

Predicting city-scale daily electricity consumption using data-driven models

Zhe Wang, Tianzhen Hong*, Han Li, Mary Ann Piette

Building Technology and Urban Systems Division, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA



ARTICLE INFO

Keywords:

City-scale electricity usage
Decomposed time-series modeling
Gradient boosting trees
Temperature-sensitive energy demand
Machine learning prediction

ABSTRACT

Accurate electricity demand forecasts that account for impacts of extreme weather events are needed to inform electric grid operation and utility resource planning, as well as to enhance energy security and grid resilience. Three common data-driven models are used to predict city-scale daily electricity usage: linear regression models, machine learning models for time series data, and machine learning models for tabular data. In this study, we developed and compared seven data-driven models: (1) five-parameter change-point model, (2) Heating/Cooling Degree Hour model, (3) time series decomposed model implemented by Facebook Prophet, (4) Gradient Boosting Machine implemented by Microsoft lightGBM, and (5) three widely-used machine learning models (Random Forest, Support Vector Machine, Neural Network). Seven models are applied to the city-scale electricity usage data for three metropolitan areas in the United States: Sacramento, Los Angeles, and New York. Results show seven models can predict the metropolitan area's daily electricity use, with a coefficient of variation of the root mean square error (CVRMSE) less than 10%. The lightGBM provides the most accurate results, with CVRMSE on the test dataset of 6.5% for Los Angeles, 4.6% for Sacramento, and 4.1% for the New York metropolitan area. These models are further applied to explore how extreme weather events (e.g., heat waves) and unexpected public health events (e.g., COVID-19 pandemic) influence each city's electricity demand. Results show weather-sensitive component accounts for 30%–50% of the total daily electricity usage. Every degree Celsius ambient temperature increase in summer leads to about 5% (4.7% in Los Angeles, 6.2% in Sacramento, and 5.1% in New York) more daily electricity usage compared with the base load in the three metropolitan areas. The COVID-19 pandemic reduced city-scale electricity demand: compared with the pre-pandemic same months in 2019, daily electricity usage during the 2020 pandemic decreased by 10% in April and started to rebound in summer.

1. Introduction

1.1. Background

City-scale electricity demand prediction can be used to assist in power generation resource planning, energy efficiency program evaluation, greenhouse gas emissions tracking, grid infrastructure analysis, and analysis of reserve requirements. Therefore, understanding building energy use at the city scale is a critical component of advancing urban sustainability, carbon reduction, and energy efficiency across the globe [1]. City-scale electricity usage is temperature sensitive. Ambient temperature, along with factors such as population and income, is a key driver of city-scale energy usage [2], because heating and cooling buildings is a major energy use in cities [3], and that is highly dependent on the outdoor air temperature. Temperature-sensitive city-scale electricity consumption analysis is becoming an important topic under the context of climate change, which is leading to more frequent, more

severe, and longer extreme weather events such as heat waves. To develop effective climate change solutions, researchers, energy planners, and policy makers need to pay attention to climate change adaptation [4]. Similarly, these stakeholders need to understand how the electricity generation and transmission infrastructure should be better prepared for high demand events, to enhance energy security and resilience as part of climate change adaptation.

Hou et al. (2014) studied how the increasing ambient temperature influenced the electricity consumption in Shanghai. They argued that the projected temperature increase implies an increasing electricity demand in summer and a decreasing demand in winter if the current energy consumption pattern does not change [5]. Similarly, it was estimated in California that the atmospheric warming and the associated peak demand increases would necessitate up to 38% of additional peak generation capacity and up to 31% additional transmission capacity by 2099 [6]. In August 2020, a heat wave in California led to a power supply shortage due to a surge of air-conditioning use, and California's

* Corresponding author.

E-mail address: thong@lbl.gov (T. Hong).

residents experienced rotating power outages. When policy makers reflected on this disruptive event, the first going-forward action mentioned in the Response Letter was to refine the electricity demand forecast by accounting for climate change, capturing extreme weather events and associated load impacts [7].

In addition to the ambient weather condition, city-scale electricity usage can also be influenced by other factors such as unexpected public health events. Studies conducted in Brazil found the COVID-19 pandemic has influenced Brazilian's electricity consumption patterns, and as a result, the electricity consumption in Brazil was reduced by 7%–20%, depending on the local economic structure: industry-dominated area was less affected [8]. Another study conducted in Europe found the strictness and intensiveness of lockdown measures influenced the society's electricity consumption. For countries with severe restrictions such as Spain and Italy, the electricity usage profiles during the pandemic are similar to the pre-pandemic weekend profiles for the same period in 2019; while for countries with less restrictive measures such as Sweden, the decrease in power consumption was lower [9]. Electricity consumption can be used as a real-time indicator of the economic effects of the lockdown. It was found that the overall electricity use decrease in Switzerland was 4.6%, while a reduction of 14.3% was observed in the Canton of Ticino where stricter curtailment measures were implemented on top of federal regulations [10].

There are two general approaches to model city-level energy usage: top-down methods and bottom-up methods [11]. Top-down models treat the city-level energy consumption at the macro scale, ignoring details of individual end uses [12]. The top-down approach is usually employed to extract the relationships between energy usage and macroeconomic and demographic factors [13]. The bottom-up approach develops simulation models for individual units and then aggregates those units to calculate the macro level energy usage [14]. As there can be millions of end users at the city scale; the simulation unit is not necessary to be individual end user, instead, it can be a cluster/group of end users with similar characteristics or patterns.

It is difficult, if not impossible, to develop a physics-based model at the macro level (e.g., because it requires a large amount of input data with high uncertainty to establish a physics-based model), so top-down modeling usually uses data-driven approaches. Conversely, the bottom-up approach has the flexibility to develop either a physics-based model or a data-driven model. Wang et al. (2015) developed a physics-based bottom-up model to estimate the state level heating energy consumption in China [15]. Kontokosta and Tull (2017) developed a data-driven bottom-up model to predict the city-scale building energy consumption in New York [1]. The bottom-up approach allows city-scale building retrofit analysis covering individual buildings with their varying characteristics and baseline performance—an approach that cannot be conducted using top-down statistical-based methods [16]. It is difficult to say whether the top-down approach or the bottom-up approach is better, as both methods have strengths and weaknesses [17]. Which approach is more suitable depends on the application and objective of the modeling (i.e., fit-for-purpose).

1.2. Objectives

City-scale electricity prediction has wide applications, and we found many studies on this topic. However, a major research gap we identified after reviewing existing studies is: different approaches have been proposed but none of the studies provide open source data, model or code for public inspection or reuse. Therefore, there is a lack of summary and comparison between those methods, especially between the conventional linear regression approach and the emerging machine learning algorithms.

In this study, we applied the top-down data-driven approach to predict the daily electricity consumption at the city scale. The contributions of this paper are twofold:

- First, we explored three different approaches to predict city-level¹ daily energy consumption: linear regression models, machine learning models for time-series data, and machine learning model for non-time-series data. We tested and compared these approaches with electricity data from three U.S. metropolitan areas: Los Angeles, Sacramento, and New York. We open sourced all the data and all the codes at Github, <https://github.com/LBNL-ETA/City-Scale-Electricity-Use-Prediction>, which enables other researchers to test and compare their models against ours.
- Second, we used the models to study both the impacts of extreme heat waves and the unexpected public health event of the 2020 COVID-19 pandemic on city-scale daily electricity demand. These results provide quantitative daily electricity usage and peak demand prediction to inform energy planning and policy making for utilities and state or local government.

The remainder of this paper is organized as follows. Section 2 introduces the three modeling approaches and how our work is built upon and different from existing studies. The data used in this study and model prediction results are presented in Section 3. Section 4 first compares the seven models in terms of accuracy and complexity to implement; and then applies the best performing models to explore how heat waves and COVID-19 influenced the city-level energy consumption. The implication and limitations of this study are discussed in Section 5 before we conclude in Section 6.

2. Methodology

In this section, we review the existing city-level electricity usage models and then introduce the models we developed. The workflow of this study is shown in Fig. 1. We developed seven data-driven models to predict the daily electricity use at city scale. These models belong to three different modeling approaches: linear model, machine learning model for time-series data, and machine learning model for tabular data. Due to the simplicity and interoperability, we first developed two linear models: ASHRAE's Five-Parameter piecewise linear regression model and Heating Cooling Degree Hour model. In terms of the Machine Learning model, we first developed three conventional models as baselines: Random Forest (RF), Support Vector Machine (SVM), and Artificial Neural Network (NN). Additionally, two machine learning models – Generalized Additive Model (GAM) and Gradient Boosting Machine (GBM), which have rarely been used in city-scale building energy prediction – were developed and compared. GAM and GBM models the electricity consumption data in distinct way. GAM models the electricity usage as time series data while GBM models the electricity usage as tabular data. The key difference between time series modeling and tabular data modeling is how the temporal information is encoded. Since electricity usage has clear weekly and yearly cycles, it is a natural idea to use time series modeling to predict electricity usage. The time series model treats the electricity use as time series data, representing temporal information using an evolving index. The three baseline machine learning models (RF, SVM, NN) process the electricity usage as tabular data. In this case, the weekly and yearly cycles need to be captured by adding extra features, such as the day of week and month of year. Tabular data model adopts a different approach to encode temporal information, i.e., representing the temporal information using extra features (such as day of year, hour of day). In other words, in time series models, the sequence of data cannot be changed because the data sequence contains temporal information which is useful for prediction. However, in tabular data models, the data sequence can be shuffled because the temporal information was captured by the extra features.

The time-series model and the tabular data model are compared with the baseline – the linear models and baseline machine learning models.

¹ City-level in this study refers to the electricity consumed by the city and its surrounding areas.

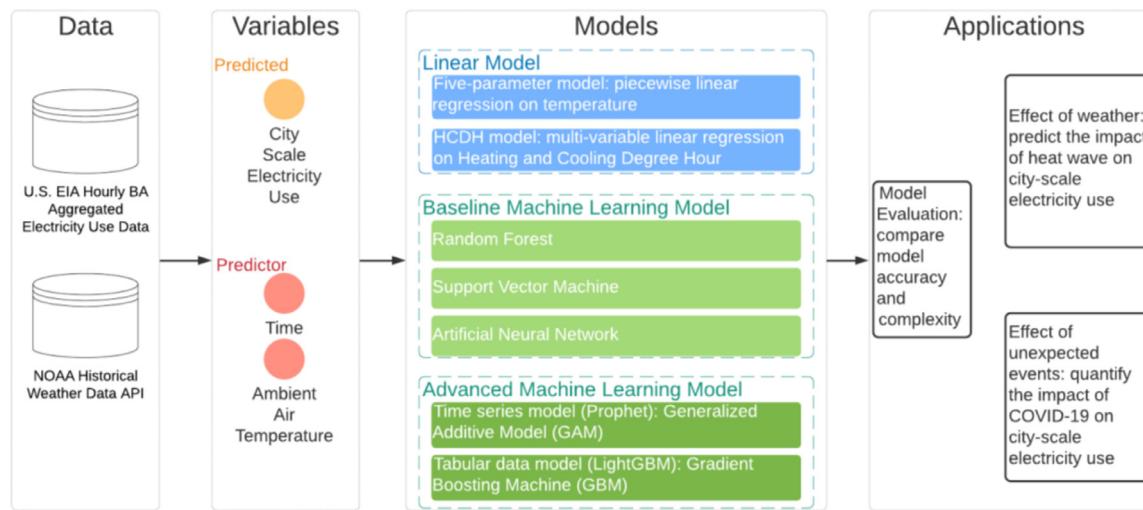


Fig. 1. Overall workflow.

The best performing models were selected to answer two practical questions: 1) how the city-scale daily electricity use would be influenced by the ambient air temperature, especially in a heat wave event; 2) how the city-scale daily electricity use would be influenced by unexpected public health events, in this case the COVID-19 pandemic.

In this study, we are predicting city-scale daily electricity use which includes electricity use for buildings, transportation (e.g., electric vehicles, public electric buses/trains), industry, and other public services inside the city and the neighboring rural areas.

2.1. Linear models

Linear models use a linear relationship to regress the observation and the independent variables. The strength of linear models lies in their simplicity and interoperability. Therefore, linear models are frequently used in city-level energy modeling. For instance, Lindsey et al. (2011) developed a linear model to predict the city-level transportation energy usage and greenhouse gas emission in Chicago [18]. Kuusela et al. (2015) developed a multi-variable linear regression model to predict the neighborhood scale energy consumption [19]. Actually, even for a relatively complicated energy system, such as a large-scale ground source heat pump, linear models can deliver decent prediction performance [20].

