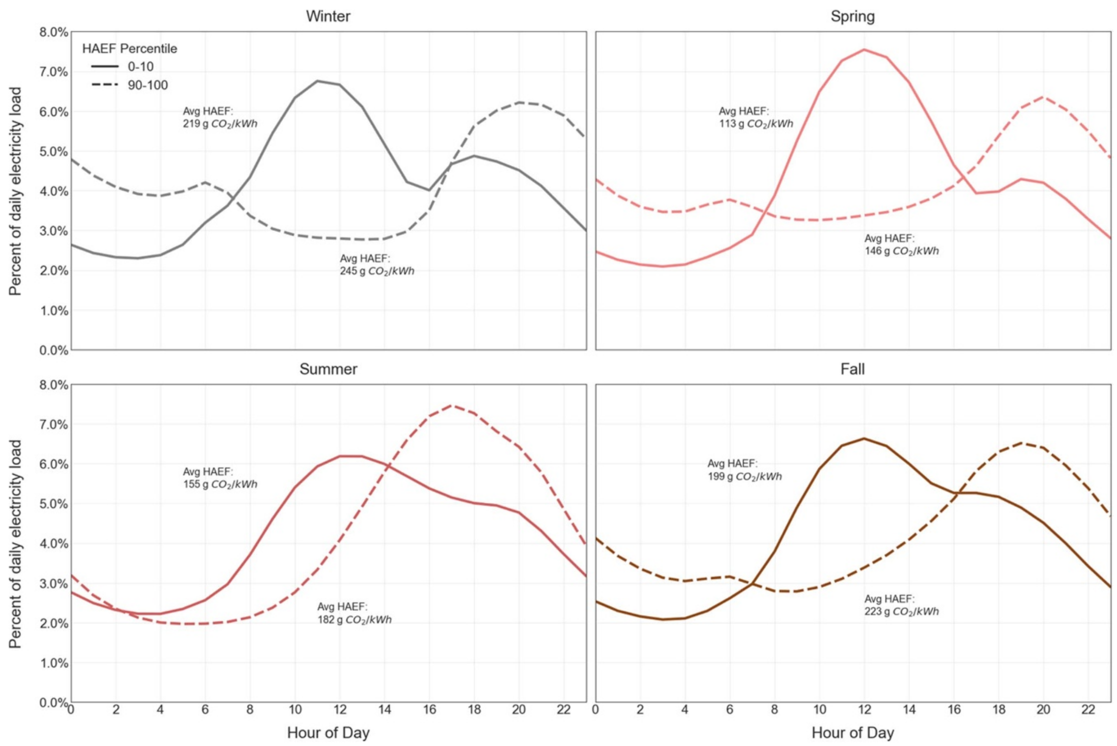

Figure 6. Normalized daily electricity load in each season of 2019 average across the top and bottom 10% of households by seasonal HAEF. Lower HAEF households have a higher fraction of consumption in the middle of the day, while high HAEF households have more of their consumption concentrated in evening hours.

