Energy & Buildings 180 (2018) 146–158

Contents lists available at ScienceDirect

Energy & Buildings

journal homepage: www.elsevier.com/locate/enbuild

The role of household level electricity data in improving estimates of the impacts of climate on building electricity use

Mo Chen, George A. Ban-Weiss*, Kelly T. Sanders*

Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, CA 90089, USA

ARTICLE INFO

Article history:
Received 12 February 2018
Revised 22 August 2018
Accepted 8 September 2018
Available online 19 September 2018

Key words:
Residential electricity consumption
Temperature indicator
Spatiotemporal resolution
Temperature dependence
Urban heat islands
Climate change

ABSTRACT

Prior studies conclude that climate plays one of the most important roles in driving variations in residential electricity consumption. While some past studies have quantified sensitivities of electricity use to ambient temperature, 1) few previous studies utilize both high temporal and spatial resolution electricity data, and 2) no research to our knowledge has investigated how the temporal and spatial resolution of electricity data, and choice of ambient temperature indicators, affects quantification of these sensitivities. In this study, we use smart meter data records of electricity use for 1245 households across California, along with hourly ambient temperature records, to compute electricity–temperature sensitivities using a segmented linear regression approach. We find that electricity use and temperature show the strongest relationships when computed using daily accumulated electricity use and daily average temperatures; using these metrics results in a mean electricity–temperature sensitivity of 0.11 kW °C^{-1}. This value is higher than corresponding sensitivities computed using spatially aggregated data, with values ranging from 0.097–0.10 kW °C^{-1} depending on the amount of spatial aggregation. Through presenting probability density functions of household-level electricity–temperature sensitivities, we illustrate insights that can be gleaned using high resolution electricity datasets such as that used here. We note that values of electricity–temperature sensitivity reported here are representative of the 1245 households under investigation.

1. Introduction

The United States (US) residential sector accounts for about 37% of total US electricity consumption—more than any other end-use sector—making it a target for energy efficiency and power grid reliability interventions in recent years [1]. Household electricity demand increased by 16.5% [2] between 2001 and 2015 in the US and is projected to increase by 8% and 11% between 2015 and 2040 with and without the Clean Power Plan, respectively [3]. Much of this increase is expected to come from increases in space cooling demand. In 2016, space cooling and heating together was the largest end use of electricity, representing nearly 18% and 7% of US residential sector electricity demand [4,5], respectively. Demand for cooling is expected to increase by 11% between 2015 and 2040 in the US, outpacing the average projected rate of increase (8%) in total electricity consumption [3]. Although these increases are significant, nearly 90% of US homes already have air-conditioning (AC), which is very high compared to other regions of the world. Exploding global demand for AC, combined with increasing urbanization, is expected to bring cooling to billions of people in the coming decades, which poses large questions regarding the impact that these new electricity demands will have on global energy demand and greenhouse gas emissions [6,7].

Although a large number of factors impact residential electricity consumption, climate has been shown to play one of the most important roles in driving variations in residential electricity consumption [5,8–11]. Because of the diverse nature of the residential sector, analyzing the sensitivity of electricity demand to ambient temperature across the residential sector presents unique challenges compared to other sectors [12,13]. Households tend to have larger spatiotemporal variations in electricity consumption compared to other sectors, driving more uncertainty in prediction [14], presumably due to factors such as highly variable housing stock characteristics, appliances and other energy consuming device selections, occupant behavioral patterns, heating and cooling sources, energy prices, demographic factors, and other socio-economic indicators, which can vary significantly across regions [15–19]. Thus, to maximize our understanding of factors affecting residential sector electricity use, energy-climate sensitivity should be derived using data at the household level so that variability across residences can be observed and analyzed.

* Corresponding authors.
E-mail addresses: chen921@usc.edu (M. Chen), banweiss@usc.edu (G.A. Ban-Weiss), ktsanders@usc.edu (K.T. Sanders).

https://doi.org/10.1016/j.enbuild.2018.09.012
0378-7788/

Coarse data resolution has been a limiting factor in the majority of prior research endeavors in this field. Although there are a few studies that have used high resolution data, past studies typically rely on daily, monthly or annual electricity consumption data that might be insufficient for resolving the relationship between climate and electricity consumption [8]. Additionally, most studies utilize relatively coarse climatic data to represent relatively broad spatial extents, often ranging from megacity- to country-wide in scale [12,20,21]. These spatial scales are not sufficient for building a highly resolved understanding of climate-driven variations in energy consumption behavior. Using datasets that have coarse spatial and/or temporal resolution can average out important insights and cause loss of valuable information, especially for studies addressing the residential sector. Confounding analysis of the residential sector is the fact that residential homes have greater daily and seasonal variations in electricity use than other sectors, distinguishing the residential sector as the most difficult to analyze due to high amounts of variability and uncertainty [12–14,19,22,23]. Relying on temporally aggregated data, in particular, diminishes the ability to gain insight on electricity consumption patterns, which can lead to uncertainties in quantifying electricity–temperature relationships. Although this issue has been partially addressed by several recent studies using either high temporal resolution data (e.g. hourly or sub-hourly) [10,24–28] or high spatial resolution data (e.g. at building level) [24,29,30], knowledge gaps still exist.

Our main insights based on a survey of existing literature in this field (detailed below in Section 2) are that previous studies: 1) use datasets that vary widely in spatiotemporal resolution, spanning hourly to yearly resolution, across various spatial regions of interest; 2) rarely utilize electricity datasets that are both highly temporally and spatially resolved, and 3) utilize different types of electricity and temperature indicators to determine electricity–temperature sensitivities (e.g. hourly temperature, daily average temperature, daily minimum or maximum temperature, cooling degree days (*CDD*) or heating degree days (*HDD*), monthly average temperature, and some other derived indicators).

Despite these large methodological differences, no research to the authors' knowledge has investigated how electricity–temperature sensitivities vary according to the spatiotemporal resolution of electricity and climate data or choice in temperature indicators. While it is straightforward to assume that increasing data resolution is valuable to establishing refined and robust functional relationships between residential electricity usage and climate parameters, these increases in data resolution can cause large increases in the computational resource requirements of analysis, so gaining insight into these tradeoffs offer merit. Thus, research questions addressed in this study are as follows:

1) How does the spatiotemporal resolution of selected datasets affect the calculated relationship between residential electricity consumption and climatic parameters, i.e., ambient temperature?
2) How does the choice of temperature indicators affect the calculated relationship between residential electricity and ambient temperature?

Understanding and quantifying the functional relationships between residential electricity consumption and climatic parameters is crucial to developing effective energy conservation, peak energy management, and climate adaptation strategies, as well as informing meaningful and cost-effective power capacity investments in the future. Establishing robust electricity–temperature sensitivities is particularly important for future studies attempting to understand the role that phenomena such as climate change and the urban heat island effect might have on the power sector.

2. Literature review

Previous studies conclude that climate plays one of the most important roles in driving variability in residential electricity consumption [5,8–11]. In an effort to improve estimates of electricity–temperature relationships (hereafter referred to as "electricity–temperature sensitivity"), we conducted a survey of existing literature on this topic. Table 1 summarizes 24 publications in the literature analyzing climate-related influences on electricity consumption. These studies come from somewhat disparate fields including grid-scale electricity demand forecasting [10,31–35], building-level energy use modeling [24,29], and assessing the impact of climate change on electricity consumption [22,25,27–30,33,36–40]. The studies investigate regions in more than 40 countries and differ significantly according to research scope and objectives, data availability, researcher preferences on data metrics, and spatio-temporal resolution. Major modeling and data selection considerations across these studies are discussed in the sections below and resulting research objectives to be explored are then identified.

2.1. Models and observations

Studies have used different methods to quantify energy-climate relationships, including statistical techniques (e.g., regression) that relate energy use and climate indicators, and physics-based building energy modeling [41]. Statistical analyses offer advantages over other methods that rely on model-simulated data since they generally make use of real historical energy use and climate data. Regression models describe the relationship between a dependent variable, usually electricity consumption, and a temporally aligned independent variable, such as ambient temperature. Other climatic parameters such as humidity, wind speed, and solar insolation, have also been used as independent variables in multivariable regression analyses. Twenty out of 24 studies summarized in Table 1, representing the vast majority of analyses in this space, use regression methods. Within these 20 studies, 13 applied linear regression models [22,26,27,31,33–36,38,40,42–44], four applied non-linear regression models [24,30,45,46], and three applied a mixture of linear and non-linear models [28,29,47].

2.2. Electricity data metrics and resolution

Studies utilizing real-world electricity data have used a variety of metrics or indicators to characterize electricity use based on source datasets with widely varying spatiotemporal resolutions. While most studies surveyed use hourly, daily, monthly, seasonally, or yearly accumulated electricity usage data [22,25–32,34–36,38,43–45,47], six use peak electricity demand (i.e. electricity load during time periods of highest demand and electricity prices) [28,33,36,40,42,43], and one uses mean electric current intensity (Amperes) [10]. Ideally, the electricity indices utilized in a particular research study should reflect the research questions under investigation. For example, electricity usage data are most suitable for predicting future energy use trends and patterns (e.g. as a result of climate change or urban heat islands) [10,22,25,27,29–31,35,38,44,46,47], while peak electricity demand data are valuable for informing grid reliability [28,33,36,40,42,43].

The underlying resolution of these datasets is also an important driver of the accuracy in computed energy use-climate relationships [48,49]. Coarse spatial resolution has been a major limitation across the majority of prior research endeavors in this field. Past studies rely on electricity data at the sub-city [10], city [25,44,45,47,50], county [26,36], state [22,28,33,34,38,40], regional [46], or country levels [27,31,32,35,42,43] since house-level data have been less commonly available. However, these spatial scales

Table 1
Literature review of studies that investigate the influence of climate on electricity consumption.