A major reason why linear models are widely used is their interoperability. The regressed coefficients (slopes and intercept) have clear implications that modelers can use to validate and explain their models. Linear models might be less accurate, as they fail to capture non-linear relationships that are common in the real world [21]. However, the simplicity and understandability of linear models makes them a good choice as the baseline model for benchmarking against more complicated non-linear models.

In this study, we developed two linear models to predict the temperature-sensitive electricity usage in city-scale. The first linear model is ASHRAE's change-point model, which was originally proposed by ASHRAE in the 1990s [22]. As shown in Fig. 2 and Eqn 1, the change point model uses five parameters (β_{base} , T_h , T_c , β_h , β_c) to characterize the relation between energy usage and ambient temperature. β_{base} is the base load. When the outdoor temperature is in the range of $[T_h, T_c]$, the energy usage is the lowest, and that is referred to as the base load. When the outdoor temperature is lower than the heating change point T_h , city-level energy usage increases as heating demand increases in response to the temperature decreasing. Conversely, when the outdoor temperature is higher than the cooling change point T_c , city-level energy usage also increases as cooling demand increases with higher tempera-

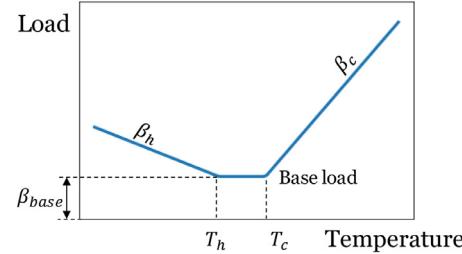


Fig. 2. ASHRAE's change point model.

ture. The slope on the cooling (β_c) and heating (β_h) sides characterize how sensitive the city-level load is related to the temperature change. Since we are interested in electricity consumption only, the β_h would be smaller than β_c , because many buildings use natural gas for heating while the majority of air conditioned buildings use electricity for cooling. ASHRAE's five-parameter change point model has been widely used to predict building-level energy consumption [23,24], and to benchmark building energy performance [25]. In this study, we applied the five-parameter change point (5-p) model to predict city-level energy usage.

$$load(T) = \begin{cases} \beta_{base} + \beta_h \times (T_h - T), & \text{if } T < T_h \\ \beta_{base}, & \text{if } T_h < T < T_c \\ \beta_{base} + \beta_c \times (T - T_c), & \text{if } T_c < T \end{cases} \quad (1)$$

The second linear model we used is the Heating/Cooling Degree Hour (HCDH) model. Heating/Cooling Degree Hour method is one of the most well-known methods used in the heating, ventilating and air-conditioning (HVAC) industry to estimate heating and cooling energy requirements [26]. Because of its significance in this field, the heating and cooling degree day was used to determine the U.S. climate zone, and has been widely used as a proxy variable to quantify the influence of climate change on electricity demand [27]. As shown in Eqn 2, heating degree hour (HDH) and cooling degree hour (CDH) are calculated as the accumulative sum of the difference between the ambient temperature and the heating (T_{b_h}) cooling (T_{b_c}) base temperatures. Once the outdoor temperature is below T_{b_h} , heating is likely to be triggered. The cumulative sum of the difference between the outdoor and the base temperature $T_{b_h} - T_i$ is a good indicator of how much heating is needed. Heating and cooling degree hours are widely used to estimate building energy demand [28], to determine the building thermal insulation [29], and for other purposes. In this study, we regressed the daily city-level

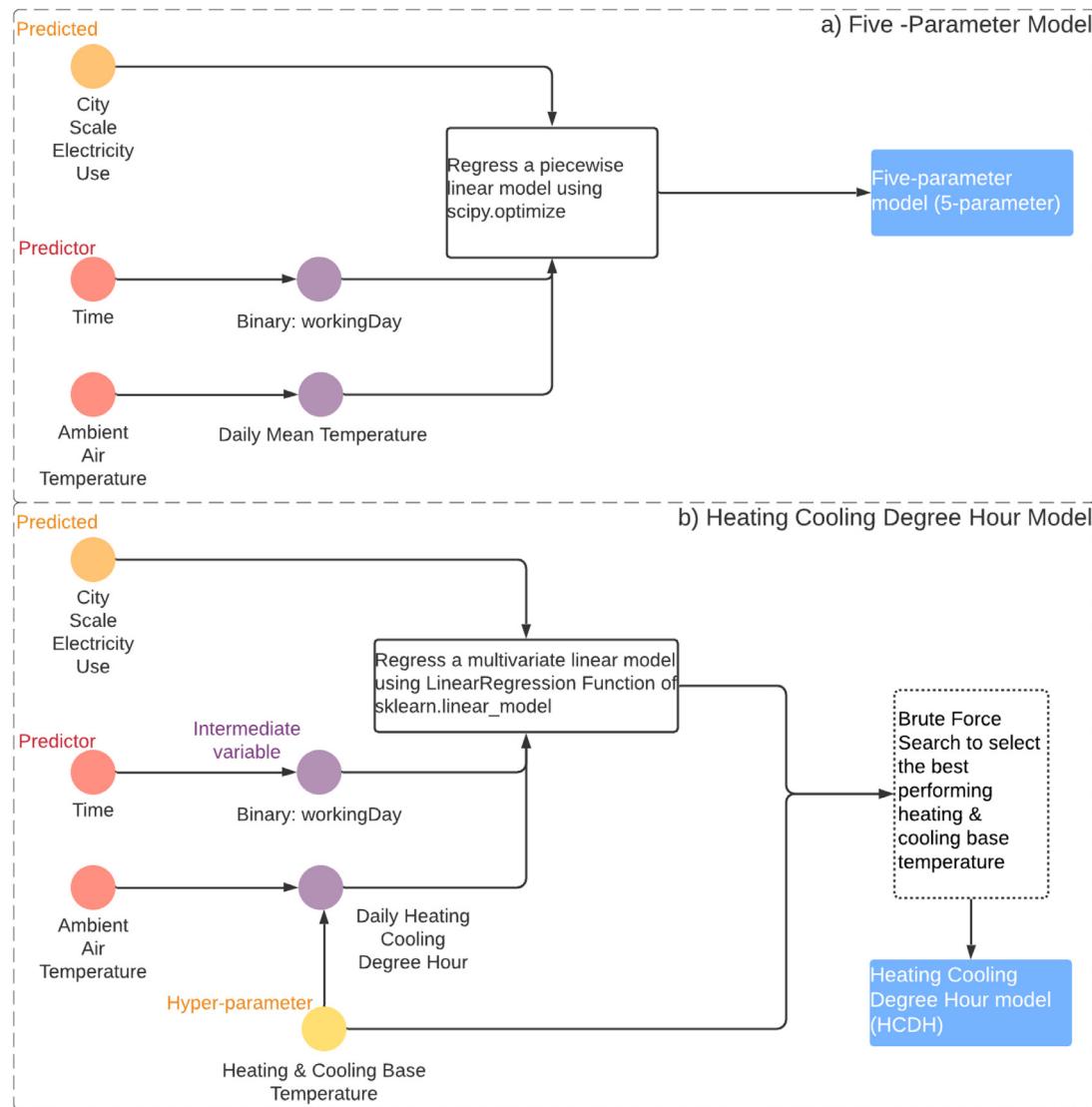


Fig. 3. Workflow to develop the linear models: top for Five-Parameter Change Point model, bottom for Heating and Cooling Degree Hour model.

energy usage as a linear function of the HDH and CDH.

$$HDH = \sum_{i=1}^{24} \max(0, (Tb_h - T_i))$$

$$CDH = \sum_{i=1}^{24} \max(0, (T_i - Tb_c))$$

$$load(T) = \beta_0 + \beta_1 \times HDH + \beta_2 \times CDH \quad (2)$$

A common problem of developing either the five-parameter or HCDH model is the need to carefully select the change temperature (T_h , T_c in the five-parameter model) and the base temperature (Tb_h , Tb_c in the HCDH model) [30]. In this study, we selected those temperatures based on which sets of change or base temperature could deliver the most accurate linear model. We used the `scipy.optimize.curve_fit` function [31] to determine the best change temperatures for the five-parameter model, and a brute force search to select the best base temperatures for the HCDH model. The workflow, input variables, intermediate variables, packages and algorithms used to develop the two linear models are illustrated in Fig. 3.

2.2. Machine learning model for time-series data

The Autoregressive Integrated Moving Average (ARIMA) is the oldest time-series modeling technique [32]. ARIMA predicts a time-series variable y_t with its own lagged values (y_{t-1}, y_{t-2}, \dots) and previous prediction error ($\varepsilon_{t-1}, \varepsilon_{t-2}, \dots$), where $\varepsilon_t = \hat{y}_t - y_t$, and \hat{y}_t is the prediction of the true value y_t . ARIMA has been widely used in our field; for instance, to predict the natural gas demand in Turkey [33] and the electricity demand in Lebanon [34]. Pappas et al. (2008) used Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) to decide the order of the ARIMA model and to validate electricity demand load forecasting models [35]. ARIMA also has been applied together with other techniques such as wavelet transform to enhance its prediction accuracy [36]. Another mainstream type of machine learning model for time series data is the Recurrent Neural Network (RNN) and its variation—Long Short Term Memory (LSTM). RNN and LSTM take a different approach to encode the time dependency: inputting a state from previous time step s_{t-1} to the neural network to predict y_t . The state s_{t-1} is a function of y_{t-1} , approximated by another neural network. With the recent advancement in deep learning algorithms and computational power, a neural network based approach has been used extensively to predict energy usage. Suganthi and Samuel's literature review paper found more than 40 papers

using neural network based approach for short, medium, and long term load forecasts [37]. In recent publications on this topic, Rahman et al. (2019) used RNN to predict the electricity consumption for commercial and residential buildings [38] and Wang et al. (2019) used LSTM to predict district-level energy demand [39]. There are some existing studies to compare the two time-series data modeling approaches—ARIMA and RNN/LSTM. For instance, Wang et al. (2019) compared ARIMA and LSTM in terms of building load forecast and found LSTM outperformed ARIMA, as LSTM is more capable of capturing non-linear relations between time series data and exogenous variables [40].

In this paper, we explore a new time-series modeling approach to predict city-level electricity usage—ARIMA and RNN. We used a decomposed time series model with three major components: exogenous variables, trend, and seasonality (weekly and yearly), as shown in Eqn 3. $f(Temp_t)$ is a linear function of heating and cooling degree hours; $g(t)$ models the trend function, which models non-periodic changes in the value of the time series; and $s(t)$ models the periodic changes, yearly and weekly seasonality in this case. We used Facebook's open-source software Prophet [41] to implement this decomposed time series model. For more details of Prophet's implementation, please refer to the paper [42].

$$load_t = f(Temp_t) + g(t) + s(t) + \epsilon_t \quad (3)$$

We selected the decomposed time-series model for two reasons. First, it has been applied to predict the temperature [43], financial markets [44], and COVID-19 daily cases in Bangladesh [45]. Second, the decomposed model can decouple the effects of different factors (such as weather-related temperature-dependent load and seasonal time-dependent periodical load), and accordingly provide us a unique opportunity to observe the “pure” effect of an unexpected public health event on city-level demand. To the best of the authors' knowledge, this is the first time it was applied to predict city-level energy consumption.

The strength of this decomposed time series model is that it can separate the variation of time series data into different components, and each component has clear implications. The temperature-sensitive load $f(Temp_t)$ is usually related to HVAC use, which accordingly is a function of temperature. The periodic load $s(t)$ captures the load variation as a function of time, for instance the influence of holiday season on city-level electricity usage. The non-periodic load $g(t)$ reflects the remaining variation of load, which could be due to short term reasons (such as the COVID-19 pandemic) or long term trends (such as the improving building thermal properties and equipment energy efficiency). The decomposition results could inform the amplitude of each component and identify the dominant driving factors.

The workflow to develop a time-series decomposed model is presented in Fig. 4. It is worth pointing out that the intermediate variable used to develop the time series decomposed model is the daily heating and cooling degree hour, rather than the daily mean temperature. Because the $f(Temp_t)$ term in the decomposed model is a monotonous function, to be more specific, a linear function. As the relation between the city-scale daily electricity use and the ambient mean temperature is U-shaped (i.e., high electricity use when ambient temperature is either low or high). The monotonous function cannot capture the U-shape relation. However, the relation between electricity-use and heating cooling degree hour is monotonous. Therefore, we use daily heating cooling degree hour as the regressor to the model.