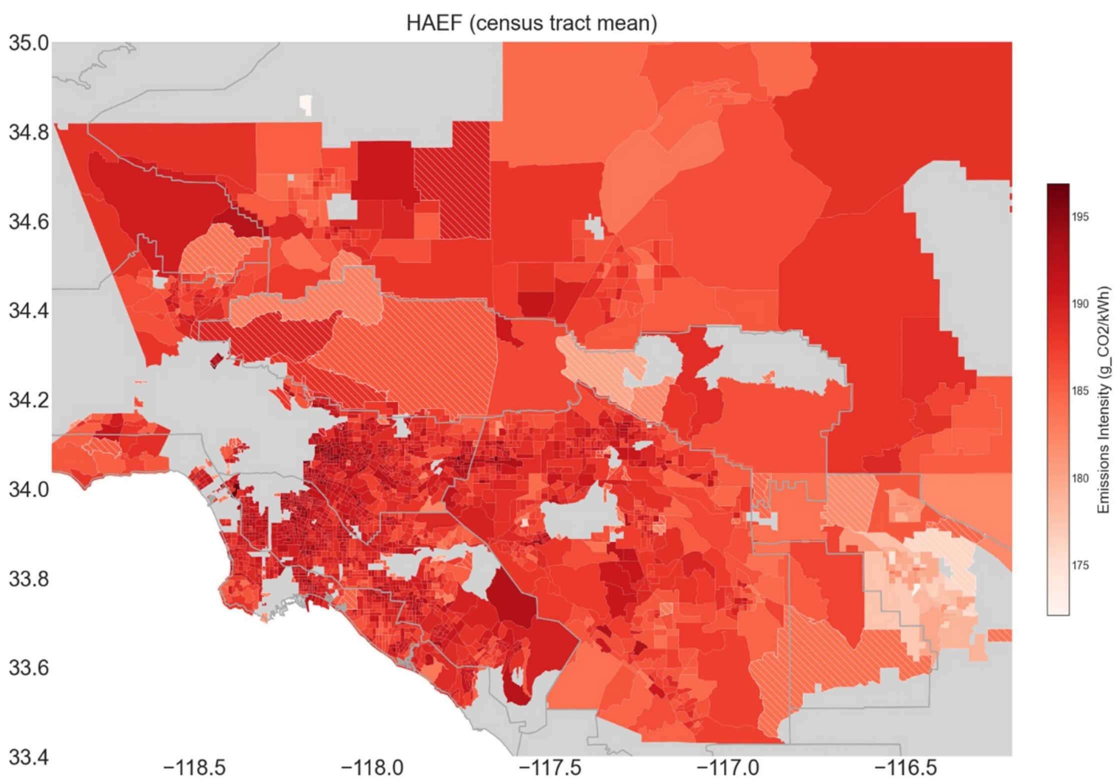

Figure 7. Average HAEF for all census tracts in the study area in 2019. Most census tracts fall near the average for the entire study region, though inland census tracts appear to have slightly lower HAEFs. Shading indicates census tracts for which fewer than 10 homes are present in our dataset.

We also examine the HAEF for AC use, which describes the average emissions intensity of electricity consumed for AC use specifically. We show the distribution of HAEF for AC alongside the previously displayed HAEFs in figure 8.