Number	Model type	Temperature indicator	Stationary point temperature	Form of electricity data	Data temporal resolution	Data spatial resolution	Derived electricity–temperature sensitivity	Region	Time period	Citation
1	Tobit model (quadratic to *CDD*)	*CDD*	18.3 °C	Air conditioning electricity load	Three min interval[a] hourly[b]	Household (metered only at air conditioner)	Not reported	Pittsburgh	2010	(Horowitz, Mauch, and Sowell 2014)[24]
2	Recurrent neural network	*Humidex index* (derived from temperature and humidity)	N/A	Mean electric current intensity	Hourly[a,b]	Sub-city (a district in Italy)	Not reported	Italy	2002–2003	(Beccali et al. 2008)[10]
3	Time-series econometric model	Hourly temperature	N/A	Hourly electricity demand	Hourly[a,b]	City	0.3–0.5% per 1% temperature increase	Singapore	2003–2012	(Doshi et al. 2012)[25]
4	Linear regression	Hourly temperature	18 °C	Hourly electricity demand	Hourly[a,b]	Grid scale (similar to county)	6% $°C^{-1}$	Sacramento County (California)	08–08–2012 (one day)	(Pomerantz et al. 2015)[26]
5	Single-variable linear regression	*CDH*[c]	24 °C	Hourly electricity demand	Hourly[a,b]	Country	Mean hourly demand: 2.4–3.5% $°C^{-1}$ peak hourly demand: 2.8–4.2% $°C^{-1}$	Thailand	2004	(Parkpoom and Harrison 2008)[27]
6	Cubic regression for daily demand and linear regression for hourly peak demand	Average daily temperature & maximum hourly temperature	N/A	Daily electricity demand and hourly peak electricity demand	Hourly[a] Daily[b] & monthly[b]	State	Annual demand: 1.4–4.4% $°C^{-1}$ [d] daily peak demand: 1.7–5% $°C^{-1}$[d]	California	2004–2005	(Franco and Sanstad 2008)[28]
7	Non-linear regression (formulas not specified)	Daily average temperature and *CDD*/*HDD*	22 °C	Daily energy demand	Hourly[a] daily[b]	City	Summer daily demand: 0.6% $°C^{-1}$ [d]	Greece	1993–2001	(Giannakopoulos and Psiloglou 2006)[45]
8	Multivariable linear regression (also a semi-parametric function)	Daily average temperature	21 °C	Daily average or peak demand	Hourly[a] daily[b]	County	Hourly load: 1.6% $°C^{-1}$ [d] Daily peak demand 1.9% $°C^{-1}$ [d]	USA	2006–2014	(Auffhammer, Baylis, and Hausman 2017)[36]
9	Multivariable non-linear regression (cubic)	Daily max temperature	N/A	Daily peak demand	Daily[a,b]	Regional	Daily peak demand: 2.3% $°C^{-1}$ [d]	Canada	1991–1995	(Colombo, Etkin, and Karney 1999)[46]
10	Linear regression	Daily max temperature (only data points above 25°C are used)	N/A	Daily peak demand	Hourly[a] daily[b]	State	August peak demand: 5.6–7% $°C^{-1}$ [d]	California	1960–1990	(Sathaye et al. 2013)[33]
11	Multivariable linear regression	*CDD* & *HDD*	18°C	Daily electricity demand	Daily[a,b]	Country	Not reported	Spain	1983–1999	(Pardo, Meneu, and Valor 2002)[35]
12	Linear regression	Temperature at 8:00 and 14:00	N/A	Summer peak-hour electricity load	Daily[a,b]	Country	Daily peak demand: 2.6–2.7% $°C^{-1}$ [d]	Israel	1987–1988	(Segal et al. 1992)[42]
13	Linear regression	Daily average temperature	N/A	Daily energy demand	Daily[a,b]	Country	Daily average demand: 0.5% $°C^{-1}$	Netherlands	1970–1999	(Hekkenberg et al. 2009)[43]
14	Multivariable linear regression	*CDD* & *HDD*	18.5°C	Monthly electricity demand	Hourly[a] Daily[b] and monthly[b]	Country	Daily demand: 1.1–1.9% $°C^{-1}$ [d]	Greece	1993–2002	(Mirasgedis et al. 2006)[31]
15	Using both linear regression and physical models	*CDD* & *HDD*	Building specific (but only average is reported)	Monthly energy demand for commercial buildings	Monthly[a,b]	Buildings sampled in 101 cities	14% $°C^{-1}$ [d]	USA	1989	(Belzer, Scott, and Sands 1996)[29]
16	log-linear specification model	Number of days per year that the mean daily temperature falls in each temperature bin (every 5°F)	N/A	Monthly electricity demand of households	Monthly[a,b]	Household	9–13% $°C^{-1}$ [d]	California	2003–2006	(Aroonruengsawat and Auffhammer 2011)[30]

(*continued on next page*)

Table 1 (*continued*)

Number	Model type	Temperature indicator	Stationary point temperature	Form of electricity data	Data temporal resolution	Data spatial resolution	Derived electricity–temperature sensitivity	Region	Time period	Citation
17	Multivariable linear regression	*CDD*	Not specified	Monthly electricity demand	Monthly[a,b]	City	Monthly demand: 7.49% $°C^{-1}$	Bangkok, Thailand	2002–2006	(Wangpattarapong et al. 2008)[44]
18	Both quadratic and linear regression	Monthly average temperature	N/A	Monthly electricity demand	Monthly[a,b]	City	domestic: 8.9% $°C^{-1}$ commercial: 3.0% $°C^{-1}$ industrial: 2.0% $°C^{-1}$	Hong Kong	1990–2004	(Fung et al. 2006)[47]
19	Multivariable linear regression	*CDD* & *HDD*	18 °C	Summer Peak electricity demand	Daily[a] Monthly[b]	State	Not reported	California	1970–2005	(Lebassi et al. 2010)[40]
20	Time series multivariable linear regression	*CDD* & *HDD* (population weighted)	State specific	Monthly electricity demand	Monthly[a,b]	State	2.54% $°C^{-1}$ [d]	USA	2008–2012	(Huang and Gurney 2016)[22]
21	Multivariable linear regression	*CDD* & *HDD*	53°F-71°F (different across fuels and sectors)	Monthly electricity demand	Monthly[a,b]	State	residential: 0.1% $°F^{-1}$ commercial: 0.04% $°F^{-1}$	Maryland	1977–2001	(Ruth and Lin 2006)[38]
22	Multivariable linear regression	Monthly average temperature	Region specific	Monthly electricity demand	Monthly[a,b]	State	Summer monthly: 5.97–32.2 kWh per capita $°C^{-1}$ $month^{-1}$	8 states in USA	1984–1993	(Sailor and Muñoz 1997)[34]
23	Panel analysis models	Average seasonal temperature	N/A	Monthly energy demand (including electricity)	Monthly[a] Seasonally[b] and yearly[b]	Country	Not reported	31 countries around the world	1978–2000	(Bigano, Bosello, and Marano 2006)[32]
24	Not specified	Annual average temperature	N/A	Grid load at a specific time	Not specified[a,b]	City	1. Peak demand: FL: 6% $°C^{-1}$ AL: 3% $°C^{-1}$ West TX: 6%$°C^{-1}$ NM: 3% $°C^{-1}$ AZ: 1% $°C^{-1}$ Southern CA: 3% $°C^{-1}$ Northern CA: 1.5% $°C^{-1}$ 2. Annual usage: FL: 3% $°C^{-1}$ AL: 1.5% $°C^{-1}$ West TX: 3% $°C^{-1}$ NM: 0.5% $°C^{-1}$ AZ: 6% $°C^{-1}$ Southern CA: 1.5% $°C^{-1}$ Northern CA: 0.5% $°C^{-1}$	Different cities in US	1986	(Akbari 1992)[50]

Note:
[a] Source data resolution
[b] Processed data resolution
[c] *CDH*: Cooling Degree Hours, defined by the cited literature as: "a short–term version of *CDD* described by: $\sum_{h=1}^{N}(T_h - T_b)$ *for* $T \geq T_b$ *and* 0 *otherwise*, where N is the number of hours in the period of interest, T is the air temperature, and T_b is the cooling base temperature, commonly taken to be 24 °C in Thailand."
[d] Value is calculated from percentage or absolute change in electricity consumption under different climate change scenarios versus a baseline period.

are not sufficient for building a highly resolved understanding of climate-driven variations in energy consumption behavior, especially for regions with large climatic variations, such as those adjacent to mountains and coasts like the Los Angeles basin.

The temporal resolution of datasets also varied considerably across the surveyed studies. For source data resolution, eight of the 24 studies use monthly aggregated electricity usage [22,29,30,32,34,38,44,47], five use daily aggregated electricity usage [35,37,40,42,46], nine use hourly electricity data [10,25–28,31,33,36,45], one uses sub-hourly meter data at the appliance level instead of the entire household [24], and one study does not specify data resolution [50]. The majority of studies have source data and processed data of the same resolution, but eight utilize processed data at a coarser resolution than source data [24,28,31–33,36,40,45], meaning that the researchers chose to aggregate their datasets prior to analysis.

Although a few studies address the importance of high resolution data [36,48,49], typically no justification is provided on why one resolution is chosen over other possible resolutions. It is assumed that temporal resolution reflects the availability of source data in most cases. Most studies rely on data being shared by utility companies or grid operators, so resolution is constrained by the data provided to researchers [10,25–28,31,33,35–37,42,44–47]. Relatively low temporal resolution data (e.g. monthly or yearly averages) have traditionally been most easily acquired from technical reports or bills [29,30]; while the widespread dissemination of smart electricity meters has enabled the collection of hourly electricity data, few studies have had access to these data for analysis [24].

2.3. Temperature data metrics

Prior studies have used a variety of indicators for characterizing climate. For example, recent studies have utilized a range of temperature metrics, including Cooling/Heating Degree Days [22,24,27,29,31,35,38,40,44], hourly temperature [25,26], daily average temperature [28,36,43,45], daily max temperature [33,46], monthly average temperature [34,47], seasonal average temperature [32], histograms of daily temperature [30,50], and other indices derived from temperature data [10]. Of the literature surveyed in Table 1, 10 out of 24 used *CDD*/*HDD* [22,24,27,29,31,35,38,40,44,45] and six used daily average or max temperature [28,33,36,43,45,46], suggesting that daily temperatures have been the most commonly utilized resolution in this body of literature.