2.3. Machine learning model for tabular data (non-time-series data)

City-level electricity consumption can also be modeled as tabular data. To capture the timing and periodic behavior of energy usage, new features need to be added. For instance, to encode the weekly and yearly cycles, two new features—day of week, and month of year—need to be added as input variables.

There are many studies that model time series energy usage data as tabular data. Machine learning algorithms for tabular data can be classified into three major categories: neural network based, decision tree based, and others. Neural network (NN) based approaches use neural networks to solve regression or classification tasks. These techniques are also known as artificial neural network (ANN), feedforward neural network (FNN), or multi-layer perception (MLP) in different studies. The strengths of an NN-based approach include that it can be parallelized easily, and can be used to solve many different tasks. Fernández et al. (2011) applied an NN-based approach to predict building load [46]. Decision tree based approaches include Classification and Regression Tree (CART), Random Forest (RF), and Gradient Boosting Machine (GBM). RF and GBM are ensemble learning techniques, which ensemble multiple decision trees to make predictions. Combining multiple trees usually outperforms a single tree in terms of model accuracy and robustness. Tso and Yau (2007) applied CART to predict building energy usage in Hong Kong [47]. Roth et al. (2019) developed RF and GBM models to predict building energy consumption in New York City [48]. In addition to the NN-based and decision tree-based approaches, other algorithms for tabular data also are being used for energy usage modeling. For instance, Li et al. (2017) used the Support Vector Machine (SVM) approach to predict community-level renewable generation [49]. Al-Qahtani and Crone (2013) applied k-Nearest Neighbors to predict electricity demand in the United Kingdom [50]. Fonseca and Schlueter (2015) applied k-means to predict district-level building energy usage in Zurich [51]. Additionally, Kontokosta and Tull (2017) found the best performing machine learning algorithm might depend on the geographical resolution: SVM provides the most accurate prediction at the building level while Linear regression model outperforms other methods at the zip code-level [52].

Four tabular data modeling algorithms are selected in this study: RF, SVM, NN, and GBM. RF, SVM, and NN are selected as baseline algorithms and GBM is introduced in greater detail here. In the recently organized building load prediction competition (the ASHRAE Great Energy Predictor III competition), all the top six teams applied GBM in their final predictors [53]. Additionally, existing studies have confirmed that GBM outperforms many other machine learning algorithms. For instance, [54] found GBM outperforms Ridge regression, Lasso regression, Elastic Net, Support Vector Machine, Random Forest, vanilla Deep Neural Network, and Long Short Term Memory in predicting building loads. In this study, we selected GBM as the state-of-art algorithm and as the representative of tabular data modeling approach.

The workflow to develop a tabular data Gradient Boosting Machine model is illustrated in Fig. 5. The first step is to use the time index to generate intermediate variables to encode temporal information. There are different implementation packages of GBM, and we used Microsoft's lightGBM [55] as it is well documented and easy to use. Hyper-parameter tuning is needed to train a GBM to avoid over-fitting or under-fitting.

3. RESULTS

This section introduces the city-level electricity data we used and the results of the seven models. A more complete model evaluation and comparison are presented in Section 5.1.

3.1. Data

This study required ambient temperature data and city-level electricity usage data. Ambient temperature data were downloaded using the U.S. National Oceanic and Atmospheric Administration (NOAA) FTP API (<ftp.ncdc.noaa.gov>). We downloaded the hourly temperature of Station 722880-23152 for the Los Angeles metropolitan area, Station 724830-23232 for Sacramento metropolitan area, and Station 725053-94728 for the New York metropolitan area. We selected these three weather stations because they have the least amount of missing data compared with nearby weather stations.

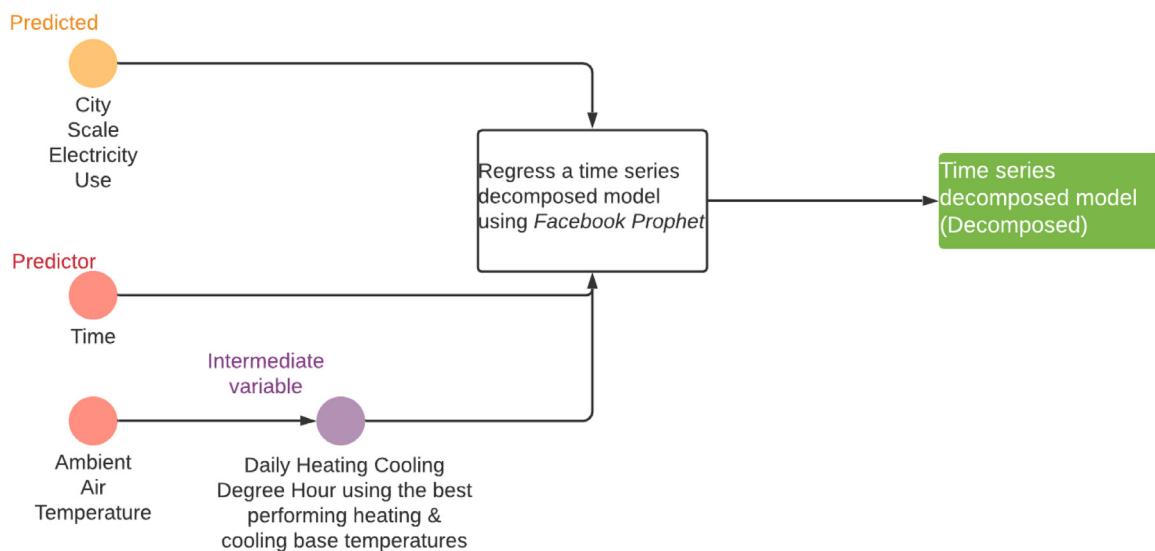


Fig. 4. Workflow to develop a time series decomposed model.

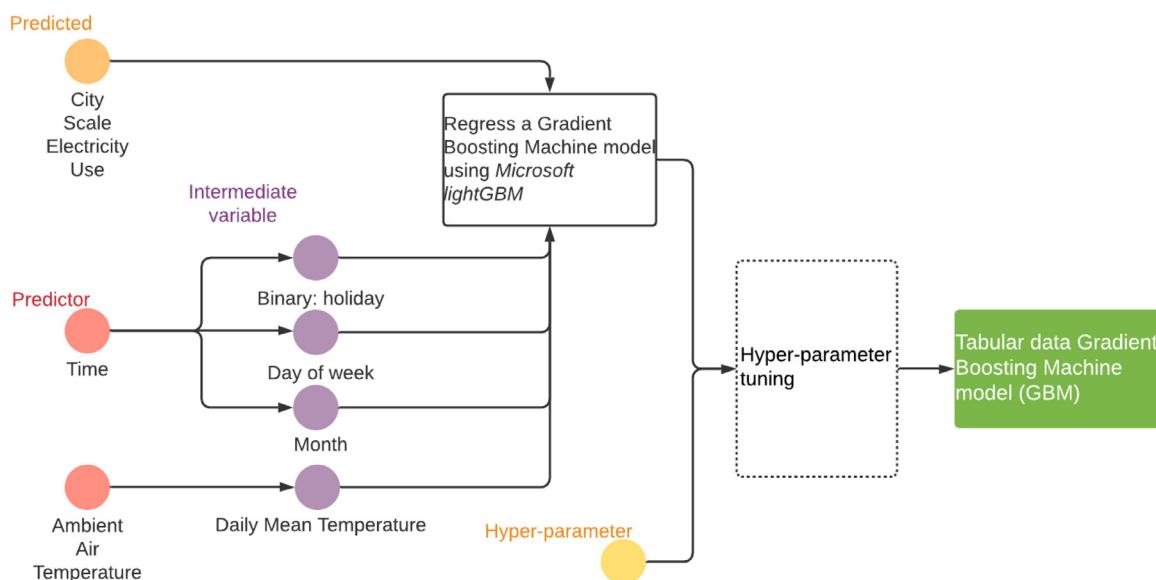


Fig. 5. Workflow to develop a time series decomposed model.

The city-level electricity usage data were downloaded from the U.S. Energy Information Administration (EIA). EIA collected and open sourced hourly demand data from all balancing authorities (BAs)² in the United States. For more details about the dataset, please refer to the paper [56] published at Nature Scientific Data. We used the data from three BAs: the Los Angeles Department of Water and Power (LADWP), Balance Authority of Northern California (BANC), and New York Independent System Operator (NYISO), representing the three metropolitan areas of Los Angeles, Sacramento, and New York, respectively. According to the ASHRAE climate zone definition [57], the three metropolitan areas belong to climate zone 3B for Los Angeles and Sacramento, and climate zone 4A for New York. The electricity usage in this study refers to the net demand to the electric grid, which is calculated as the net generation (NG) minus total interchange (TI) of each BA [58]. Therefore,

the city-level electricity consumption includes the electricity usage for industry, transportation (e.g., electric vehicles), buildings, and possibly agriculture in the cities and neighboring rural areas. For instance, the BANC dataset includes the electricity usage of the Modesto Irrigation District nearby. In this paper, we use "LA," "Sac," and "NY" as the abbreviations for the three balancing authorities. The raw electricity data are plotted in Fig. 6.

Fig. 7 shows the daily and weekly cycles of the electricity usage between July 2015 and September 2020. The daily demand peaks between 5 and 7 p.m. and bottoms at 3–5 a.m. The peak demand can be as much as twice the base load. The partial peak hours last longer in New York, compared with the other two metropolitan areas. In terms of the weekly cycle, a clear weekday-weekend pattern can be observed.

Fig. 8 plots the hourly electricity usage in four seasons between 2015 and 2020. Summer has a very high peak about 6 p.m., while the other three seasons have a relatively smooth load curve. The grid is more stressed during hot summer afternoons compared with other seasons. The load difference between working and non-working days is lowest during the nighttime, and highest during the peak hours. The Sacra-

² For the concept of Balancing Authorities, please refer to the website of U.S. Energy Information Administration: <https://www.eia.gov/todayinenergy/detail.php?id=27152>.

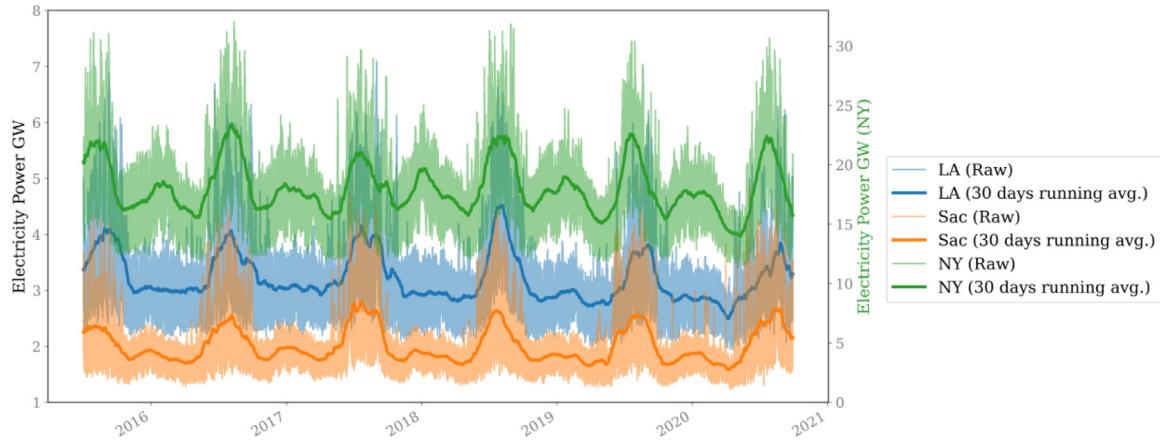


Fig. 6. Hourly electricity usage in the LA, Sac, and NY metropolitan areas.