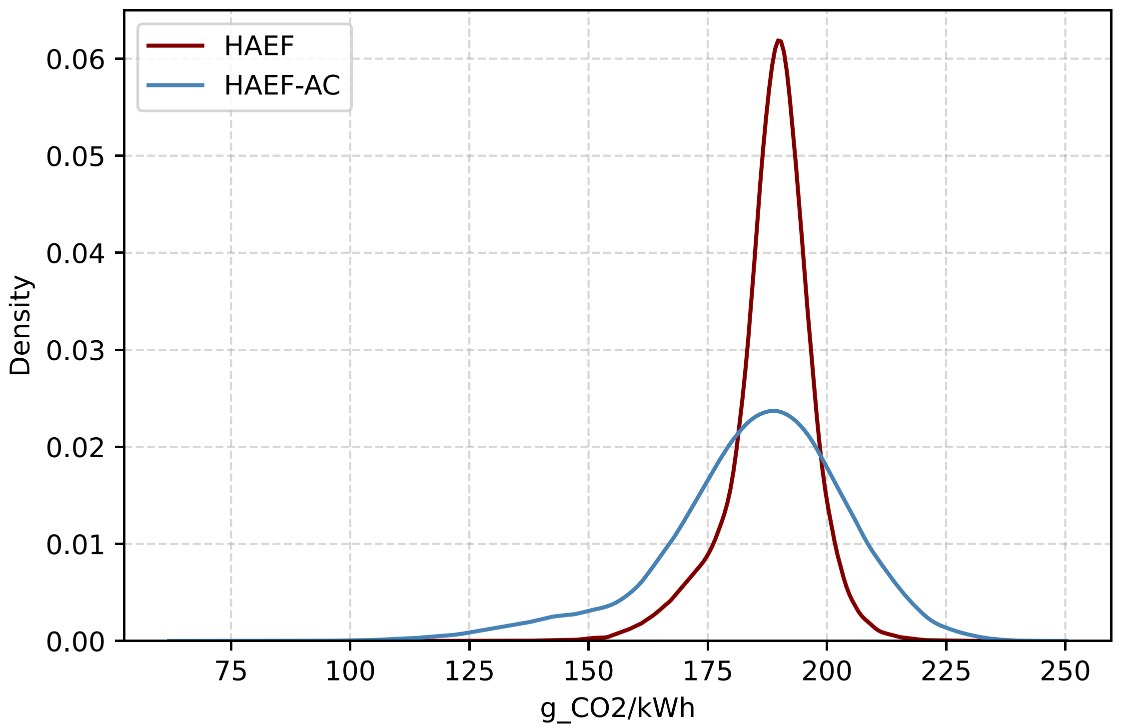

Figure 8. Distribution of the emissions intensity for all household electricity consumption and AC-specific electricity consumption (for all homes classified as having AC) in 2019. Electricity for AC has a much larger range of emissions intensity across households, and the high HAEF for AC households are potential targets for AC load shifting.

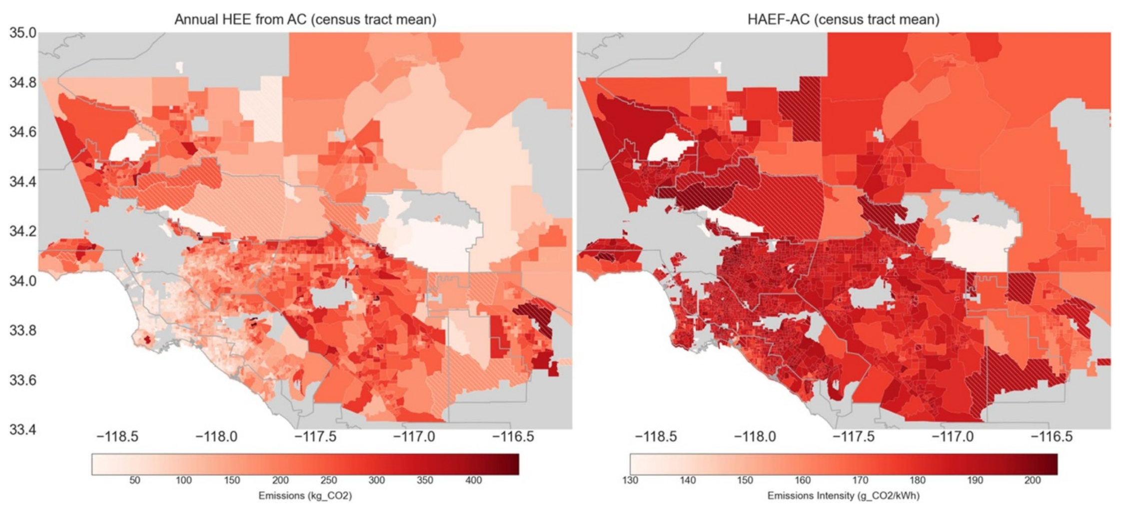

Figure 9. Map of census tract mean annual HEE and annual HAEF for electricity consumed by air conditioning in 2019. The census tract mean Annual HEE from AC is calculated using all households in the dataset. The census tract mean annual HAEF for AC is calculated only for households identified as having AC. Shading indicates census tracts for which fewer than 10 homes are present in our dataset.

Figure 8 shows that AC consumption is on average slightly less emissions intense than non-AC consumption, with a mean HAEF for AC of 185 gCO2/kWh^{-1}, versus 189 gCO2/kWh^{-1} for general consumption (a small but statistically significant result at the $p = .001$ level both pairwise and en masse). However, we observe a much larger spread in HAEF for AC. This means that there is a substantial fraction of households who consume electricity for AC in a particularly emissions-intense way, making them strong candidates for DSM that targets the timing of AC use. We also show the mean annual HEE from AC across census tracts in figure 9.