2.4. Stationary point temperatures

Studies utilizing *CDD* and *HDD* (see Section 3.2 and Eq. (2) for more details on *CDD*/*HDD*) as a temperature indicator need to choose a pre-defined, fixed threshold temperature to calculate this metric. The threshold (also sometimes called a "stationary point temperature" or "base temperature") refers to the temperature below (above) which no cooling (heating) is needed (discussed in more detail in Section 3.2). In past studies, 18 °C is the most common threshold temperature, chosen by five out of 13 studies that use *CDD* and/or *HDD* [24,26,31,35,40]. 60 °F (15.6 °C), 21 °C, 22 °C, and 24 °C are also used in past studies [27,36,38,45]. Three studies assign specific stationary point temperatures to different buildings or regions [22,29,34]. One study does not specify stationary point temperature [44]. Several methods have been applied to set a stationary point temperature, including: 1) choosing the temperature threshold arbitrarily; 2) referencing a previous study in the same or neighboring region; and 3) extracting it from a preliminary electricity–temperature plot. Only one study analyzes the impact of setting a region-specific stationary point temperature using a segmented regression model, but the study calculates this point at the state level only [22].

3. Methods

To address the research questions presented above, this study utilizes a dataset representing the hourly electricity consumption of 1245 households across California for a one-year period. We also utilize data from a network of 145 weather stations to assess hourly temperatures in locations adjacent to each home. A segmented linear regression model is applied to assess the electricity–temperature sensitivity of each household. The electricity data are spatially and temporally aggregated in various ways (i.e. both before and after computing electricity–temperature sensitivity) to assess how data resolution impacts electricity–temperature sensitivity. In addition, the dependence of chosen temperature indicators on computed sensitivities is assessed. The dataset used here includes only residential homes and thus varies from many previous studies using spatially aggregated datasets, which would also include commercial and industrial buildings.

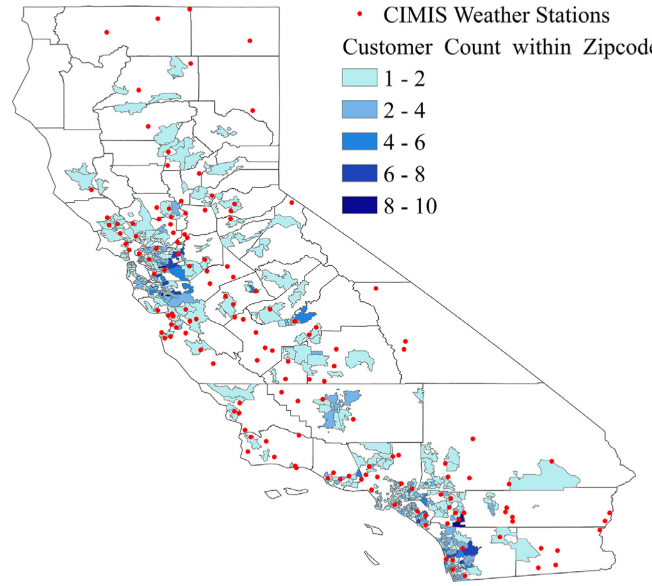

Fig. 1. Map showing locations of the 1245 residential electricity customers (shown as number of households per zip code) and 145 CIMIS weather stations considered in this study. Each household was linked to a weather station based on shortest distance.

3.1. Datasets

Hourly smart meter data records of electricity usage at the household level from 1245 residential customers (after data cleaning and screening) across California were analyzed. These households reflect utility customers that voluntarily downloaded an energy-related smart phone app for tracking their electricity use. Since this sample is likely biased towards energy-conscious households, this paper focuses on comparing methods for computing energy-temperature sensitivity but does not claim that computed values are representative of the general population of California cities. Only zip code information for each household included in the dataset was provided to protect customer privacy. Several procedures for data cleaning and screening were carried out. First, to fully capture the year-round relationship between residential electricity consumption and ambient temperature, only customers with one full year of electricity data (05/18/2015–05/17/2016) were included in the study. Second, households that could be identified as energy generators (e.g., with solar photovoltaic installations) were removed from the dataset to reduce the impacts of these households on load curves. We defined these generating households as those that had a negative value of electricity usage at any time; thus, households whose onsite generation never exceeded their own energy use through the period of study would not be flagged and is a limitation of the study.

The households included in the dataset spanned 41 counties, 549 zip codes, and 15 of 16 climate zones in California [51]. These climate zones were established by the California Energy Commission (CEC) and specifically characterize building energy use under various climate characteristics [51].

Hourly ambient temperature data over the year under investigation were retrieved from the California Irrigation Management Information System (CIMIS) [52], which includes a network of over 145 automated weather stations in California covering most of the state's population centers. Geospatial analysis was executed with ArcGIS (10.4.1, ESRI, Redlands, CA, USA) to map each household with an electricity record to its nearest weather station. Fig. 1 illustrates the number of homes per zip code and location of weather stations of this study.

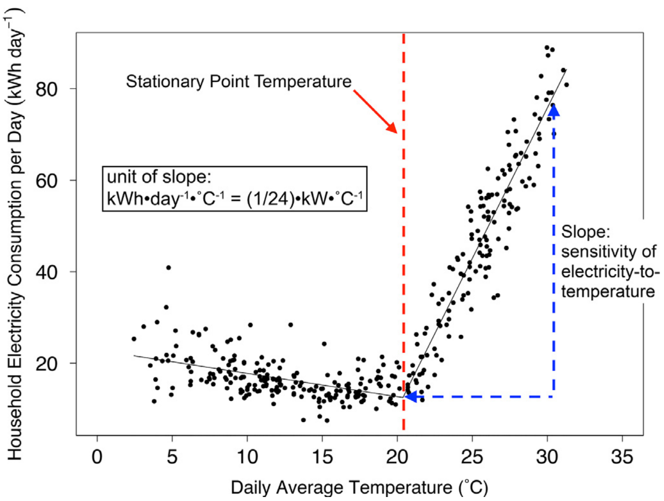

Fig. 2. Example of a home in Clovis, CA illustrating the stationary point temperature (*SPT*) and electricity–temperature sensitivity through a segmented linear regression method.

It is important to note that this dataset, which includes utility records averaging 30 households per county and less than three households per zip code, is not statistically representative of the population from a spatial perspective. Accordingly, the objective of this study is not to define the effects of climate on electricity across regional boundaries (i.e. city, county, climate zone); rather, the goal is to assess the influence of spatiotemporal electricity data resolution and climate indicators on derived electricity–temperature sensitivity (see research questions in the Introduction).

3.2. Statistical models

The nonlinearity of relationships between building energy consumption and ambient temperature has been established in previous studies [28,34,45–47]. Nonlinear regression models (i.e. polynomial functions) have been developed that can achieve a good fit among variables. However, sophisticated models may have issues with overfitting and can thus fail to generalize trends under investigation and are also not applicable to other regions [53]. To address the nonlinear relationship between electricity use and ambient temperature in residential homes, while avoiding overfitting, a segmented linear regression model proposed by [54] will be utilized in this analysis.

The segmented linear regression reveals two important pieces of information. The first is the stationary point temperature (*SPT*), which sits at the stationary point of the piece-wise linear function and can be thought of as analogous to the base temperature in the *CDD* method. In other words, stationary point temperature is the temperature at which household electricity consumption reaches a minimum, with the assumption that no cooling or heating is needed at this temperature. In the segmented regression model, the stationary point is calculated iteratively to determine the best overall piece-wise linear fit of the original dataset. The second is the slope of the linear regression to the right of the stationary point temperature (referred to in this analysis as the "electricity–temperature sensitivity"), representing the change in electricity consumption that corresponds to a change in ambient temperature of one degree Celsius. Electricity–temperature sensitivity can be affected by factors like house size, insulation, behavior, etc., since these factors also affect air-conditioning use.

Fig. 2 shows an example segmented regression for a home in Clovis, California. Daily aggregated electricity usage is plotted against daily average temperature, and stationary point temperature and electricity–temperature sensitivity are illustrated. The plot in Fig. 2 is thus divided into two regimes: (1) strong positive sensitivity between electricity use and temperature to the right of the stationary point temperature, and (2) electricity use that is relatively insensitive to temperature change to the left of the stationary point temperature. In California, cooling energy demand from air conditioning is driven by electricity while heating is mainly supported by natural gas [55], which is why there is not a strong increase in electricity use as temperatures decrease below the stationary point. For the same reason, only one stationary point is identified in the segmented model, whereas in some regions there might be two (e.g., in the case of Israel described in [56]).

To address research question 1, stationary point temperature and electricity–temperature sensitivity are computed for different spatial aggregation levels using a segmented regression defined as:

$$\frac{\boldsymbol{E_{s,t}}}{t} = \begin{cases} \alpha_1 + \beta_1 \times \boldsymbol{T_{s,t}} \pm \epsilon_1, & \boldsymbol{T_{s,t}} < SPT_{s,t} \\ \alpha_2 + \beta_2 \times \boldsymbol{T_{s,t}} \pm \epsilon_2, & \boldsymbol{T_{s,t}} \geq SPT_{s,t} \end{cases} \tag{1}$$

where $\boldsymbol{E_{s,t}}$ is a vector of residential electricity consumption over a period of time t (vertical axis in Fig. 2). $\frac{\boldsymbol{E_{s,t}}}{t}$ is expressed in units of electric power (kW). $\boldsymbol{T_{s,t}}$ is a vector of near-surface ambient temperatures (in the units of °C) over the same period of time (horizontal axis in Fig. 2). The first row of Eq. (1) describes the relationship between $\frac{\boldsymbol{E_{s,t}}}{t}$ and $\boldsymbol{T_{s,t}}$ for $\boldsymbol{T_{s,t}} < SPT_{s,\,t}$. The second row of Eq. (1) describes the relationship between $\frac{\boldsymbol{E_{s,t}}}{t}$ and $\boldsymbol{T_{s,t}}$ for $\boldsymbol{T_{s,t}} \geq SPT_{s,\,t}$. The electricity–temperature sensitivity, $S_{s,t}$, is defined as the slope of the regression line above the *SPT* (i.e., β_2), with units of kW °C^{-1}. *SPT* is calculated by iteratively locating the intersection of the two linear regions to maximize the model's overall coefficients of determination (r^2). Thus, in Eq. (1), $\boldsymbol{E_{s,t}}$ and $\boldsymbol{T_{s,t}}$ are inputs to the segmented regression and all other variables are outputs. (Note: α_1, α_2, and β_1 are additional regression coefficients and ϵ is the error term.)