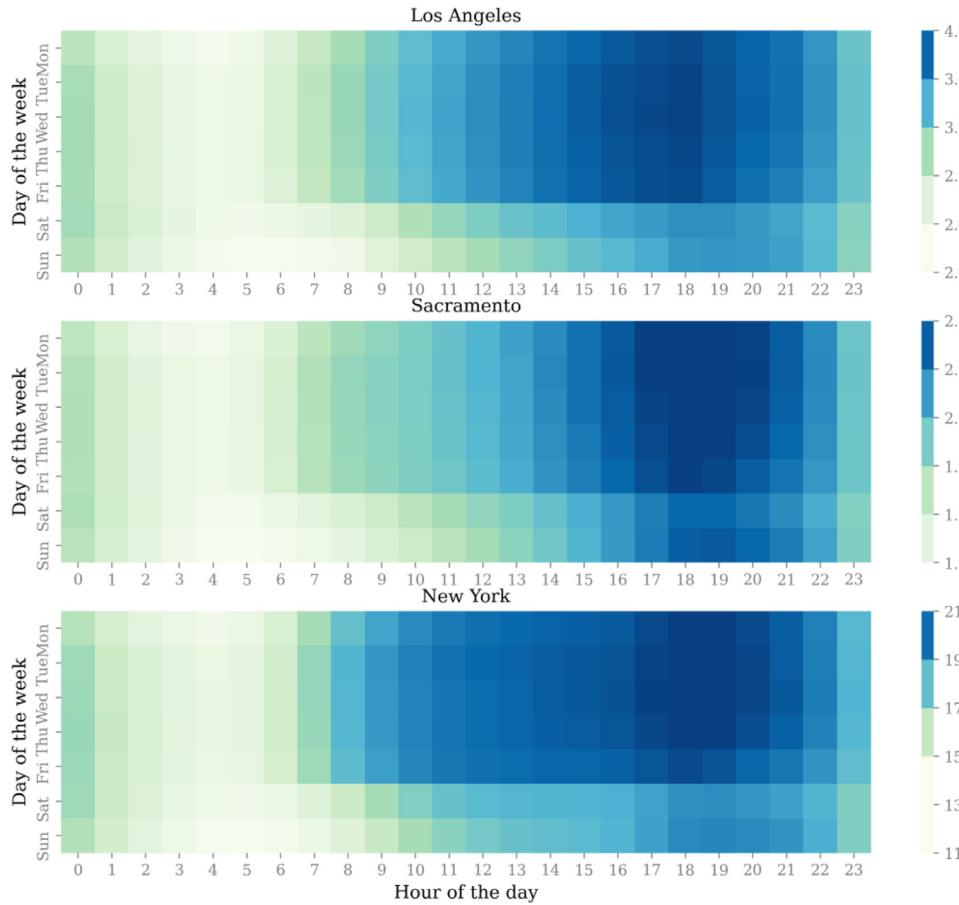


Fig. 7. Hourly electricity usage (GW) in the LA, Sac, and NY metropolitan areas.

metropolitan area has the largest peak-to-base ratio in summer, while the New York metropolitan area experienced the largest difference between working and non-working days.

These models can be used potentially for higher temporal resolution prediction (e.g., hourly load prediction) through adjustment of some hyper-parameters.

3.2. Linear models

3.2.1. Five-parameter change point model

As we observed a clear working day and non-working day difference in Fig. 8, we developed the five-parameter change point model separately for working and non-working days. The optimal change-point

temperature was selected by using the `scipy.optimize.curve_fit` function [31]. The input data points and the piecewise linear regression models are plotted in Fig. 9 and recorded in Table 1. The simple linear model can fit the daily energy consumption well. The R-squared values of all the three areas are above 0.82.

The regressed patterns are very similar in the Los Angeles and Sacramento metropolitan areas; the heating change point is about 14 °C, and the cooling change point is about 18 °C. In the New York metropolitan area, the heating change point and cooling change point are very close to each other, and accordingly the horizontal portion is very short especially during working days, indicating a short shoulder season in New York. There are two possible explanations: first, it could be because New York has a relatively short transition season, and second, a higher pro-

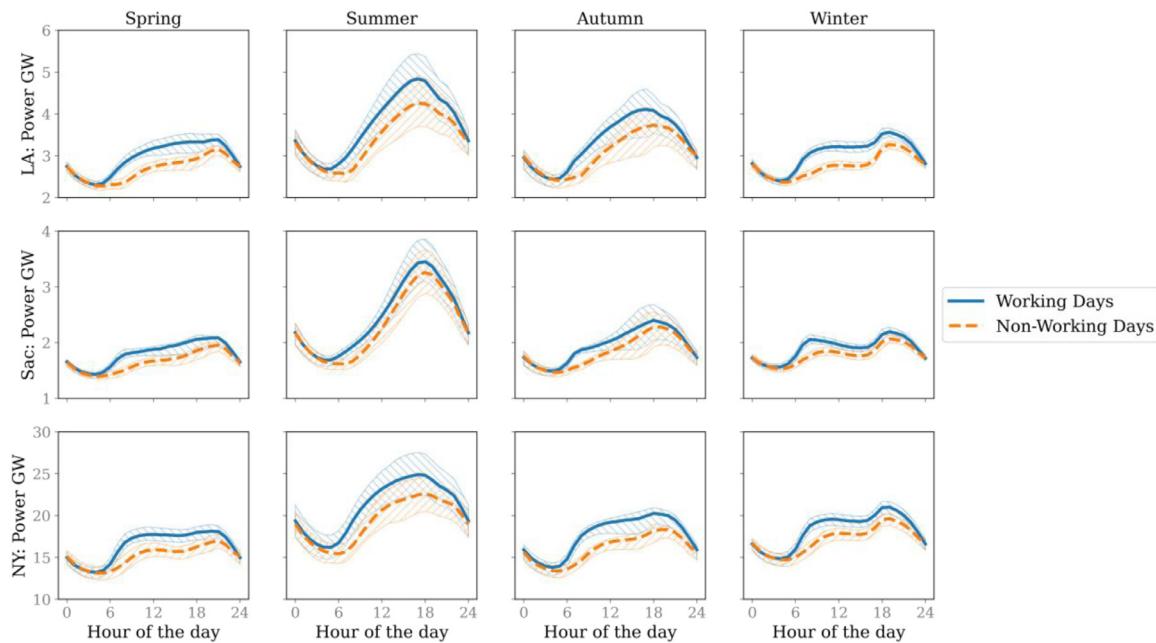


Fig. 8. Hourly electricity usage in the four seasons: Spring (Mar., Apr., May), Summer (Jun., Jul., Aug.), Autumn (Sep., Oct., Nov.), and Winter (Dec., Jan., Feb.).

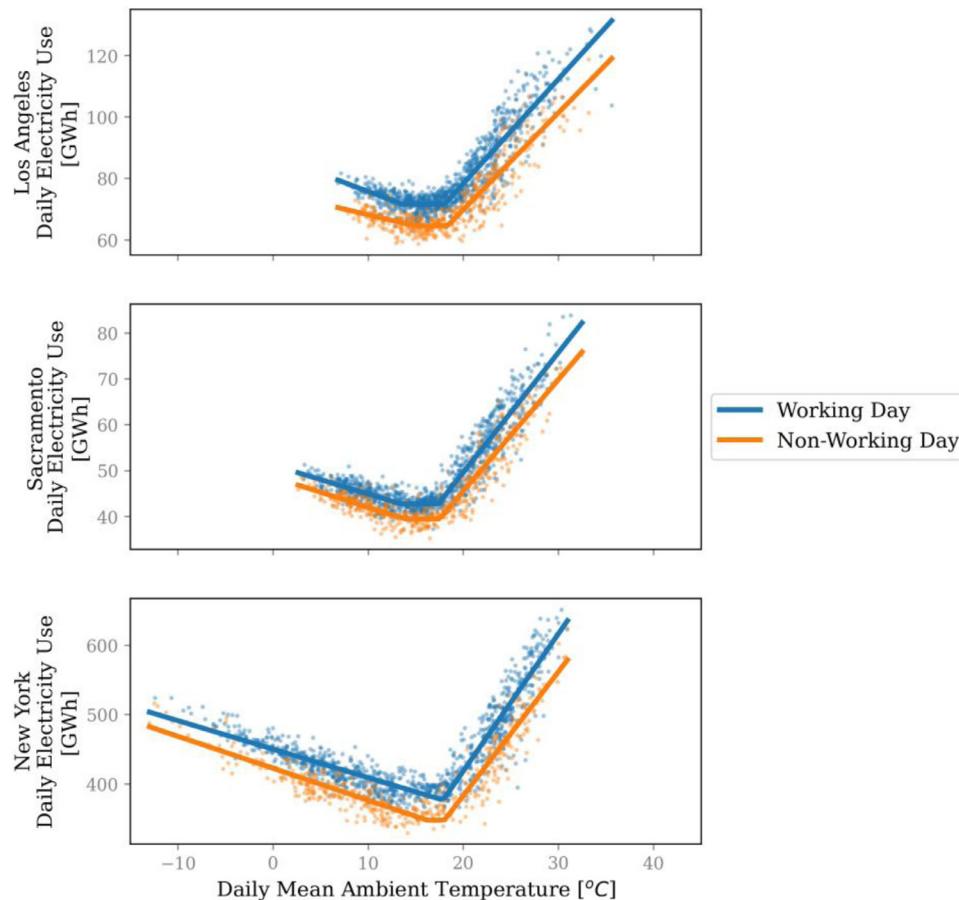


Fig. 9. Five-parameter change-point model of city-scale daily electricity usage.

portion of buildings in New York have adopted mechanical ventilation systems, which require either cooling or heating. Conversely, in the Los Angeles and Sacramento areas, a higher proportion of buildings might adopt natural ventilation or use free cooling for longer periods of time. In terms of the heating and cooling slopes, 1°C of ambient temperature increase leads to about 5% more daily energy usage compared with the

base load: 4.7% in Los Angeles, 6.2% in Sacramento, and 5.1% in New York. The heating slope is about one-third of the cooling slope in all three areas. A major reason for this is that we only considered electricity usage in this study. Almost every building uses electricity for cooling, while a significant number of buildings use other energy resources (such as natural gas or oil) for heating.

Table 1

Five-parameter change point model results.

		Regression Coefficients					Accuracy Metrics	
		Daily base load (GWh)	Heating base temp. (°C)	Heating slope (GWh/°C)	Cooling base temp. (°C)	Cooling slope (GWh/°C)	RMSE (GWh)	R-squared
Los Angeles	Working day	71.63	13.58	-1.16	18.02	3.40	4.73	0.85
	Non-working day	64.6	14.93	-0.73	18.32	3.15	4.65	0.82
Sacramento	Working day	42.74	13.43	-0.63	17.36	2.62	2.8	0.88
	Non-working day	39.45	13.9	-0.66	17.49	2.43	2.82	0.86
New York	Working day	376.73	17.88	-4.11	17.88	19.81	21	0.87
	Non-working day	347.66	16.25	-4.61	18.01	17.86	20.76	0.84

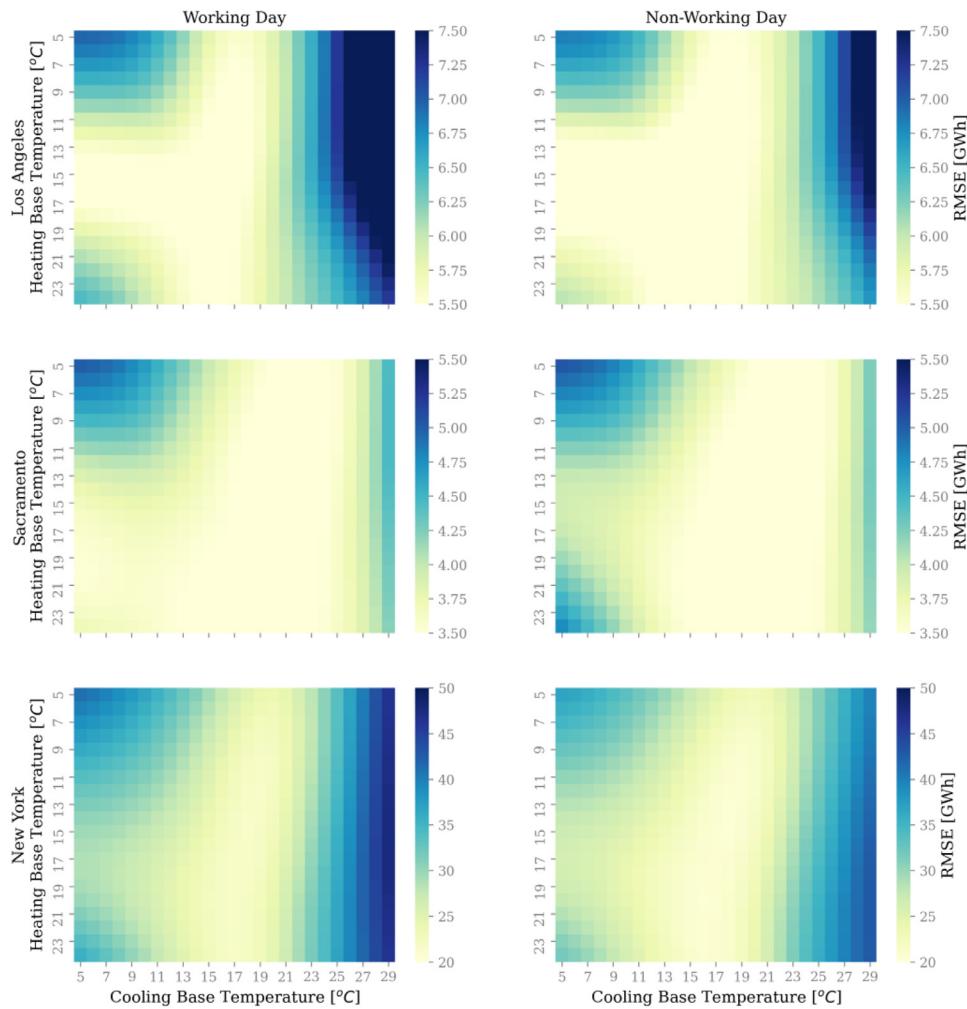


Fig. 10. Brute force search for the best performing heating and cooling base temperatures: the lighter color indicates a smaller prediction error, therefore corresponding to the heating and cooling base temperatures that deliver a more accurate prediction.