In general, there are lower AC-related emissions in the coastal regions, where cooler temperatures reduce the need for AC ownership and operation. These results are in line with previous studies of cooling demand in Southern California which found lower rates of AC ownership in the cooler, coastal regions and higher

Table 1. Summary of regression of 2019 census tract mean annual HEE on sociodemographic variables and building climate zone. This model achieves an R^2 value of 0.72 with all variables significant at the 0.05 level.

	Dependent variable-Census tract mean annual HEE				
Variable	Coef	Std Err	p-value	0.025	0.975
Constant	724.9	34.36	0.000	657.5	792.3
Mean income	1793	67.52	0.000	1661	1926
Education	−307.2	42.44	0.000	−390.4	−224.0
Percent over 65	461.4	59.84	0.000	344.1	578.8
Average occupancy	1010	49.64	0.000	912.5	1107
Unemployment	96.73	41.86	0.021	14.65	178.8
Percent Black	−94.17	43.04	0.029	−178.6	−9.765
Percent Asian	−342.9	33.92	0.000	−409.4	−276.4
Percent Latino	−321.0	46.85	0.000	−412.9	−229.1
CZ 6	−293.9	15.68	0.000	−324.7	−263.2
CZ 8	−156.9	12.56	0.000	−181.5	−132.3
CZ 10	201.5	14.38	0.000	173.3	229.7
CZ 14	296.5	22.96	0.000	251.5	341.6
CZ 15	584.1	30.29	0.000	524.7	643.5
CZ 16	88.61	44.89	0.049	0.579	176.6

rates in the inland and desert regions [3, 47]. Interestingly, the higher emissions from AC electricity consumption in the desert areas appear to be mitigated by a less emissions-intense pattern of use. This may occur because these regions feature high daytime heat but relatively low nighttime heat, so the majority of the cooling need is in the middle of the day when solar penetrations are high.

4.4. Influence of household and census-tract variables on HEE

In section 4.1, we observed that there is a large spread in the HEE values with some weak spatial trends observed in figure 3. Our processed dataset includes information about the square footage, building age, and the number of CDD experienced by each household (calculated individually for each residence using the nearest weather station). To better understand the influence of these variables on HEE, we regress annual HEE on these factors, and a constant. Prior to regression, we normalize our independent variables to better understand the relative influence of each feature. We find a strong positive relationship between annual HEE and square footage, but surprisingly find a negative relationship between building age and annual HEE, and no statistically significant relationship with cooling degree days. However, the low R^2 value of 0.17 indicates that the majority of variation in household-level emissions is not predicted by these variables and suggests that these conclusions may not be reliable. The full results of this regression are shown in the SI.

Given the weak relationship at the building level, we instead use the aggregated results to analyze HEE at the census tract level and examine the influence of socio-demographic variables on emissions. Here we regress the census tract average annual HEE on mean income, education level, percent of population over 65, average occupancy, unemployment rate, percent of population identifying as Black, Asian, and Latino, building climate zone, and a constant. To ensure that census tracts are well represented, we require a minimum of 10 homes per census tract for inclusion in the regression. We treat building climate zone as a categorical variable, and again normalize the remaining variables prior to performing the regression. To avoid correlation between variables, we drop building climate zone 9 prior to regression. The coefficient on the climate zones can then be interpreted as having results relative to climate zone 9. The results of the census tract level regression are displayed in table 1.

The regression between sociodemographic variables and mean census tract annual HEE explains a significant portion of the variance, with an R^2 value of 0.72. Mean census tract income and average occupancy (persons per household) have a large, statistically significant, positive effect on annual HEE. Having older occupants, who are perhaps more likely to be home throughout the day, also has a positive effect on emissions. Higher education tends to somewhat lower emissions, while unemployment has a small positive impact. Census tracts with high Asian, Black, or Latino identifying populations generally have slightly lower annual HEE. Regarding climate zones, the inland, warmer climate zones of 10, 14, and 15 have higher HEE, suggesting that in a well-specified model there is a relationship between climate and emissions.