The spatial and temporal aggregations of data represented in vectors $\boldsymbol{E_{s,t}}$ and $\boldsymbol{T_{s,t}}$, as well as scalars $S_{s,t}$, and $SPT_{s,\,t}$, are indicated by subscripts s and t, respectively. Values of subscript s in this study include household, city, county, and climate zone, and t can be hourly or daily. For example, if $s =$ household for $\boldsymbol{T_{s,t}}$, then $\boldsymbol{T_{s,t}}$ corresponds to the observed temperature at the nearest weather station for that home, while if $s =$ city, $\boldsymbol{T_{s,t}}$ corresponds to the population weighted spatial mean observed temperature for that city. By "population", we mean "number of homes", so "population-weighted spatial mean" means we take the average of temperature readings from multiple weather stations in the area, weighted according to how many homes are assigned to each weather station. (Note that we discuss various daily temperature metrics below when discussing research question 2.)

For $s =$ household (i.e. no spatial aggregation), segmented linear regression is conducted separately for each of 1245 homes using their hourly or daily aggregated electricity consumption (i.e. depending on t). The mean values of both stationary point temperatures ($\overline{SPT}_{s=\text{household},t}$) and sensitivities ($\bar{S}_{s=\text{household},t}$) are computed by taking the mean over all 1245 households of computed stationary point temperatures and sensitivities. For $s =$ city, county, or climate zone, the segmented linear regressions are carried out using spatially averaged electricity consumption over spatial extent s, along with population-weighted temperature to mimic studies using more spatially aggregated data to compute sensitivities. The mean values $\bar{S}_{s,t}$ and $\overline{SPT}_{s,t}$ are then computed by taking the population-weighted average of all city, county, or climate zone level sensitivities and stationary point temperatures (i.e. depending on s), respectively. For example, to compute $\bar{S}_{s=\text{city},t=\text{daily}}$, we first compute spatially aggregated electricity use for each city, and then the city level (hourly) electricity data is accumulated to daily

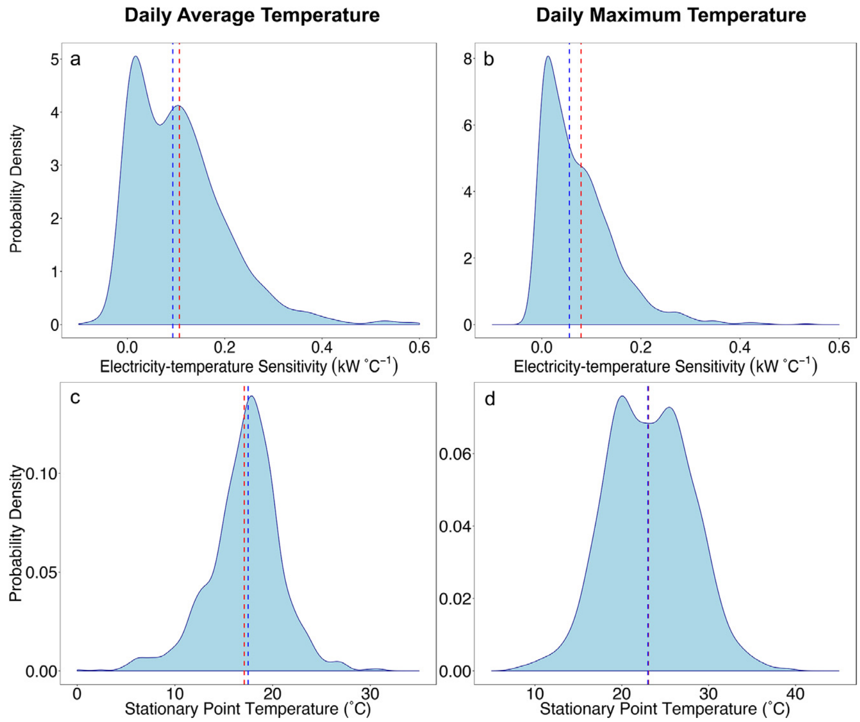

Fig. 3. Probability density distributions of electricity–temperature sensitivities (a, b) and stationary point temperatures (c, d) of all 1245 homes in this study's dataset. Red dashed lines indicate the mean values of all 1245 homes. Blue dashed lines indicate the median values of all 1245 homes (Note that in panel d, the blue dashed line partially overlaps the red dashed line.) Sensitivity in units of kW °C^{-1} can be converted to kWh day^{-1} °C^{-1} by multiplying by a factor of 24. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

resolution. For temperature, first the metric of choice (see the next paragraph) is computed for each weather station (i.e. daily minimum, average, or maximum), and then city population-weighted averages are computed. Then, segmented linear regressions are applied using the averaged data per city to compute city-level stationary point temperatures and electricity–temperature sensitivities. Lastly, $\overline{S}_{s=\text{city},t=\text{daily}}$ and $\overline{SPT}_{s,t}$ are computed by taking the mean of city-level values.

To address research question 2, the relationship between residential electricity use and various temperature indicators (i.e., hourly temperature, daily average temperature, daily maximum temperature, daily minimum temperature, and *CDD*) are explored. To quantify the effect of utilizing different temperature indicators on computed electricity–temperature sensitivities, we carry out the segmented linear regression using hourly temperature (i.e. $t=\text{hourly}$), daily maximum, daily minimum, and daily average temperature (i.e. for $t=\text{daily}$). This comparison is carried out for both a typical household in San Jose, and also for all households (within our dataset) in the City of San Jose. We also compute sensitivity using *CDD* assuming a uniform base temperature $T_b = 18$ °C for all homes. Since *CDD* calculations already include base temperatures, a standard linear regression model is applied rather than the segmented regression; the slopes of these linear regressions represent the electricity-*CDD* sensitivity. The coefficient of determination (r^2) values of these regression models are compared to assess the quality of fit.

CDD is computed as:

$$CDD = \frac{\int_{day} \max\left[0,\ T_{s,t=\text{hourly}}(h) - T_b\right] dh}{24} \tag{2}$$

$T_{s,t=\text{hourly}}(h)$ is the hourly ambient temperature for hour h expressed in °C and T_b is the base temperature (i.e. 18.0 °C in this study). The daily value of *CDD* can be obtained by integrating $T_{s,t=\text{hourly}}(h)$ over each day as done in [57]. Physically, the base temperature is the ambient temperature at which a building's heat loss and heat gain reaches an equilibrium, such that cooling is not needed. The base temperature is often chosen based on previous studies that focus on a similar geographical zone or is set arbitrarily. Due to different climate zones, building characteristics, and occupant behavior patterns, the base temperature can vary significantly among spatial areas [22,34]. This issue has been identified by several previous studies [10,22,29,34,38] included in Table 1. In our study, *CDD* is calculated as the cumulative degrees beyond 18.0 °C for each hour on a daily basis. We note that computing electricity-*CDD* sensitivities using linear regression is analogous to that of electricity–temperature sensitives using segmented linear regression with a fixed stationary point temperature.

4. Results

4.1. Sensitivity and stationary point temperature distribution

One of the biggest advantages to using household level electricity consumption data to derive electricity–temperature relationships is that these data illustrate home-to-home variability in terms of (a) the ambient temperatures at which homes start using increased electricity (i.e. stationary point temperature), and (b) the amount of additional electricity that homes use as temperatures increase beyond the stationary point. Probability density distributions of electricity–temperature sensitivities (top row) and stationary point temperatures (bottom row) for all households investigated here are plotted in Fig. 3. Subplots *a* and *c* in this figure use daily average temperature, while subplots *b* and *d* use daily maximum temperature. Both distributions of sensitivities are skewed to the right.

The long tails of the probability density distributions of electricity–temperature sensitivities represent homes that have large increases in electricity consumption as temperatures increase above the stationary point (Fig. 3 (a) and (b)). When daily maximum temperature is used in the segmented regression, 47% of households in the dataset have a sensitivity less than 0.05 kW °C^{-1}, while 24%, 24%, and 6% of households have a sensitivity value of 0.05 to 0.1, 0.1 to 0.2, and over 0.2 kW °C^{-1}, respectively. For daily average temperature, 34% of households in the dataset have a sensitivity less than 0.05 kW °C^{-1}, while 19%, 32%, and 15% of households have a sensitivity value of 0.05 to 0.1, 0.1 to 0.2, and over 0.2 kW °C^{-1}, respectively. Both temperature indicators have similar distribution shapes, but daily average temperature leads to overall higher sensitivity values than daily maximum temperature. In other words, daily electricity consumption at the household level is generally more sensitive to daily average temperature than daily maximum temperature.

If daily maximum temperature is used, the stationary point temperatures of the 1245 homes are distributed within a range from about 10–35 °C with a mean value of 23.1 ± 4.9 °C (73.6 ± 8.9 °F) (Fig. 3(d)). For daily average temperature, the distribution of stationary point temperatures is almost normal within a range from about 5–25 °C and more concentrated to the mean value, which is 17.1 ± 3.9 °C (62.8 ± 7.0 °F) (Fig. 3(c)). It is interesting that a small percentage of homes have negative or zero sensitivity values in Fig. 3(a) and (b). This can be attributed to a lack of cooling devices in these homes, or on-site energy generation (e.g., solar photovoltaics). Also, a small number of homes have stationary point temperatures less than 10 °C in Fig. 3(c) and (d), which appears anomalously low. One possible explanation is differences between ambient and indoor air temperature due to solar heating; in this case, indoor temperatures may be higher than ambient, causing inhabitants to turn on air conditioners at lower ambient temperatures than expected. More information about building design is needed to further explore this possibility. Homes without cooling devices could also be the cause, with lower than expected stationary point temperatures being identified for reasons other than increasing cooling energy use.