3.2.2. Heating and Cooling Degree Hour (HCDH) model

The first question we needed to answer before developing the HCDH model is which base temperatures (T_{bh} and T_{bc}) should be selected. ASHRAE Standard 90.1 recommended 18 °C as a heating base temperature and 10 °C as a cooling base temperature [59]. However, these two temperatures did not provide satisfactory accuracy for the Sacramento area, possibly because those base temperatures were selected for buildings that do not necessarily apply to the weather-sensitive energy consumption pattern at city scale. Therefore, we conducted a brute force search: i.e., testing different combinations of T_{bh} and T_{bc} to see which had the least root mean squared error (RMSE). Fig. 10 shows the brute force search results.

The regression result for the HCDH model with the best performing heating and cooling base temperatures is presented in Fig. 11 and Table 2. In Sacramento, a wider range of heating and cooling base tem-

perature could deliver a more accurate city-scale electricity use prediction. However, the base temperature needs to be carefully selected for Los Angeles and New York. Even though the best performing base temperatures are different for these three regions, the heating base temperature between 15 and 19 °C and the cooling base temperature between 13 and 19 °C can produce decent predictions. Unlike in the five-parameter model, the heating and cooling slopes are all positive values, as higher HDH and CDH always lead to higher electricity usage. Conversely, in the five-parameter model, lower temperatures lead to higher heating demand. HCDH models for all the three regions are not as accurate as the five-parameter model. We discuss the model accuracy in greater detail in Section 5.1. The take-home message is that the five-parameter change model is a better linear model, compared with the HCDH model, in terms of city-level electricity modeling. The selection of base temperature varies significantly for different cities. ASHRAE's recommendation

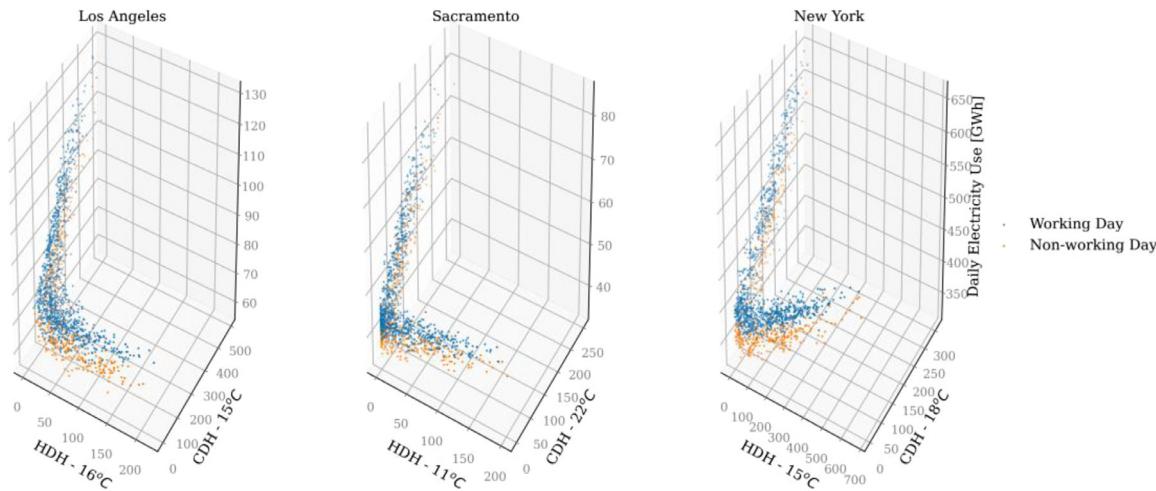


Fig. 11. Heating and Cooling Degree Hour models with the best performing heating and cooling base temperatures.

Table 2

Heating and Cooling Degree Hour model results.

		Regression coefficients					Accuracy metrics	
		Daily base load (GWh)	Heating base temp. (°C)	Heating slope (GWh/h °C)	Cooling base temp. (°C)	Cooling slope (GWh/h °C)	RMSE (GWh)	R-squared
Los Angeles	Working day	62.56	16	0.08	15	0.13	5.35	0.8
	Non-working day	56.21	16	0.08	15	0.12	5.22	0.77
Sacramento	Working day	42.82	11	0.03	22	0.18	3.34	0.83
	Non-working day	39.66	11	0.04	22	0.16	3.27	0.82
New York	Working day	376.4	15	0.21	18	0.83	21.63	0.86
	Non-working day	348.24	14	0.23	18	0.73	20.66	0.84

is not the optimal choice for either of the three metropolitan areas we tested because the ASHRAE recommendation was designed for building-scale energy usage but not for city-scale usage, as the latter includes other usage such as industry, transport, etc.

3.3. Machine learning model for time-series data

We selected the decomposed model implemented by Facebook Prophet as the time series modeling approach. We used four full years of data, from July 2015 to June 2019. We trained our model with the first three years and kept the last year for validation. We did not use 2020 data because the COVID-19 curtailment situation biased the electricity usage behavior. We discuss this issue in detail in Section 4.3. As we can observe from Fig. 12, the decomposed model can capture the general trend of city-scale daily electricity use in all three regions. However, the decomposed model tends to underestimate electricity use when the usage is high and overestimate the electricity use when the usage is low. This is because Prophet used a Fourier Series to model the periodic effects, as shown in Eqn 4 [42]. To avoid overfitting, the Fourier Series was usually truncated at $N = 3$ for weekly seasonality and at $N = 10$ for yearly seasonality [42]. Truncating the Fourier Series is like applying a low-pass filter to the seasonality term. As the result, the predicted electricity is smoothed. We can use a larger value of N to mitigate this problem, however at the risk of overfitting.

$$s(t) = \sum_{n=1}^N \left(a_n \cos\left(\frac{2\pi nt}{P}\right) + a_n \sin\left(\frac{2\pi nt}{P}\right) \right) \quad (4)$$

The strength of the decomposed model is that we can compare the magnitudes of different influential factors, as plotted in Fig. 13. For all the three metropolitan areas, the non-periodic trends of electricity usage are almost flat, which means, excluding the effects of temperature change and weekly and yearly periodical fluctuation, the city-scale daily

electricity usage did not change since 2015. A clear weekly cycle of electricity usage was observed in all three metropolitan areas: the daily electricity usage during weekdays was about 10% higher than that during weekends (4 out of 45 GWh in Sacramento, 10 out of 80 GWh in Los Angeles, and 40 out of 450 GWh in New York). An obvious yearly cycle was observed in all three cities. It is worth noting that these yearly cycles excluded the weather effect in the decomposed model. There are two yearly peaks in all three metropolitan areas. The summer load peak was significantly higher than the winter peak in both Sacramento and Los Angeles, while in New York the summer peak and winter peak were similar. Another difference was the relative magnitude of the yearly cycle, which was 1.5 times that of the weekly cycle in Los Angeles, 2 times in Sacramento, and 1.5 times in New York. The last component is the weather-related term. The temperature-sensitive electricity usage was about 30 GWh in Sacramento, 40 GWh in Los Angeles, and 200 GWh in New York. The proportion of weather-related load, compared with the base load, was similar in all three regions.

3.4. Machine learning models for tabular data

We applied four tabular data modeling approaches for city-level electricity usage prediction (Gradient Boosting Machine, Random Forest, Support Vector Machine, and Neural Network). Similar to the previous section, we used four full years of data, from July 2015 to June 2019. We trained our model with the first three years and kept the last year for validation. The result is shown in Fig. 14. We observe that the electricity usage in all the three metropolitan areas are well captured by the Gradient Boosting Machine models. Unlike the linear models, it is challenging to record the decomposed and Gradient Boosting models using functions and coefficients, as shown in Tables 1 and 2, because of the higher model complexity. Due to the space constraint, we presented

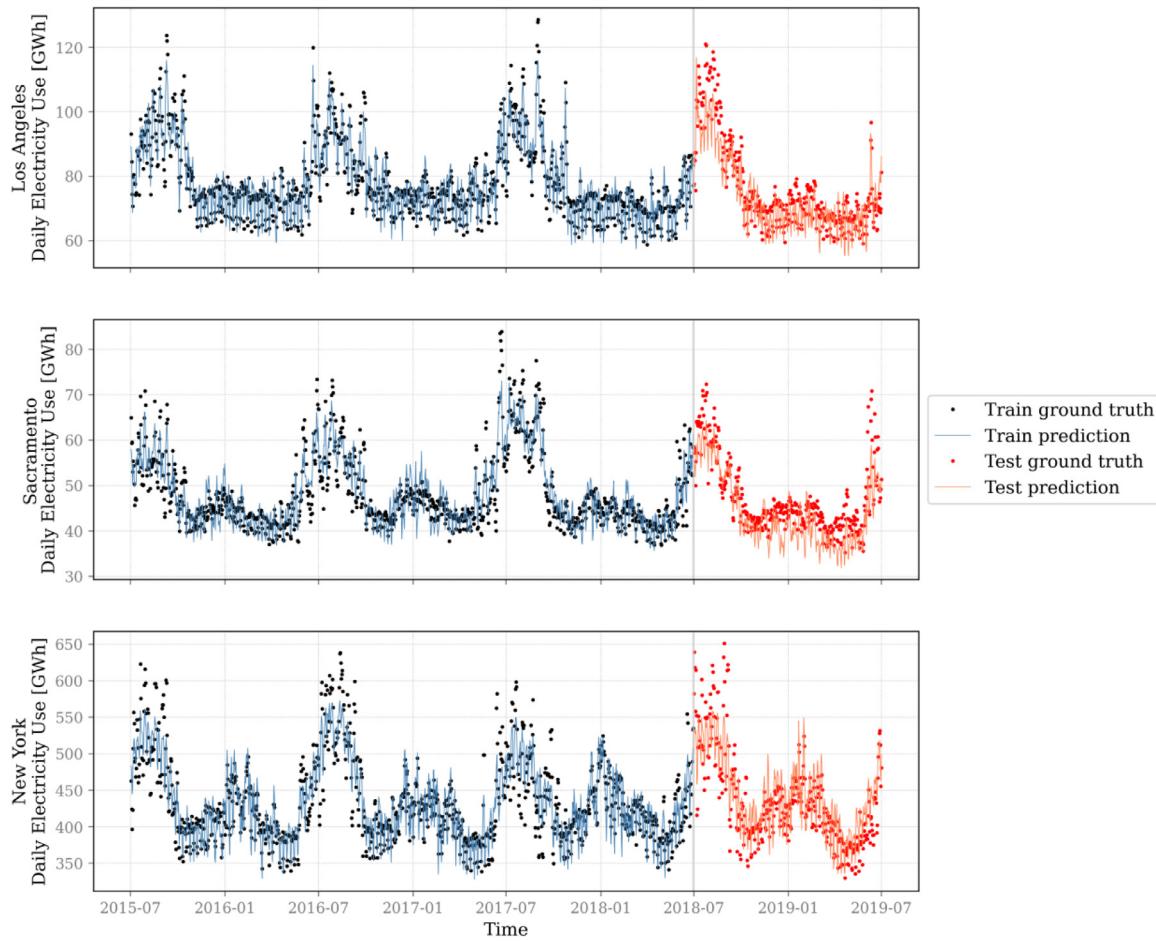


Fig. 12. Predictions from the time-series decomposed model and ground truth. From top to bottom – the Los Angeles, Sacramento, and New York metropolitan areas.