Lastly, we perform the same regression but use mean census-tract annual HAEF as the dependent variable. While there was limited variation in HAEF across census tracts, we were still able to identify statistically significant influences for many of the independent variables, as shown in table 2. We again use only census tracts with more than 10 households.

Table 2. Summary of regression of 2019 census tract average HAEF on sociodemographic variables and building climate zone. This model achieves an R^2 value of 0.66 with 13 out of 15 variables significant at the 0.05 level.

	Dependent variable: census tract mean annual HAEF				
Variable	Coef	Std Err	*P*-value	0.025	0.975
Constant	190.1	0.297	0.000	189.5	190.7
Mean income	−2.179	0.584	0.000	−3.324	−1.034
Education	3.146	0.367	0.000	2.426	3.865
Percent over 65	−2.856	0.517	0.000	−3.871	−1.842
Average occupancy	−4.449	0.429	0.000	−5.290	−3.607
Unemployment	−0.379	0.362	0.295	−1.089	0.331
Percent Black	2.235	0.372	0.000	1.505	2.964
Percent Asian	3.064	0.293	0.000	2.489	3.639
Percent Latino	3.067	0.405	0.000	2.273	3.861
CZ 6	−0.054	0.136	0.691	−0.320	0.212
CZ 8	−0.400	0.109	0.000	−0.613	−0.187
CZ 10	−1.753	0.124	0.000	−1.997	−1.509
CZ 14	−3.532	0.198	0.000	−3.921	−3.143
CZ 15	−10.99	0.262	0.000	−11.50	−10.48
CZ 16	−1.995	0.388	0.000	−2.756	−1.234

We observe an R^2 value of 0.66 for the model of census tract average annual HAEF and sociodemographic variables. In general, census tracts with higher income, higher occupancy, and higher fraction of individuals over 65 have slightly lower HAEF (i.e. slightly less emissions-intense patterns of consumption). Census tracts with higher education, as well as a higher percent of population identifying as Asian, Black, and Latino, tend to have slightly more emissions-intense electricity consumption. Census tracts in building climate zone 15 tend to have less emissions-intense consumption, which could occur due to the magnitude and timing of their AC consumption. The high daytime and low nighttime temperatures in climate zone 15 may lead to a high fraction of electricity consumption coming from daytime AC use, during which solar production is often high and the grid is relatively clean.

Generally, we see that the variations in census tract mean annual HEE and mean annual HAEF can be fairly well predicted by geographical location and sociodemographic factors. These results could be used to identify candidates for emissions reducing strategies. For example, census tracts with high income households have higher annual HEE (more emissions), but lower annual HAEFs (less emissions intense consumption), suggesting that the issue is simply the magnitude of consumption. On the other hand, census tracts in building climate zone 6 generally have lower annual HEE, but higher HAEFs than every building climate zone other than 9. For this area, a program focused on changing the timing of consumption may be more successful for reducing emissions.

5. Conclusion

In our analysis of HEE, we observe significant inequity at the household level, with the top 20% of households responsible for more HEE than the bottom 60% over the course of 2019. While homes with larger square footage and more occupants tend to emit more, we also find that households in census tracts with higher mean incomes, a larger percentage of 65+ residents, or a larger percentage of white residents are more likely to be high emitting. Though these higher-emitting households typically consume large amounts of electricity, our analysis also revealed that some households demand electricity in a significantly more emissions-intense way over the course of a year than others do, meaning they are responsible for more emissions per kWh consumed (by up to a factor of 1.7x). This is especially true in the spring and summer months when there is more variation in consumption patterns and in the emissions intensity of the grid. Therefore, DSM strategies for reducing household emissions should focus not only on reducing the magnitude of demand, but should also target households that have particularly emissions-intense patterns of electricity consumption. We also found that AC is responsible for as much as 72% of annual HEE, and that the emissions-intensity of AC use varies more across households than non-AC electricity consumption, implying that specific households may be able to make notable reductions in emissions by treating AC as a flexible load and relying on their households' thermal mass. As electricity grids achieve higher penetrations of carbon-free generation and experience more fluctuations in net load, electricity demand and emissions will continue to decouple. Targeted DSM can be an important tool in the battle to mitigate climate change and reduce inequity in emissions. We also observe that the average customer in our dataset generally consumes electricity in a less emissions-intense way than the average customer on CAISO. In future studies, the emissions intensity of other sectors, and of residential customers for other California utilities, could be tested to identify which sectors and utilities should prioritize load-shifting with the goal of reducing emissions.