4.2. Impact of temperature indicators on computed electricity–temperature sensitivity

The impact of using various temperature indicators on computed electricity–temperature sensitivity is illustrated using (a) electricity data for a typical household in San Jose, and (b) averaged electricity consumption for all households for which we have data in San Jose (Fig. 4). The city of San Jose is chosen because our dataset includes a relatively large number of homes ($n = 80$) compared to other cities. In Fig. 4, the first row shows hourly electricity consumption versus hourly temperature. The other rows show daily accumulated electricity consumption versus daily minimum temperature, daily average temperature, daily maximum temperature, and cooling degree days at 18 °C (*CDD18C*) both including and excluding days with *CDD18C* = 0. (Days with *CDD18C* = 0 occur when hourly temperatures remain below 18 °C.)

The temperature indicator utilized significantly affects the coefficient of determination (r^2) and the computed electricity–temperature sensitivity. Overall, using hourly electricity and temperature data shows weak coefficients of determination relative to the daily metrics. Among the daily metrics (i.e. daily minimum, average, and maximum temperature), daily average temperature is shown to lead to (a) the highest coefficients of determination for both the typical home and all homes in San Jose, and (b) highest sensitivity. Daily maximum temperature and daily minimum temperature lead to the second and third highest sensitivities among the daily metrics. Linear regressions of daily aggregated electricity use versus *CDD18C* (including days with *CDD18C* = 0) show similar sensitivity as daily average temperatures for both the individual home and city of San Jose. The coefficient of determination and sensitivity increases when days with *CDD18C* = 0 are removed from the regression.

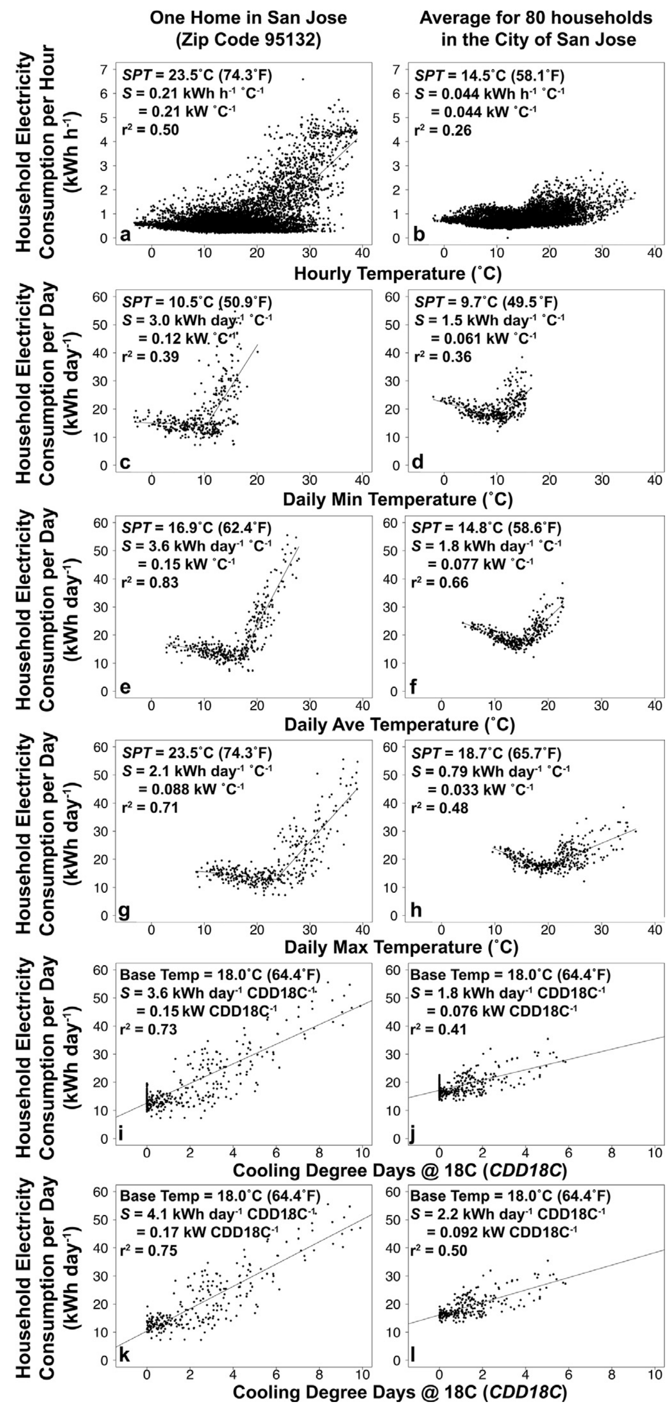

Fig. 4. Segmented linear regression applied to a single household (left column) and the average of households in our dataset ($n = 80$) within the City of San Jose, California (right column) using various temperature indicators: hourly temperature (a, b), daily minimum temperature (c, d), daily average temperature (e, f), daily maximum temperature (g, h), *CDD18C* including days with *CDD18C* = 0 (i, j), and *CDD18C* without days where *CDD18C* = 0 (k, l). *SPT* corresponds to Stationary Point Temperature and *S* corresponds to electricity–temperature sensitivity. In panel (i–l), Base Temp corresponds to the base temperature, which can be seen as a prescribed stationary point temperature.

Table 2
Electricity–temperature sensitivities and stationary point temperatures computed with various spatial aggregations using electricity consumption data for 1245 California households.

Resolution of electricity use data, $\boldsymbol{E}_{s,t}$, used in the segmented regression analysis	Mean value of electricity–temperature sensitivity $\bar{S}_{s,t=\text{daily}}$ (kW °C^{-1})	Standard deviation of electricity–temperature sensitivity (kW °C^{-1})	Mean value of stationary point temperature $\overline{SPT}_{s,t=\text{daily}}$ in °C (°F)	Standard deviation of stationary point temperature in °C (°F)
Using daily maximum temperature				
s = household[a]	0.079[f]	0.18	23.1 (73.6)	4.93 (8.87)
s = city[b]	0.066	0.058	22.4 (72.3)	4.46 (8.03)
s = county[c]	0.063	0.040	22.2 (71.9)	3.21 (5.78)
s = climate zone[d]	0.063	0.036	22.3 (72.1)	3.30 (5.93)
s = state[e]	0.064	N/A	22.1 (71.8)	N/A
Using daily average temperature				
s = household	0.11[g]	0.10	17.1 (62.8)	3.86 (6.95)
s = city	0.098	0.082	17.1 (62.8)	3.50 (6.30)
s = county	0.098	0.058	17.2 (63.0)	2.08 (3.74)
s = climate zone	0.097	0.040	17.2 (63.0)	2.64 (4.75)
s = state	0.10	N/A	17.0 (62.6)	N/A

[a] Segmented regression was performed for each household, and then sensitivity and stationary points per home were averaged for all homes in California.
[b] Electricity data were averaged by city, segmented regression was performed for each city, and then sensitivity and stationary points were population-weighted averaged for all cities.
[c] Electricity data were averaged by county, segmented regression was performed for each county, and then sensitivity and stationary points were population-weighted averaged for all counties.
[d] Electricity data were averaged by climate zone, segmented regression was performed for each climate zone, and then sensitivity and stationary points were population-weighted averaged for all climate zones.
[e] Electricity data were averaged for entire state of California and then segmented regression was performed for state-averaged data
[f] Equivalent to 10.5% change in electricity consumption °C^{-1} (% change means the relative change in electricity consumption per °C increase using the consumption at the stationary pointe temperature as a baseline).
[g] Equivalent to 15.3% change in electricity consumption °C$^{-1.}$

4.3. Impact of spatial aggregation on computed electricity–temperature sensitivity

Mean values of stationary points ($\overline{SPT}_{s,t=\text{daily}}$) and sensitivities ($\bar{S}_{s,t=\text{daily}}$) for California derived using data with different spatial aggregation levels (i.e. household, city, county, climate zone, and state) are displayed in Table 2. Sensitivity values are calculated using both daily maximum temperature and daily average temperature for comparison purposes.

When daily maximum temperature is used, the mean value of sensitivities calculated using household level electricity data (i.e. no spatial aggregation) $\bar{S}_{s=\text{household},t=\text{daily}}$ is 0.079 kW °C^{-1}, about 19% higher than computing sensitivities where s is spatially aggregated to the city, county, climate zone or state-level, which range from 0.063 to 0.066 kW °C^{-1}, depending on the level of spatial aggregation (Table 2). A similar phenomenon is also observed using daily average temperature. The mean value of sensitivities computed for the 1245 homes using household level electricity data is 0.11 kW °C^{-1}, higher than that using aggregated data, which ranges from 0.097 to 0.10 kW °C^{-1}.

The level of spatial aggregation affects electricity–temperature sensitivity more than stationary point temperature (up to 19% for sensitivity vs. 4% for stationary point temperature by using daily maximum temperature, and up to 6% vs. 1% by using daily average temperature). Using daily maximum temperature, the mean value of computed stationary point temperature $\overline{SPT}_{s=\text{household},t=\text{daily}}$ for all 1245 households is 23.1 °C (73.6 °F). Using electricity data that are spatially aggregated, stationary point temperatures $\overline{SPT}_{s=\text{city/county/climate zone/state},\ t=\text{daily}}$ are slightly lower, ranging from 22.1 to 22.4 °C (71.8 to 72.3 °F). Using daily average temperature, the stationary point temperature is 17.0–17.2 °C (62.6–63.0 °F), regardless of level of aggregation.

5. Discussion

5.1. Advantages of utilizing high spatiotemporal resolution data

Using high spatiotemporal resolution electricity and climate data to investigate the effects of climate variability on energy consumption offer advantages over using aggregated data. From a research perspective, having access to household-level data enables the ability to investigate how data resolution influences computed electricity–temperature interactions. Table 2 indicates that computed electricity–temperature sensitivity is dependent on the level of spatial aggregation of the data used in the segmented linear regressions. For example, our research suggests that computing the electricity–temperature sensitivity using household data and then averaging all households in a state results in a different sensitivity value than computing the sensitivity using state-mean electricity data, as illustrated in Table 2. In addition, using electricity data at the household level is ideal for most accurately calculating electricity–temperature sensitivities given that more representative temperature data for each household can be used in the analysis. This is especially important for cities like Los Angeles that have strong spatial variability in climate.