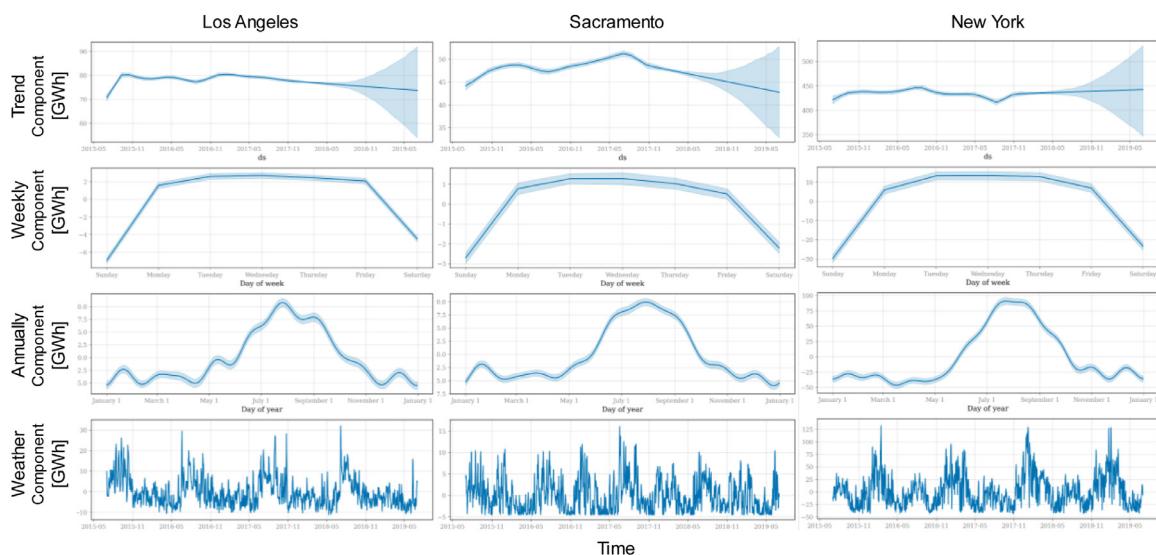


Fig. 13. Decomposed time series model results. From left to right – the Los Angeles, Sacramento, and New York metropolitan areas. From top to bottom – general trend (non-periodic change), weekly cycle, yearly cycle, and influence of extra regressor (heating and cooling degree hours).

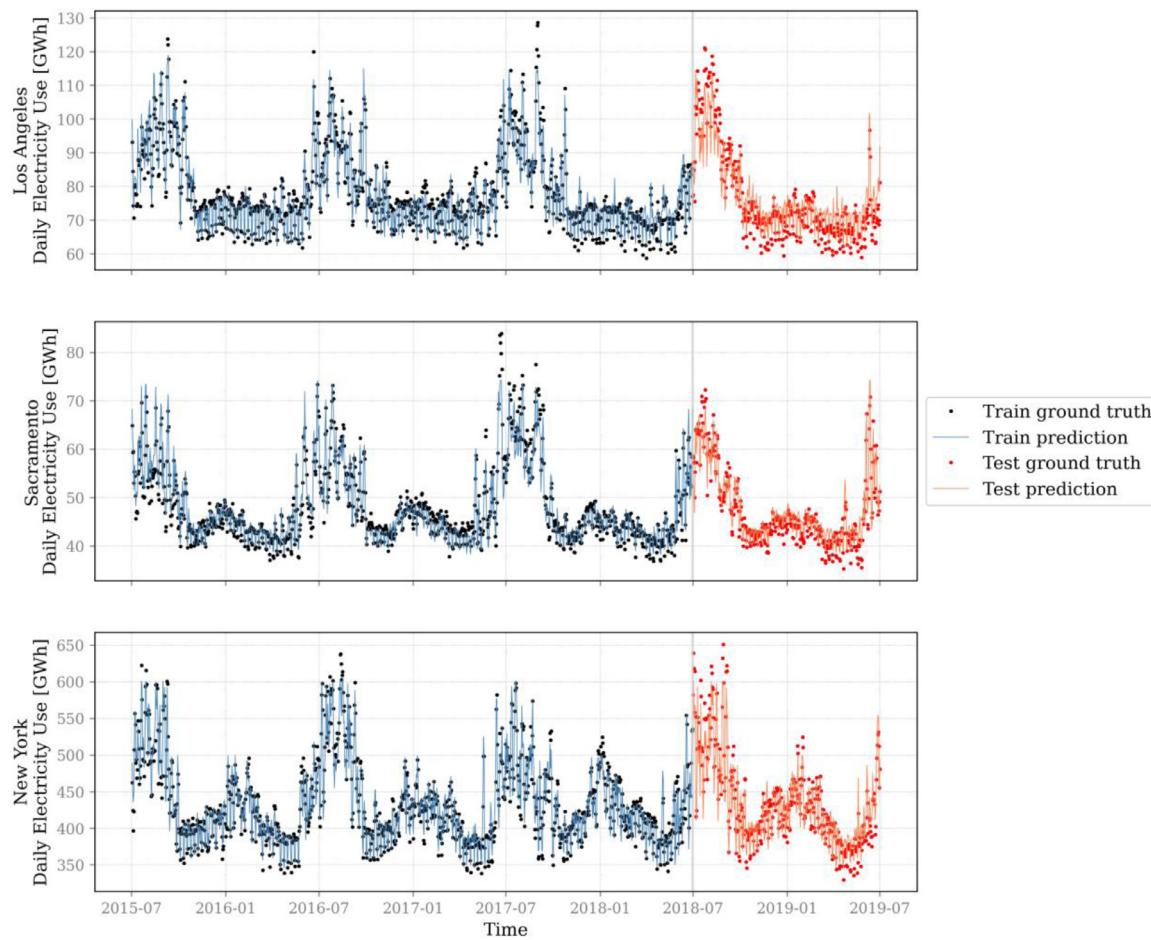


Fig. 14. Predictions from Gradient Boosting Machine model and ground truth. From top to bottom – Los Angeles, Sacramento, and New York metropolitan areas.

the results of the other three baseline models (Random Forest, Support Vector Machine, Neural Network) in the Appendix.

4. Applications

A natural question after we developed seven data-driven models is which model perform the best. In this section, we compare the seven models first and used the best performing model to answer two practical questions: a) what is the impact of ambient temperature on city-scale daily electricity use, especially during a heat wave event; b) how unexpected public health event, for instance the COVID-19 pandemic, would influence the city-scale daily electricity use.

4.1. Model comparison

Model accuracy is one of the top, if not the most important, criteria to compare different data-driven models. To facilitate the model comparison, we trained the model using three years of data (July 2015 to June 2018) and used the final year (July 2018 to June 2019) for model validation. We did not use 2020 data, as we wanted to decouple the effect of unexpected events from the model comparison. We developed seven data-driven models: two linear models (five-parameter change point model, Heating and Cooling Degree Hour model), three conventional machine learning models (Random Forest, Support Vector Machine, Artificial Neural Network), time-series decomposed model, and Gradient Boosting Machine model. The model was trained on the

training set only and then evaluated on the test set. The prediction accuracy is shown in Table 3. Three metrics were used for comparison: mean absolute error (MAE), root mean squared error (RMSE), and cross validation root mean squared error (CVRMSE). Additionally, we plotted the CVRMSE of the seven algorithms of the three metropolitan areas in Fig. 15.

As shown in Table 3, the top-down data-driven models can provide accurate city-level electricity demand. All of the seven data-driven models can predict the city-scale daily electricity use with an accuracy higher than 90%. Overfitting is not a problem for Sacramento or New York. But in Los Angeles, the model performance on the test dataset deteriorated more significantly than it did for the other two regions; while in the other two regions, the models performed almost equally well on the test and train dataset. There could be two potential reasons for this. First, the electricity usage behavior of Los Angeles changed in the test dataset, and second, some other hidden factors that significantly influence Los Angeles's electricity demand was identified and included in the data-driven model.

Simple linear regression models (piecewise linear regression and multivariate linear regression) is able to provide decent prediction. Among the two linear models, the five-parameter change point model outperformed the Heating and Cooling Degree Hour model in all three metropolitan areas. CVRMSE of five-parameter model is in the range of 5.2%–8.2%; in comparison, CVRMSE of HCDH model is in the range of 5.4%–8.8%. Additionally, the five-parameter change point model is easier to implement because it does not need to determine the best-performing heating and cooling base temperatures,

Table 3
Model comparison in terms of prediction accuracy.

	MAE_train [GWh]	MAE_test [GWh]	RMSE_train [GWh]	RMSE_test [GWh]	CVRMSE_train (%)	CVRMSE_test (%)
Los Angeles metropolitan area						
5-Parameter	3.01	4.97	4.08	6.38	5.25	8.20
Degree Hour	3.72	5.27	4.74	6.85	6.10	8.80
Decomposed	2.58	3.97	3.45	5.34	4.43	6.86
lightGBM	1.99	4.16	2.68	5.06	3.44	6.50
RF	2.23	4.61	3.08	5.70	3.95	7.33
SVM	2.16	4.43	3.19	5.53	4.10	7.11
NN	2.85	5.01	3.66	6.07	4.70	7.81
Sacramento metropolitan area						
5-Parameter	2.04	2.25	2.79	2.95	5.82	6.16
Degree Hour	2.49	2.66	3.32	3.34	6.93	6.97
Decomposed	2.28	3.43	2.99	4.26	6.24	8.89
lightGBM	1.42	1.69	2.03	2.21	4.23	4.60
RF	1.52	2.05	2.12	2.72	4.43	5.67
SVM	1.48	1.69	2.31	2.13	4.81	4.44
NN	1.74	2.08	2.39	2.51	4.98	5.23
New York metropolitan area						
5-Parameter	14.78	17.36	20.26	22.57	4.66	5.19
Degree Hour	15.03	17.90	20.83	23.22	4.79	5.34
Decomposed	16.89	20.92	22.42	27.16	5.16	6.25
lightGBM	9.74	13.26	13.61	17.85	3.13	4.10
RF	11.34	16.27	15.56	22.10	3.58	5.08
SVM	16.68	20.42	26.73	31.80	6.15	7.31
NN	11.20	14.05	15.20	18.68	3.50	4.30

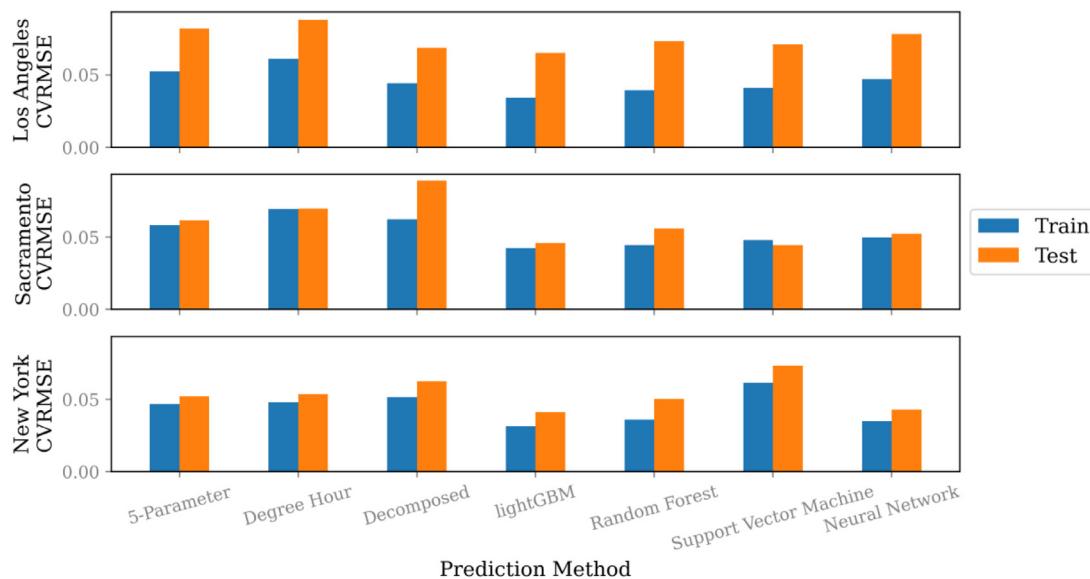


Fig. 15. Coefficient of Variation of Root Mean Squared Error (CVRMSE) of the seven algorithms for the three metropolitan areas. From top to bottom – the Los Angeles, Sacramento, and New York metropolitan areas.

which may vary by cities with different weather and electricity use behavior.