There were several limitations and future improvements to this work that should be addressed. First, while our dataset spans many census tracts, we estimate that we only have statistically significant results for a subset of these census tracts, as was shown in figures 3, 6, and 8. When determining the influence of sociodemographic variables on HEE and HAEF, the lack of household-level data required us to perform regressions at the census tract level. As a result, we can only draw conclusions at the census tract level (e.g. higher-income census tracts tend to contain households with higher emissions than lower-income census tracts). When analyzing the HAEF for AC, we were limited by the lack of ground truth regarding the presence of an AC in a home, and instead relied on an established methodology for estimating the presence of an AC. This methodology was checked against survey data for the region with encouraging results but was not able to be verified at the household level. Lastly, this analysis compares the emissions associated with AC use and non-AC use but does not examine the distribution of other specific electricity consuming technologies, such as electric heating. We believe an analysis comparing the emissions intensity of electricity consumption for different household technologies assuming typical use patterns would be highly useful for programs designed to reduce emissions through DSM.

The implications of this work are especially important in the face of rising electricity demand in the residential sector, a consequence of many climate change mitigation and adaptation strategies. As we aim to minimize the emissions associated with this rising electricity consumption, our results help define which groups of users are high-emitting and the cause of their electricity-associated emissions to identify appropriate strategies for reducing emissions.

Data availability statement

The data cannot be made publicly available upon publication because they contain sensitive personal information. The data that support the findings of this study are available upon reasonable request from the authors.

Funding

- NSF CBET-1845931
- NSF CBET-1752522
- JPL (Sponsor: NASA) Neighborhood-Scale Extreme Humid Heat Health Impacts Grant

Conflict of interest

We declare no conflicts of interest. No parts of this manuscript have been published in any other format.

Please note that the views herein do not necessarily reflect the views of any current or past employer.

Sincerely,

Dr Stepp Mayes, Dr McKenna Peplinski, and Dr Kelly Sanders

Department of Civil and Environmental Engineering, University of Southern California

ORCID iDs

Stepp Mayes https://orcid.org/0000-0002-5762-8291
McKenna Peplinski https://orcid.org/0000-0002-6353-6211
Kelly T Sanders https://orcid.org/0000-0003-4466-0054

References

[1] Energy Information Administration (EIA) 2021 How much of US carbon dioxide emissions are associated with electricity generation *2022%2C the U.S. electric,sector accounted for about 31%25* (available at: www.eia.gov/tools/faqs/faq.php?id=77%26t=3#:[?]:text=In)

[2] E. I. A. (EIA) U.S. energy facts explained 2023 (available at: www.eia.gov/energyexplained/us-energy-facts)

[3] Peplinski M, Dilkina B, Chen M, Silva S J, Ban-Weiss G A and Sanders K T 2024 A machine learning framework to estimate residential electricity demand based on smart meter electricity, climate, building characteristics, and socioeconomic datasets *Appl. Energy* **357** 122413

[4] Huebner G, Shipworth D, Hamilton I, Chalabi Z and Oreszczyn T 2016 Understanding electricity consumption: a comparative contribution of building factors, socio-demographics, appliances, behaviours and attitudes *Appl. Energy* **177** 692–702

[5] Jones R V, Fuertes A and Lomas K J 2015 The socio-economic, dwelling and appliance related factors affecting electricity consumption in domestic buildings *Renew. Sustain. Energy Rev.* **43** 901–17