Two case studies are presented here to explicitly illustrate how different electricity–temperature sensitivities can arise using household level versus aggregated data (see Fig. 5). In Case I, Household A (zip code 94583, San Ramon) has large daily non-cooling loads (i.e. electricity use to the left of the stationary point) and a large sensitivity. Household B (zip code 90504, Torrance) has relatively small daily non-cooling loads and a small corresponding sensitivity. A smaller sensitivity value is calculated if we take the mean value of sensitivities computed per household (i.e. $\bar{S}_{s=\text{household},t=\text{daily}}$) compared to performing the segmented regressions after aggregating the electricity consumption of two households (i.e. $\bar{S}_{s=\text{city},t=\text{daily}}$). In the latter case, the average sensitivity will be weighted towards Household A because of its higher electricity use and thus $\bar{S}_{s=\text{household},t=\text{daily}} > \bar{S}_{s=\text{city},t=\text{daily}}$. In Case II, Household C (zip code 92571, Perris) has overall high daily electricity use with relatively small sensitivity while Household D (zip code 92069, San Marcos) has small daily electricity consumption with relatively high sensitivity. In this case, a higher sensitivity value is calculated if we take the mean value of sensitivities computed per household compared to performing the segmented regression after aggregating the electricity consumption of two households ($\bar{S}_{s=\text{household},t=\text{daily}} < \bar{S}_{s=\text{city},t=\text{daily}}$). In either case, computing the mean sensitivity after spatial aggregation will weight

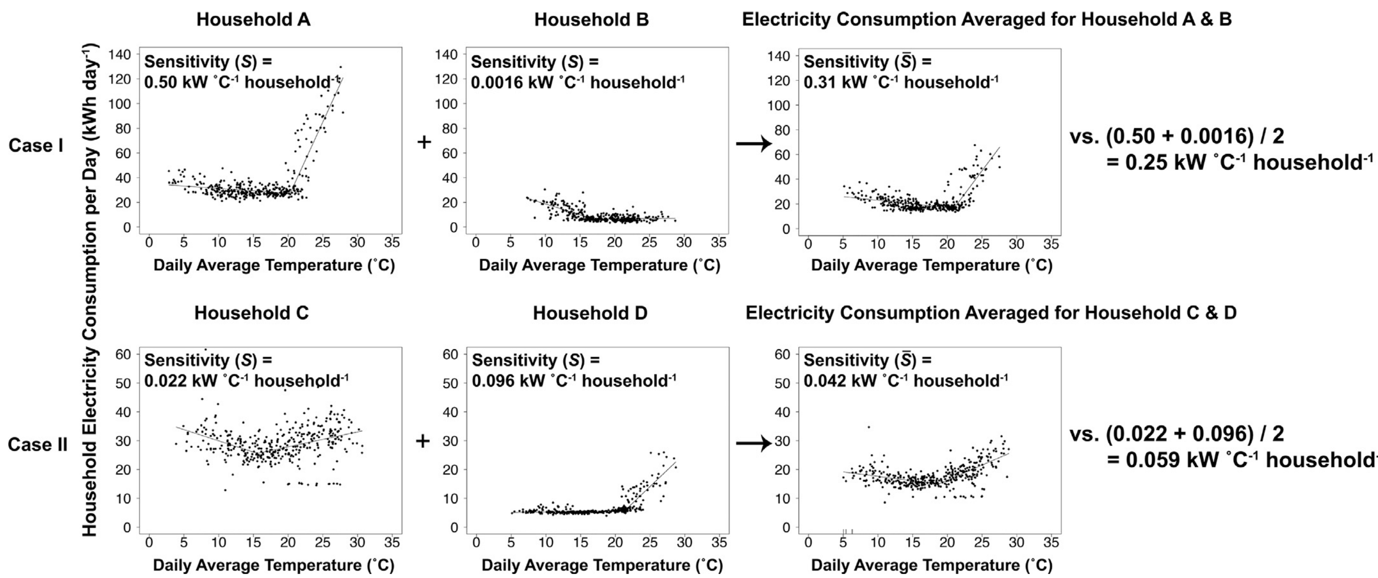

Fig. 5. Two case studies illustrating two households with different non-cooling electricity usages and sensitivities. In Case I (top row), Household A has higher daily electricity use and a larger electricity–temperature sensitivity value than Household B. In Case II (bottom row), Household C has a higher daily electricity use, but a smaller sensitivity than Household D. In either case, the average sensitivity for the two households, if calculated based on aggregated electricity use, will more heavily weight the larger electricity consumer.

big electricity consumers more heavily, and give less weight to smaller consumers regardless of their sensitivity values.

A second advantage to using high spatiotemporal resolution data is that they offer the ability to investigate the distribution of energy use patterns among different households, which in this study is reflected by stationary point temperatures and sensitivities. Fig. 3 illustrates that households in this sample have a wide distribution of sensitivities and stationary point temperatures. Variability in sensitivities are likely a result of variations in occupant behavior patterns, building and HVAC system characteristics, and climate zone. More information about building characteristics at the household level is needed to further quantify the relative importance of these causal factors of variability in sensitivity. We hypothesize that a large number of households show small sensitivity to ambient temperature change due to lack of air conditioning equipment presumably concentrated in coastal locations, and possibly also due to homes with relatively low square footage and/or occupants that cannot afford air conditioning. These hypotheses should be validated with additional datasets in future analyses. Fig. 3 also illustrates that different households have unique stationary point temperatures, which is an important distinction between the method used in this study and previous studies that assume fixed base temperatures that are not necessarily computed based on the dataset (e.g. *CDD18C*). Spatial variations in stationary point temperatures reflect building characteristics, occupant behavior, and climate variability [22]. If aggregated electricity data are used, only one stationary point temperature for the entire region can be computed and used in regressions.

5.2. Roles of temperature indicators

As indicated in Fig. 4 and Table 2, electricity–temperature sensitivities are dependent on the temperature indicator used in the regression. We suggest that the following considerations be used to help decide which temperature indicator is of interest.

First, electricity and temperature show the strongest relationships when computed using data at daily temporal resolution. Regressions between hourly household electricity consumption and hourly temperature result in relatively low coefficients of determination (r^2). This can be partially explained by the daily electricity

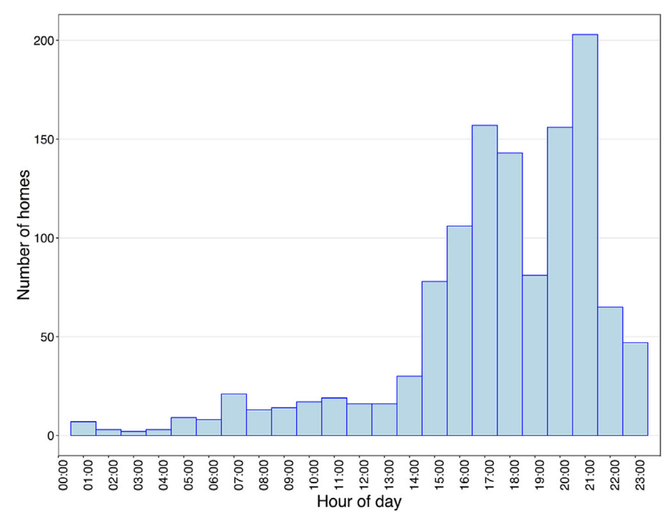

Fig. 6. Histogram of time of day corresponding to peak energy for each household during summer (July, August, September). Peak hourly electricity use occurs in the late afternoon to early evening for the majority of households in this study.

use patterns of residential homes, which can be heavily affected by household energy consumption behavior. For example, energy use patterns in many cases will not directly follow hourly temperatures since occupants that go to work during normal business hours may peak in their electricity usage in the evening when ambient temperatures are not at their daily peak. There can also be a timing lag between ambient outside temperature rise and its impact on indoor temperature (and thus, air conditioning usage). This reasoning can also be partially observed by Fig. 6, which presents a histogram of the hour of day at which summertime (defined as July, August, September) peak electricity consumption occurs for each household. (In other words, the height of each bar represents the total number of households that have summertime peak electricity consumption at that time of day.) The hour of day corresponding to peak electricty consumption per household represents the most frequently occurring daily peak time over the summertime period. We observe in this study that the timing of most households' peak electricity use does not correspond to daily peak ambient temperature (usually during mid afternoon). By contrast, a large number of households have peak energy use in the late afternoon through early evening. Although we currently lack data to calculate how much of this peak energy use is driven by air conditioners, it is reasonable to assume that air conditioners are a major driver of evening electricity consumption since a large fraction of occupants are home from work during this period and might choose to cool their homes for occupant comfort.

Second, the choice of whether to use daily average or maximum temperature depends on the research questions under investigation. For example, most research analyses assessing the impacts of climate change on electricity utilize daily average temperature, since this indicator is what is estimated most commonly in global climate modeling studies [22,28,29,32,37–40]. On the other hand, daily maximum temperature is often used to predict future peak electricity demand, which is driven instantaneously by extreme heat during the day [28,33,46]. While it should be noted that total electricity usage is dependent on many factors, in this study, daily average temperature shows the best segmented linear relationship with electricity use relative to other temperature indicators (i.e. hourly, and daily minimum and maximum temperature). One of the driving reasons for this trend is likely due to nature of temperature fluctuations across differing climates, which can change the need for cooling throughout the day. For example, while a coastal home may experience similar daily average temperature (e.g. 30 °C) with an inland home in a dry desert region, the diurnal temperature range that each home experiences can be vastly different (e.g., coastal daily temperature range: 28–32 °C vs inland: 20–40 °C). In this example, the maximum daily (or minimum daily) temperature is vastly different in each region, even when the daily average temperature is the similar. While one might assume that total daily electricity consumption might scale with maximum temperature, the inland home would experience a great deal more nighttime cooling than the coastal home; this nighttime cooling might attenuate the need for some daytime air-conditioning use since it experiences pre-cooling. On the other hand, while the coastal home might not be subjected to extreme maximum temperatures, it also experiences less cooling relief during the evening in this example.