In the category of Machine Learning models, conventional ML models (RF, SVM, and NN) deliver similar accuracy. Even though the same hyper-parameter and model architecture is applied, the performance of the ML models depend on individual data set: SVM performs the best in Los Angeles and Sacramento but performs the worst in New York. In this regard, the model generalizability is in question.

The GBM model outperforms the decomposed model and other baseline machine learning models (RF, SVM, and NN) in all three metropolitan areas with a decrease in CVRMSE between 0.4% (Los Angeles) to 4.3% (Sacramento). Electricity usage might not be necessarily modeled as time-series data if the temporal information is carefully encoded (in

this study, encoded by two variables: the month of year and day of week). Compared with time-series modeling, the tabular data model is more robust to missing data, because time-series modeling uses the sequential ordering of input data to represent temporal information. It is a problem if some data is missing, and data imputation is needed in this case, which adds another layer of complexity in terms of data preprocessing. However, tabular data models encode temporal information using extra features (such as day of week) and therefore do not suffer if some data is missing.

Comparing the winner of linear model (5-parameter) with the winner of machine learning model (lightGBM), the lightGBM model can improve the model accuracy with a margin of 1.1% to 1.7%.

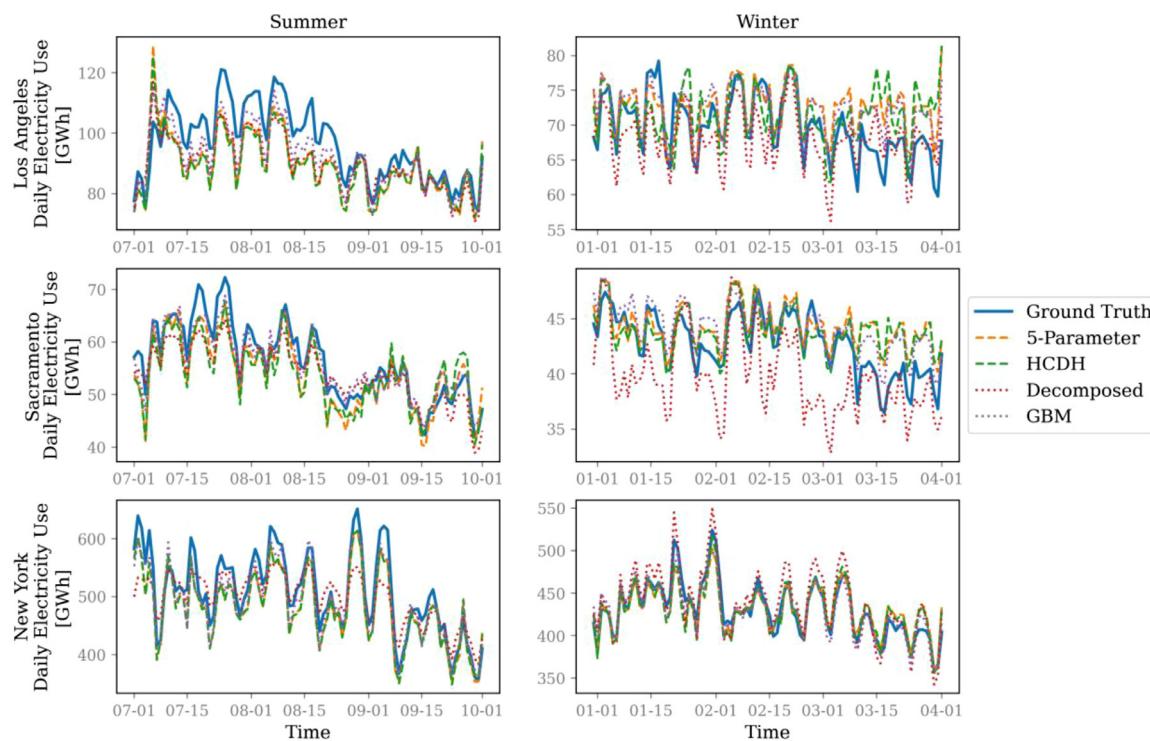


Fig. 16. Model prediction of city-scale daily electricity consumption (GWh) on the test dataset (July 2018–June 2019). From top to bottom – the Los Angeles, Sacramento, and New York metropolitan areas. Left is summer season, and right is winter season.

Next, we compared the model prediction with the ground truth on the test dataset, as shown in Fig. 16: the ground truth is plotted in solid line, the two linear models are plotted in dashed lines, and the two machine learning models are plotted in dotted lines. To avoid making the plot too distracting, we only compared the results from the two linear models and the two machine learning models, the Decomposed model and GBM model.

As can be observed from Fig. 16, the general trend of city-scale daily electricity usage can be predicted by all of the four data-driven models. The only exception is the winter case of Sacramento region. The five-parameter, HCDH and GBM models predict the electricity use well before mid-March and overestimate the electricity use after mid-March. Contrarily, the decomposed model underestimates the load before mid-March but captures the load trend in late March. It can also be observed that the data driven models work well during some periods in some cities (e.g., New York) but not in other cases (e.g., Summer in Los Angeles and Winter in Sacramento).

In addition to the model accuracy, we also care about the models' interoperability and generalizability. The GBM method can deliver very accurate predictions; however, the results are not interpretable. It should be very careful to extrapolate the GBM models. Linear model and decomposed time series might be less accurate; however, their results have clear physics implications. For instance, the results of linear models are easy to understand and explain, while the results of a decomposed model can be used to examine the relative magnitudes of different influencing factors.

As discussed above, different models have their own strengths and weaknesses, because of their unique mathematical structures. As the well-recognized British statistician George E. P. Box once said, “all models are wrong, but some are useful.” Which model should be developed depends on how you are going to use your model. For instance, the decomposed time-series model is the best candidate to understand the relative importance of different influencing factors, but it might not be a good choice to explore the sensitivity and elasticity to a specific vari-

able, as you need to organize your numerical experiment in the format of time-series data.

4.2. Heat wave impact on city electricity usage

As a result of climate change, heat waves happen at an increasing frequency worldwide [60]. Since cooling is a major electricity consumer, heat waves lead to more frequent use of air conditioning, and thus higher electricity usage. The increasing demand during a heat wave poses extra challenges to the grid infrastructure. In this section, we explore the impact of increasing ambient temperature on city-level electricity demand during heat waves.

The five-parameter change point model provides the most straightforward way to understand how higher temperature drives up the electricity demand, as the regressed cooling slope (β_c) depicts the sensitivity of electricity demand on ambient temperature. Gradient Boosting Machine is another good candidate because it accepts tabular data as model inputs. We can examine the temperature sensitivity by observing how the outputs (electricity demand) change with different inputs (ambient temperature). On the other hand, a decomposed time-series model is not a good choice because it requires time-series data as inputs. It is challenging to organize the numerical experiments of varying ambient temperature in the format of time-series data. Therefore, in this subsection, we applied a five-parameter linear model and Gradient Boosting Machine to explore the gradient of city-level electricity demand to ambient temperature.

The relation between the electricity demand and ambient temperature is plotted in Fig. 17. The prediction of five-parameter model is a monotonically increasing linear line, as we expected. The daily electricity demands increase at the rate of 2.5 GWh/°C, 3.3 GWh/°C, and 19.1 GWh/°C in the Sacramento (6.2%/°C), Los Angeles (4.7%/°C), and New York (5.1%/°C) metropolitan areas, respectively. The daily peak hourly loads increase by 0.19 GW/°C, 0.22 GW/°C, and 1.04 GW/°C in the Sacramento, Los Angeles, and New York metropolitan areas, respec-

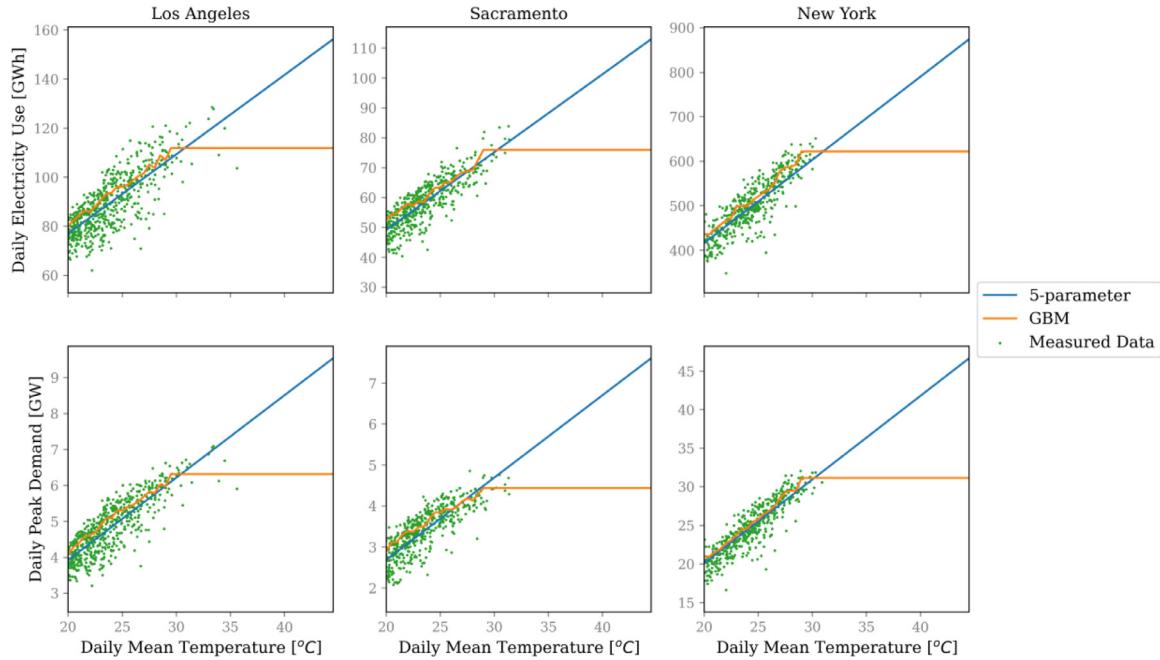


Fig. 17. The impact of ambient temperature on daily total electricity consumption and peak demand. From left to right – the Los Angeles, Sacramento, and New York metropolitan areas. The top figures are daily total consumption, and the bottom figures are daily peak demand.

tively. The prediction of the lightGBM aligned with the five-parameter model until the daily mean ambient temperature was above about 30 °C. When the ambient temperature was above 30 °C, the prediction of lightGBM was a constant value. This is because the GBM model has a poor generalizability. Its prediction is not trustworthy if the input data are outside the space covered by the training datasets.

4.3. COVID-19 pandemic impact on city electricity usage

The second question we wanted to answer is how the measures government and individuals took (e.g., shelter in place, stay-at-home, business shutdown, or reduced operation) to mitigate the impact of the COVID-19 pandemic in U.S. cities influenced city-level electricity consumption. As discussed in [9], city-scale electricity use can be an real time indicator of the strictness of shutdown measures implemented.

To evaluate the impact of COVID-19 on city-scale electricity consumption, we used the GBM model due to its high accuracy compared with other data-driven models. We use the pre-pandemic data (2015 Jul. to 2020 Mar.) to train the GBM model, and then apply the trained model to predict the electricity use after the pandemic. Our hypothesis is: if the electricity use patterns changed due to the COVID-19 pandemic, the model trained with the pre-pandemic data could not be used to accurately predict the electricity use after the pandemic. We plotted our results in Fig. 18.

Our hypothesis was validated by results in Fig. 18: the model trained with the pre-pandemic data tend to overestimate the electricity use after mid-March when lockdown measures were put in place in majority of U.S. cities: the predicted electricity usage after the lockdown (green line) was constantly above the actual electricity usage (dotted orange line). Conversely, the electricity demand prediction before the curtailment (blue line) was close to the actual demand (dotted orange line). This discrepancy between the forecasted and the real electricity usage after 2020 mid-March indicates that the COVID-19 curtailment changed electricity demand behavior. Another interesting finding was the discrepancy between the forecasted and real electricity usage peaks in April. In New York, the projected and real electricity usage matched again after June, which is consistent with the time COVID-19 was controlled there.