[6] Zohrabian A, Mayes S and Sanders K T 2023 A data-driven framework for quantifying consumption-based monthly and hourly marginal emissions factors *J. Clean. Prod.* **396** 136296
[7] Ryan N A, Johnson J X and Keoleian G A 2016 Comparative assessment of models and methods to calculate grid electricity emissions *Environ. Sci. Technol.* **50** 8937–53
[8] Mayes S, Klein N and Sanders K T 2024 Using neural networks to forecast marginal emissions factors: a CAISO case study *J. Clean. Prod.* **434** 139895
[9] Romero-Jordán D and Del Río P 2022 Analysing the drivers of the efficiency of households in electricity consumption *Energy Policy* **164** 112828
[10] Lyubich E *et al* 2020 The race gap in residential energy expenditures *Energy Inst. Work. Pap.*
[11] Huang W H 2015 The determinants of household electricity consumption in Taiwan: evidence from quantile regression *Energy* **87** 112828
[12] Lyons S, Pentecost A and Tol R S J 2012 Socioeconomic distribution of emissions and resource use in Ireland *J. Environ. Manage.* **112** 186–98
[13] Cetin K S, Siemann M and Sloop C 2016 Disaggregation and future prediction of monthly residential building energy use data using localized weather data network *ACEEE Rep*
[14] Luo Z *et al* 2022 Demand flexibility of residential buildings: definitions, flexible loads, and quantification methods *Engineering* **16** 176–86
[15] Today's Outlook *CAISO* (available at: www.caiso.com/todays-outlook/demand)
[16] Peplinski M, Mayes S and Sanders K T 2024 Revealing spatial and temporal patterns of residential cooling in Southern California through combined estimates of AC ownership and use
[17] Dyson M E H, Borgeson S D, Tabone M D and Callaway D S 2014 Using smart meter data to estimate demand response potential, with application to solar energy integration *Energy Policy* **73** 607–19
[18] Geng Y, Chen W, Liu Z, Chiu A S F, Han W, Liu Z, Zhong S, Qian Y, You W and Cui X 2017 A bibliometric review: energy consumption and greenhouse gas emissions in the residential sector *J. Clean. Prod.* **159** 301–16
[19] Chen H, Du Q, Huo T, Liu P, Cai W and Liu B 2023 Spatiotemporal patterns and driving mechanism of carbon emissions in China's urban residential building sector *Energy* **263** 126102
[20] Zhu D *et al* 2013 Temporal and spatial trends of residential energy consumption and air pollutant emissions in China *Appl. Energy* **106** 17–24
[21] Baynes T, Lenzen M, Steinberger J K and Bai X 2011 Comparison of household consumption and regional production approaches to assess urban energy use and implications for policy *Energy Policy* **39** 7298–309
[22] Mattinen M K, Heljo J, Vihola J, Kurvinen A, Lehtoranta S and Nissinen A 2014 Modeling and visualization of residential sector energy consumption and greenhouse gas emissions *J. Clean. Prod.* **81** 70–80
[23] Danylo O, Bun R, See L and Charkovska N 2019 High-resolution spatial distribution of greenhouse gas emissions in the residential sector *Mitig. Adapt. Strateg. Glob. Change* **24** 941–67
[24] Liu L *et al* 2020 Household CO_2 emissions: current status and future perspectives *Int. J. Environ. Res. Public Health* **17** 7077
[25] Goldstein B, Gounaridis D and Newell J P 2020 The carbon footprint of household energy use in the United States *Proc. Natl Acad. Sci. USA* **117** 19122–30
[26] Zen I S, Al-Amin A Q, Alam M M and Doberstein B 2021 Magnitudes of households' carbon footprint in Iskandar Malaysia: policy implications for sustainable development *J. Clean. Prod.* **315** 128042
[27] Wilson J, Spinney J, Millward H, Scott D, Hayden A and Tyedmers P 2013 Blame the exurbs, not the suburbs: exploring the distribution of greenhouse gas emissions within a city region *Energy Policy* **62** 1329–35