Third, setting a uniform, pre-defined base temperature as is done in the *CDD* calculation is not as good as computing household level stationary point temperatures. Using pre-defined base temperatures can lead to inaccuracies in regressions when occupant behaviors lead to the AC turning on at ambient temperatures below the threshold. This effect can be observed in Fig. 4(i) and (j), illustrated by data points with $CDD18C = 0$. Including these zero values affects the regression slope (i.e. sensitivity) relative to excluding the zeros (see Fig. 4(k) and (l)). In addition, using *CDD18C* as the indicator (with linear regression) leads to coefficients of determination that are smaller than when using daily average temperature (with segmented regression). Thus, using daily average temperatures with segmented regression may be best for studies that investigate sensitivities of daily electricity use (as opposed to peak energy use) rather than *CDD18C*.

5.3. Comparing computed electricity–temperature sensitivity to previous studies

Using the dataset described in this study, the computed electricity–temperature sensitivity $\bar{S}_{s=\text{household},t=\text{daily}}$ is 0.079 kW °C^{-1} using daily maximum temperature and 0.11 kW °C^{-1} using daily average temperature. However, previous studies commonly present electricity–temperature sensitivity in units of percentage change in electricity consumption per °C increase in ambient temperature (% °C^{-1}). Thus, to be comparable with sensitivity values from past studies, we also computed electricity–temperature sensitivity in units of percentage change in electricity consumption per °C increase in ambient temperature (% °C^{-1}). These sensitivities ($\bar{S}_{s=\text{household},t=\text{daily}}$) were 10.5% °C^{-1} using daily maximum temperature and 15.3% °C^{-1} using daily average temperature in this dataset. (All of these values are computed by calculating the sensitivities in percent units for each household and then averaging over all households.) Among previous studies surveyed, three report electricity–temperature sensitivity values computed using hourly or monthly average temperature data, presented as 6% °C^{-1} [26], 8.9% °C^{-1} [47], and 9-13% °C^{-1} [30], which are similar in magnitude to those computed in this study.

Calculating electricity–temperature sensitivity in units of kW °C^{-1} versus % °C^{-1} presents tradeoffs in terms of insights gained. Sensitivities in units of kW °C^{-1} will be highest for households with high cooling loads regardless of the magnitude of non-cooling loads, while reporting in units of % °C^{-1} is dependent on the magnitude of cooling loads versus non-cooling loads. Thus, a household with small non-cooling loads would have a higher percentage increase in cooling load per unit temperature rise relative to a household with high non-cooling loads, even if the cooling load increase in kW °C^{-1} are equal; yet, reporting the percentage of cooling load increase is insightful for understanding trends such as the relative increases in electricity costs for different socio-economic populations.

In addition to these considerations regarding selected units, several caveats of such comparisons in sensitivities between this and prior studies should be noted: 1) the sample size of this study is not statistically representative of electricity users in the state of California; and 2) electricity–temperature sensitivities can be driven by numerous factors, e.g. occupant behavior patterns, climate zones, housing characteristics, etc. Neither this study nor previous studies have revealed enough detailed information to explain these differences in sensitivities, but will be the focus of future research.

6. Conclusion

Despite a growing body of literature utilizing various types of electricity usage and temperature source data across a wide range of spatiotemporal resolutions, no research to our knowledge has focused on assessing the impacts of data resolution and choice of temperature metrics on computed functional relationships between electricity usage and ambient temperature. To address this, we use hourly energy use records from 1245 customers across California along with corresponding hourly ambient temperature data to investigate the dependence of spatiotemporal data resolution and temperature metrics on computed electricity–temperature sensitivities. Sensitivities are computed using a segmented linear regression model. We use this regression model with input data at various resolutions to emulate source data of spatial resolutions including household, city, county, and climate zone, and temporal resolutions including hourly and daily. In addition, we compare the impacts on computed electricity–temperature sensitivity of using hourly, daily minimum, daily mean, daily maximum, or cooling degree days as temperature indicators in the regression model.

Results indicate that the strongest relationships between electricity consumption and temperature, as indicated using the coefficients of determination, are computed when using data at daily temporal resolution (i.e. daily accumulated electricity consumption and daily average temperature), even when compared to those relationships computed using more resolved hourly electricity consumption and temperature data. This finding indicates that increasing the temporal resolution of electricity data to increments smaller than daily do not translate to higher regression model performance. By contrast, increasing the spatial resolution of electricity data improved the accuracy of computed electricity-temperature sensitivity (i.e., since ambient temperatures experienced by the house can be more accurately determined), and elucidated new trends masked by using spatially aggregated data as well.

The choice of temperature indicator can also impact the computed electricity–temperature sensitivities and stationary point temperature values. In this study, regression models utilizing daily average temperature offered higher coefficients of determination than those using daily minimum temperature, daily maximum temperature, or cooling degree days, thus showing the best segmented linear relationship with electricity usage. Moreover, we find that computing a unique household level stationary point temperature is superior to setting a uniform, pre-defined base temperature as is done in standard *CDD* calculations. Having access to household level data enables the calculation of a household specific base temperature (assuming that base temperature is effectively the equivalent of stationary point temperature), which can enhance the accuracy of computed sensitivities. In addition, having household level electricity data allows for determining more representative temperatures for that household, which can improve the accuracy of computed sensitivities especially in places like the Los Angeles Basin where temperature variations are significant. However, it should be noted that the choice of using daily average or maximum temperature depends on the research questions under investigation, and there are cases where daily average temperature might not be the indicator of choice.

To summarize, the take-away points of this study are:

- Sensitivities between residential electricity consumption and ambient temperature are best computed using daily data, as indicated using the coefficient of determination from a segmented linear regression model. Daily data led to improved regressions even when compared to hourly electricity consumption and temperature data. Daily data refers to daily accumulated electricity usage or daily average power consumption. The choice of whether to use daily average versus maximum temperature depends on the research question under investigation (see next bullet).
- Daily average temperature is the best choice for exploring general relationships between residential electricity consumption and ambient temperature. While use of daily average temperatures led to the highest coefficients of determination, use of daily maximum temperature can be more appropriate for investigating certain research questions (e.g. relationships between peak electricity use versus temperature).
- Having access to household level data can enhance the accuracy of computed sensitivities, and elucidate new trends (e.g., household-to-household variations in electricity–temperature sensitivity and stationary point temperature, which can be thought of as the ambient temperature at which the AC is turned on) masked by using spatially aggregated data.

For future research, a more statistically representative dataset of electricity consumption records is needed for better quantifying and understanding the relationship between residential electricity consumption and climatic parameters, which is essential for investigating effective energy conservation, peak energy management, and climate adaptation strategies.

Conflict of interest

None.

Acknowledgments

This work was funded in part by the Teh Fu “Dave” Yen Fellowship in Environmental Engineering awarded to Mo Chen at the University of Southern California; This work was also funded in part by the National Science Foundation under grants CBET-1512429, CBET-1752522, EAGER-1632945; we also acknowledge the valuable and helpful suggestions and comments offered by all reviewers for this work.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.enbuild.2018.09.012.

References

[1] EIA, U.S. Energy Information Administration (EIA), 2015. http://www.eia.gov/.
[2] U.S. Energy Information Administration, Electricity data browser - retail sales of electricity, 2016. http://www.eia.gov/electricity/data/browser/.
[3] U.S. Energy Information Administration, Annual energy outlook 2016, 2016. Washington D.C. http://www.eia.gov/forecasts/aeo/.
[4] U.S. Energy Information Administration, How is Electricity Used in U.S. Homes? - FAQ, U.S. Energy Information Administration (EIA), 2016 https://www.eia.gov/tools/faqs/faq.cfm?id=96&t=3 (accessed January 1, 2016).
[5] B.E. Psiloglou, C. Giannakopoulos, S. Majithia, M. Petrakis, Factors affecting electricity demand in Athens, Greece and London, UK: A comparative assessment, Energy 34 (2009) 1855–1863, doi:10.1016/j.energy.2009.07.033.
[6] Energy Information Administration, About the Residential Energy Consumption Survey (RECS): fuels used and end uses in U.S. homes by housing unit type, 2015, 2015 (accessed August 5, 2017). https://www.eia.gov/consumption/residential/data/2015/hc/php/hc7.8.php .
[7] L. Pérez-Lombard, J. Ortiz, C. Pout, A review on buildings energy consumption information, Energy Build. 40 (2008) 394–398, doi:10.1016/j.enbuild.2007.03.007.
[8] A. Kavousian, R. Rajagopal, M. Fischer, Determinants of residential electricity consumption: using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior, Energy 55 (2013) 184–194, doi:10.1016/j.energy.2013.03.086.
[9] J.C. Lam, Climatic and economic influences on residential electricity consumption, Energy Convers. Manage. 39 (1998) 623–629, doi:10.1016/S0196-8904(97)10008-5.
[10] M. Beccali, M. Cellura, V. Lo Brano, A. Marvuglia, Short-term prediction of household electricity consumption: assessing weather sensitivity in a Mediterranean area, Renew. Sustain. Energy Rev. 12 (2008) 2040–2065, doi:10.1016/j.rser.2007.04.010.
[11] M. Ranjan, V.K. Jain, Modelling of electrical energy consumption in Delhi, Energy 24 (1999) 351–361, doi:10.1016/S0360-5442(98)00087-5.
[12] D.J. Sailor, Relating residential and commercial sector electricity loads to climate - evaluating state level sensitivities and vulnerabilities, Energy 26 (2001) 645–657, doi:10.1016/S0360-5442(01)00023-8.
[13] Y. Hirano, T. Fujita, Evaluation of the impact of the urban heat island on residential and commercial energy consumption in Tokyo, Energy 37 (2012) 371–383, doi:10.1016/j.energy.2011.11.018.
[14] J.C. Lam, H.L. Tang, D.H.W. Li, Seasonal variations in residential and commercial sector electricity consumption in Hong Kong, Energy 33 (2008) 513–523, doi:10.1016/j.energy.2007.10.002.
[15] M.A. Badri, Analysis of demand for electricity in the United States, Energy 17 (1992) 725–733, doi:10.1016/0360-5442(92)90080-J.
[16] T.F. Sanquist, H. Orr, B. Shui, A.C. Bittner, Lifestyle factors in U.S. residential electricity consumption, Energy Policy 42 (2012) 354–364, doi:10.1016/j.enpol.2011.11.092.
[17] A. De Almeida, P. Fonseca, B. Schlomann, N. Feilberg, Characterization of the household electricity consumption in the EU, potential energy savings and specific policy recommendations, Energy Build. 43 (2011) 1884–1894, doi:10.1016/j.enbuild.2011.03.027.
[18] C. Bartusch, M. Odlare, F. Wallin, L. Wester, Exploring variance in residential electricity consumption: Household features and building properties, Appl. Energy 92 (2012) 637–643, doi:10.1016/j.apenergy.2011.04.034.
[19] S. Kelly, M. Shipworth, D. Shipworth, M. Gentry, A. Wright, M. Pollitt, D. Crawford-Brown, K. Lomas, Predicting the diversity of internal temperatures from the English residential sector using panel methods, Appl. Energy 102 (2013) 601–621, doi:10.1016/j.apenergy.2012.08.015.
[20] Y.Y. Yan, Climate and residential electricity consumption in Hong Kong, Energy 23 (1998) 17–20, doi:10.1016/S0360-5442(97)00053-4.
[21] M. Bessec, J. Fouquau, The non-linear link between electricity consumption and temperature in Europe: a threshold panel approach, Energy Econ. 30 (2008) 2705–2721, doi:10.1016/j.eneco.2008.02.003.
[22] J. Huang, K.R. Gurney, Impact of climate change on U.S. building energy demand: sensitivity to spatiotemporal scales, balance point temperature, and population distribution, Clim. Change (2016) 171–185, doi:10.1007/s10584-016-1681-6.