[28] VandeWeghe J R and Kennedy C 2007 A spatial analysis of residential greenhouse gas emissions in the toronto census metropolitan area *J. Ind. Ecol.* **11** 133–44
[29] Nässén J, Andersson D, Larsson J and Holmberg J 2015 Explaining the variation in greenhouse gas emissions between households: socioeconomic, motivational, and physical factors *J. Ind. Ecol.* **19** 480–9
[30] Papachristos G 2015 Household electricity consumption and CO_2 emissions in the Netherlands: a model-based analysis *Energy Build.* **86** 403–14
[31] Zaidan E, Abulibdeh A, Jabbar R, Onat N C and Kucukvar M 2024 Evaluating the impact of the COVID-19 pandemic on the geospatial distribution of buildings' carbon footprints associated with electricity consumption *Energy Strateg. Rev.* **52** 101350
[32] Pachauri S 2014 Household electricity access a trivial contributor to CO 2 emissions growth in India *Nat. Clim. Change* **4** 1073–6
[33] Karimu A, Krishnamurthy C K B and Vesterberg M 2022 Understanding hourly electricity demand: implications for load, welfare and emissions *Energy J.* **43** 161–90
[34] Nilsson A, Stoll P and Brandt N 2017 Assessing the impact of real-time price visualization on residential electricity consumption, costs, and carbon emissions *Resour. Conserv. Recycl.* **124** 152–61
[35] Miller G J, Novan K and Jenn A 2022 Hourly accounting of carbon emissions from electricity consumption *Environ. Res. Lett.* **17** 044073
[36] Kopsakangas-Savolainen M, Mattinen M K, Manninen K and Nissinen A 2017 Hourly-based greenhouse gas emissions of electricity—cases demonstrating possibilities for households and companies to decrease their emissions *J. Clean. Prod.* **153** 384–96
[37] California Irrigation Management Information Sytem (CIMIS) n.d. *CIMIS Station Reports [Online]* (available at: https://cimis.water.ca.gov/Stations.aspx)
[38] National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI) n.d. *NOAA NCEI Local Climatological Data (LCD) [Online]* (available at: https://www.ncei.noaa.gov/cdo-web/datatools/lcd)
[39] Environmental Protection Agency (EPA) n.d. *EPA Air Quality System (AQS) [Online]* (available at: https://aqs.epa.gov/aqsweb/airdata/download_files.html)
[40] Assessor Parcel Data 2016 *County of Los Angeles Open Data* (available at: https://data.lacounty.gov/browse?category=Property%2FPlanning&utf8=✓)
[41] San Bernardino County Assessor's Property Characteristics 2016 *Office of San Bernardino County Assessor-Recorder-Clerk* (available at: https://sbcountyarc.org/services/property-information/)
[42] Riverside County Assessor's Property Characteristics 2016 *County of Riverside Assessor-County Clerk-Recorder* (available at: https://www.rivcoacr.org/obtaining-record-copies)
[43] Office of Environmental Health Hazard Assessment 2018 *C.E.P.A. CalEnviroScreen 3.0* (available at: https://oehha.ca.gov/calenviroscreen/report/calenviroscreen-30)
[44] U. C. Bureau 2019 2019 American community survey (available at: www.census.gov/programs-surveys/acs)

[45] US Census Bureau 2016 Cartographic boundary shapefiles-census tracts (available at: https://census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html)

[46] California Energy Commision 2015 California building climate zone areas (available at: http://energy.ca.gov/maps/renewable/building_climate_zones.html)

[47] Chen M, Sanders K T and Ban-Weiss G A 2019 A new method utilizing smart meter data for identifying the existence of air conditioning in residential homes *Environ. Res. Lett.* **14** 094004

[48] Gastwirth J L 1972 The estimation of the Lorenz curve and Gini index *Rev. Econ. Stat.* **54** 306