[23] Energy Information Agency (EIA), Homes Show Greatest Seasonal Variation in Electricity Use - Today in Energy, U.S. Energy Information Administration (EIA), 2013 https://www.eia.gov/todayinenergy/detail.php?id=10211 (accessed August 7, 2017).

[24] S. Horowitz, B. Mauch, F. Sowell, Forecasting residential air conditioning loads, Appl. Energy 132 (2014) 47–55, doi:10.1016/j.apenergy.2014.06.029.

[25] T.K. Doshi, P. Fellow, A. Rohatgi, N. Bin Zahur, E. Analyst, Y.K. Hung, Impact of climate change on electricity demand of Singapore, 2012. https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=SERC2013&paper_id=405.

[26] M. Pomerantz, P.J. Rosado, R. Levinson, A simple tool for estimating city-wide annual electrical energy savings from cooler surfaces, Urban Clim. 14 (2015) 315–325, doi:10.1016/j.uclim.2015.05.007.

[27] S.J. Parkpoom, G.P. Harrison, Analyzing the impact of climate change on future electricity demand in Thailand, IEEE Trans. Power Syst. 23 (2008) 1441–1448, doi:10.1109/TPWRS.2008.922254.

[28] G. Franco, A.H. Sanstad, Climate change and electricity demand in California, Clim. Change 87 (2008) 139–151, doi:10.1007/s10584-007-9364-y.

[29] D.B. Belzer, M.J. Scott, R.D. Sands, Climate change impacts on U.S. commercial building energy consumption: an analysis using sample survey data, Energy Sources 18 (1996) 177–201, doi:10.1080/00908319608908758.

[30] A. Aroonruengsawat, M. Auffhammer, Impacts of climate change on residential electricity consumption evidence from billing data, in: Econ. Clim. Chang. Adapt. Past Present, University of Chicago Press, 2011, pp. 311–342.

[31] S. Mirasgedis, Y. Sarafidis, E. Georgopoulou, D.P. Lalas, M. Moschovits, F. Karagiannis, D. Papakonstantinou, Models for mid-term electricity demand forecasting incorporating weather influences, Energy 31 (2006) 208–227, doi:10.1016/j.energy.2005.02.016.

[32] A. Bigano, F. Bosello, G. Marano, Energy demand and temperature: a dynamic panel analysis, Fond. Eni Enrico Mattei. (2006) FEEM Working Paper No. 112.06., doi:10.2139/ssrn.928798.

[33] J.A. Sathaye, L.L. Dale, P.H. Larsen, G.A. Fitts, K. Koy, S.M. Lewis, Estimating impacts of warming temperatures on California's electricity system, Glob. Environ. Change 23 (2013) 499–511, doi:10.1016/j.gloenvcha.2012.12.005.

[34] D.J. Sailor, J.R. Muñoz, Sensitivity of electricity and natural gas consumption to climate in the U.S.A. - methodology and results for eight states, Energy 22 (1997) 987–998, doi:10.1016/S0360-5442(97)00034-0.

[35] A. Pardo, V. Meneu, E. Valor, Temperature and seasonality influences on Spanish electricity load, Energy Econ. 24 (2002) 55–70, doi:10.1016/S0140-9883(01)00082-2.

[36] M. Auffhammer, P. Baylis, C.H. Hausman, Climate change is projected to have severe impacts on the frequency and intensity of peak electricity demand across the United States, Proc. Natl. Acad. Sci. (2017) 114. 10.1073/pnas.1613193114/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1613193114.

[37] M. Hekkenberg, H.C. Moll, A.J.M.S. Uiterkamp, Dynamic temperature dependence patterns in future energy demand models in the context of climate change, Energy 34 (2009) 1797–1806, doi:10.1016/j.energy.2009.07.037.

[38] M. Ruth, A.C. Lin, Regional energy demand and adaptations to climate change: methodology and application to the state of Maryland, USA, Energy Policy 34 (2006) 2820–2833, doi:10.1016/j.enpol.2005.04.016.

[39] M. Auffhammer, A. Aroonruengsawat, Simulating the impacts of climate change, prices and population on California's residential electricity consumption, Clim. Change 109 (2011) 191–210, doi:10.1007/s10584-011-0299-y.

[40] B. Lebassi, J.E. Gonzaĺez, D. Fabris, R. Bornstein, Impacts of climate change in degree days and energy demand in Coastal California, J. Sol. Energy Eng. 132 (2010) 031005-031005-9, doi:10.1115/1.4001564.

[41] L.G. Swan, V.I. Ugursal, Modeling of end-use energy consumption in the residential sector : a review of modeling techniques, Renew. Sustain. Energy Rev. 13 (2009) 1819–1835, doi:10.1016/j.rser.2008.09.033.

[42] M. Segal, H. Shafir, M. Mandel, P. Alpert, Y. Balmor, Climatic-related evaluations of the summer peak-hours' electric load in Israel, J. Appl. Meteorol. 31 (1992) 1492–1498, doi:10.1175/1520-0450(1992)031⟨1942:CREOTS⟩2.0.CO;2.

[43] M. Hekkenberg, R.M.J. Benders, H.C. Moll, A.J.M. Schoot Uiterkamp, Indications for a changing electricity demand pattern: the temperature dependence of electricity demand in the Netherlands, Energy Policy 37 (2009) 1542–1551, doi:10.1016/j.enpol.2008.12.030.

[44] K. Wangpattarapong, S. Maneewan, N. Ketjoy, W. Rakwichian, The impacts of climatic and economic factors on residential electricity consumption of Bangkok Metropolis, Energy Build 40 (2008) 1419–1425, doi:10.1016/j.enbuild.2008.01.006.

[45] C. Giannakopoulos, B.E. Psiloglou, Trends in energy load demand for Athens, Greece: weather and non-weather related factors, Clim. Res. 31 (2006) 97–108, doi:10.3354/cr031097.

[46] A.B. Colombo, D. Etkin, B.W. Karney, Climate variability and the frequency of extreme temperature events for nine sites across canada : implications for power usage, J. Clim. (1999) 2490–2503, doi:10.1175/1520-0442(1999)012⟨2490:CVATFO⟩2.0.CO;2.

[47] W.Y. Fung, K.S. Lam, W.T. Hung, S.W. Pang, Y.L. Lee, Impact of urban temperature on energy consumption of Hong Kong, Energy 31 (2006) 2287–2301, doi:10.1016/j.energy.2005.12.009.

[48] D.J. Sailor, C. Vasireddy, Correcting aggregate energy consumption data to account for variability in local weather, Environ. Model. Softw. 21 (2006) 733–738, doi:10.1016/j.envsoft.2005.08.001.

[49] T. Ihara, Y. Genchi, T. Sato, K. Yamaguchi, Y. Endo, City-block-scale sensitivity of electricity consumption to air temperature and air humidity in business districts of Tokyo, Japan, Energy 33 (2008) 1634–1645, doi:10.1016/j.energy.2008.06.005.

[50] H. Akbari, Cooling our Communities: A Guidebook On Tree Planting And Light–Colored Surfacing, 1992. ISBN: 0-16-036034-X.

[51] California Energy Commission, California building climate zone areas, 2015 (accessed June 1, 2017). http://www.energy.ca.gov/maps/renewable/building_climate_zones.html .

[52] California Department of Water Resources, The California Irrigation Management Information System, 2017 http://www.cimis.water.ca.gov/WSNReportCriteria.aspx (accessed June 1, 2017).

[53] H.S. Hippert, D.W. Bunn, R.C. Souza, Large neural networks for electricity load forecasting: are they overfitted? Int. J. Forecast. 21 (2005) 425–434, doi:10.1016/j.ijforecast.2004.12.004.

[54] V.M.R. Muggeo, Regression models with breakpoints/changepoints estimation, RDoc (2015).

[55] Energy Information Agency, Household Energy Use in California, 2009 https://www.eia.gov/consumption/residential/reports/2009/state_briefs/pdf/ca.pdf (accessed August 9, 2017).

[56] M. Beenstock, E. Goldin, D. Nabot, The demand for electricity in Israel, Energy Econ. 21 (1999) 168–183, doi:10.1016/S0140-9883(98)00005-X.

[57] G. Ban-Weiss, C. Wray, W. Delp, P. Ly, H. Akbari, R. Levinson, Electricity production and cooling energy savings from installation of a building-integrated photovoltaic roof on an office building, Energy Build. 56 (2013) 210–220, doi:10.1016/j.enbuild.2012.06.032.