Fig. 19 summarizes the changes in monthly electricity consumption for the three metropolitan areas. The COVID-19 curtailment in the United States started in the mid-March 2020. The electricity usage reduction peaked in April 2020, reaching more than 10% in all three metropolitan areas. With the economy gradually opening up in the summer, electricity usage increased. The effect of the COVID-19 curtailment on electricity usage was highest in the Los Angeles metropolitan area and the smallest in the Sacramento metropolitan area, probability because the travel and entertainment industries, which were much more influenced by the COVID-19 pandemic, account for a higher weight of Los Angeles's economy and thus electricity use.

5. Discussion

5.1. Implications

City-level electricity prediction has wide applications and is important to enhance energy security and resilience. In this study we developed data-driven models to predict daily electricity consumption at city or metropolitan scale. We open sourced the data and the code to help researchers to a) reproduce our work, b) develop their own data-driven models using our code as the starting point, and test and compare our models with theirs, and c) apply our models to other cities' data and do the evaluations. The major contributions of this study include:

First, we applied data-driven models to study how city-scale electricity consumption can be influenced by extreme heat waves and the unexpected public health event, which can enhance our understanding of city-scale energy use dynamics and support grid operators on load forecast and generation resources planning. Such predictions can inform utilities and state or city governments to secure energy supply and avoid or minimize power outages during extreme weather events (e.g., heat waves).

Second, we implemented and evaluated seven data-driven models including the five-parameter change-point model, the Heating/Cooling Degree Hour model, the time series decomposed model implemented by Facebook Prophet, the Gradient Boosting Machine implemented by Microsoft lightGBM, and three widely-used machine learning models (Random Forest, Support Vector Machine, Neural Network).

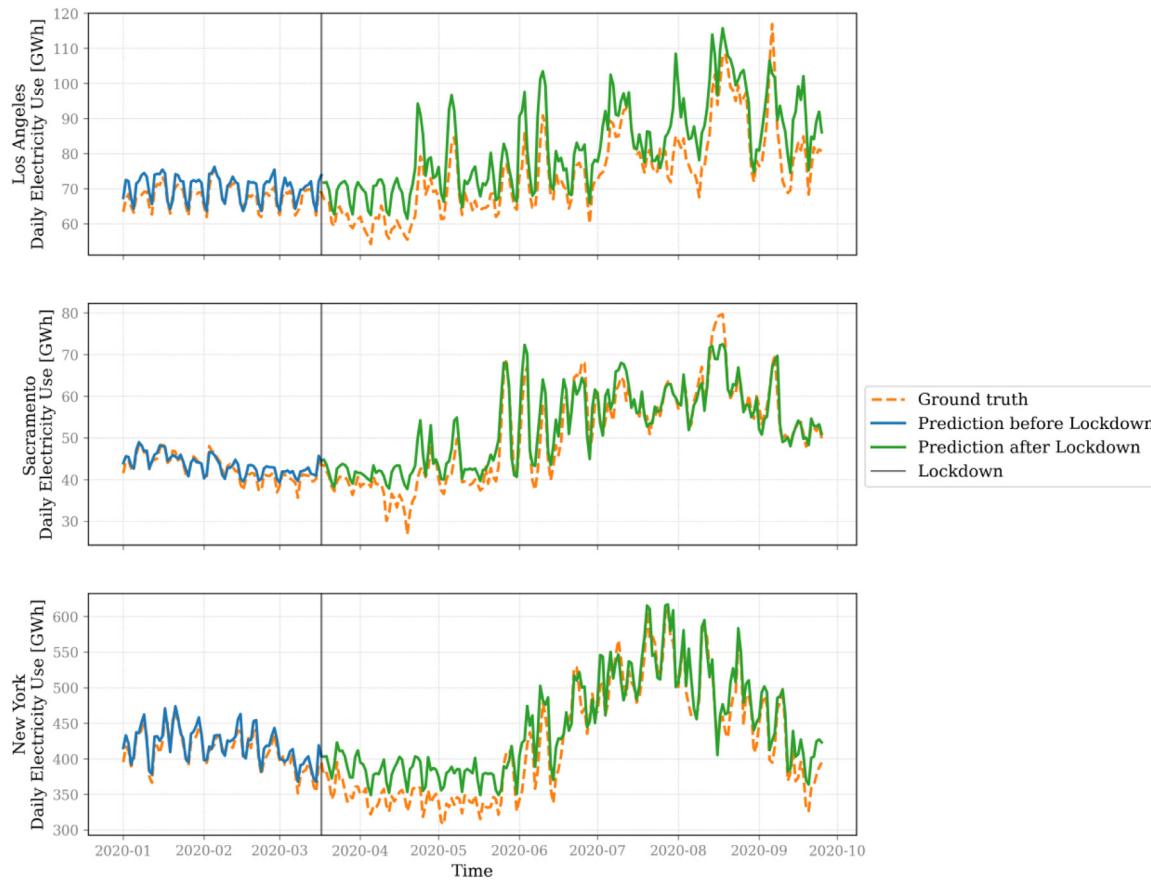


Fig. 18. Predicted vs. real city-scale daily electricity consumption before and after the COVID-19 lockdown. From top to bottom – the Los Angeles, Sacramento, and New York metropolitan areas.

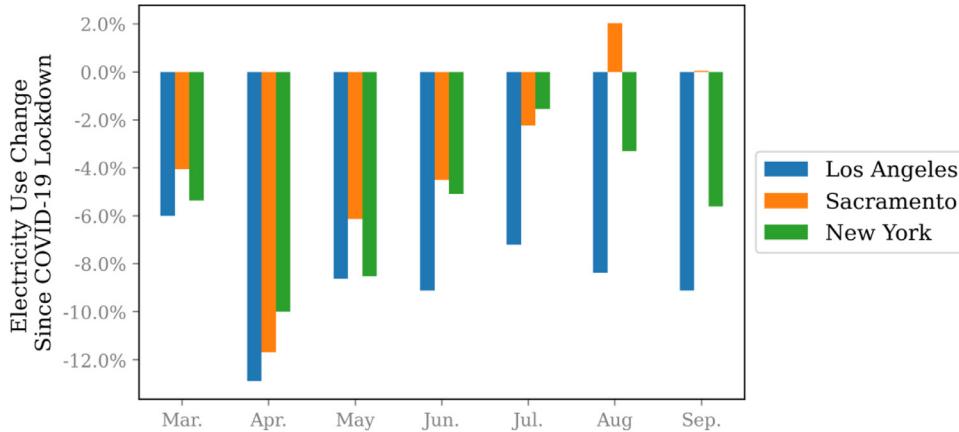


Fig. 19. Monthly electricity use changes since COVID-19 pandemic for the three metropolitan areas.

Third, we answered some key questions in developing city-scale electricity prediction models, including whether to model the electricity use as time series data or as tabular data, how to select the heating and cooling base temperature to achieve high accuracy prediction, and what is the performance margin of machine learning models compared with straightforward linear regression models.

The implications of our findings are as follows:

First, city-scale energy usage is highly correlated with ambient temperature, as heating and cooling load accounts for a high proportion of city-level energy usage. Under the current weather conditions, weather-sensitive electricity consumption accounts for 30%–50% of total electricity usage. Every degree Celsius ambient temperature increase in summer leads to about 5% more electricity usage compared with the base

load. As a result of climate change, heat waves will happen more frequently, leading to higher electricity demand in summer. The power generation and transmission infrastructure needs to be prepared to meet this new challenge.

Second, during the COVID-19 pandemic, the city-scale electricity consumption dropped by more than 10% in April 2020 and recovered in summer when the pandemic was started getting in control. The impact of COVID-19 on city-scale electricity use depends on the strictness of the lockdown measures governments took as well as the local economic structure.

Third, a linear model is a simple but powerful tool that can provide decent city-level electricity usage prediction with a CVRMSE between 5% and 10%. Other benefits of the linear model include understand-

able results and good interoperability. The Five-Parameter Change Point model slightly outperforms the Heating Cooling Degree Hour model in all the three metropolitan areas.

Fourth, when applying the Heating Cooling Degree Hour model, the heating and cooling change points and the base temperature need to be carefully selected. There is no one set of change points or base temperature that can be applied universally for every city. The ASHRAE recommended values are not a one-size-fits-all solution. Even though the best performing base temperature varies for these three regions, the heating base temperature between 15 and 19 °C and the cooling base temperature between 13 and 19 °C can produce decent predictions.

Fifth, machine learning models might not necessarily outperform simple linear models. For instance, the time-series decomposed model failed to generate more accurate prediction in Sacramento and New York regions. However, the decomposed model provides us with a unique tool that can be used to compare the relative importance of different components, such as the general trend, weekly and yearly cycle, and weather-related demand.

Sixth, Gradient Boosting Machine performs the best compared with the other models tested in this study. The prediction error of GBM could be as low as 4%–6% on the test dataset. However, users should be cautious when using those black-box models for new data, especially data with trends or patterns not covered by the training dataset.

Lastly, in addition to accuracy, we also compared the modeling complexity, computational complexity and interpretability of the models. Linear models and lightGBM are easier to implement because the software dependency is simpler and those algorithms are less sensitive to missing data. On the contrary, time-series models need to handle missing data first. In terms of computational complexity, linear models are the fastest, followed by the GBM models, while time-series models are the slowest. As for the interpretability, the linear models and time-series decomposed model have clearer implications and are easier to understand compared with the GBM model.

5.2. Limitations

A key concern about the data-driven approach is whether the conclusion drawn from data-driven methods can be extrapolated to cases where the data are not available. For instance, can we assume the slope between electricity demand and the ambient temperature is constant when the ambient temperature increases further? Will the consumers' energy consumption behavior change over time? For instance, Auffhammer and Mansur (2014) argued that climate change will affect energy consumption by changing how consumers respond to weather shocks (the intensive margin) in the short run, but in the long run people might adapt to it (the extensive margin) [61]. Unfortunately, the data-driven approach used in this study is unable to answer those questions.

The conclusion of this article is based on the analysis of city-level electricity usage in three U.S. cities (Los Angeles, Sacramento, and New York). Though these three cities are diversified (with different climate, scale and economy structure), it should be noted that the findings may not be directly applicable to other cities. This is a common problem for data-driven modeling. To help address this problem, we open sourced the code so that other researchers can retrain the model with data of other cities and evaluate the prediction performance.

Another limitation of this study is that, although urban and rural areas may have different microclimates, we only used ambient temperature data from one NOAA listed weather station per city. This was because we do not have information about the geographical distributions of the energy consumers (e.g. residential and commercial buildings, factories, transportation, and other infrastructure). So there is no hint as to which weather stations should be used. The models we proposed in this study potentially can provide more granular predictions should the distribution information become available in the future.

6. Conclusions

City-level electricity usage is temperature sensitive, because heating and cooling is a major energy consumer. As a result of climate change, extreme weather events happen more frequently. A more accurate electricity demand forecast, accounting for extreme weather events and associated load impacts, is needed to enhance energy security and resilience of the electric grid.

A literature review identified three common approaches to model the city-level electricity usage: linear models, machine learning models for time series data (Autoregressive integrated moving average, Recurrent Neural Network/Long Short Term Memory), and machine learning models for tabular data (neural network-based, decision tree based, and others). In this study, we developed and compared seven data-driven models: a five-parameter change-point model, a Heating Cooling Degree Hour model, a decomposed time series model implemented by Facebook Prophet, a Gradient Boosting Trees model implemented by Microsoft lightGBM, and three conventional machine learning models (Random Forest, Support Vector Machine, and Neural Network). The decomposed model has rarely been used in this field; however, lightGBM has been proven as a top performer in city-scale energy demand prediction.

We tested seven models with the city-level (including the city and surrounding rural area) electricity usage data from three metropolitan areas in the United States: Sacramento, Los Angeles, and New York. All the models can predict the city-level electricity demand well, with a CVRMSE less than 10%. The five-parameter model outperforms the HCDH model. Gradient Boosting Machine is the most accurate model among the seven. The CVRMSE of lightGBM on the test dataset was 6.5% for Los Angeles, 4.6% for Sacramento, and as low as 4.1% for New York metropolitan area. Though not as accurate as lightGBM, the decomposed time series model provides us a unique chance to decouple and compare the effects of different driving factors (weather-related, yearly and weekly cycle, general trend) of energy demand.

We applied the best performing models to explore how extreme weather events (heat wave) and unexpected public health events (COVID-19 curtailment) influenced the city-level electricity demand. Under the current weather condition, weather-related electricity consumption accounts for 30%–50% of daily electricity usage of a city. Every degree Celsius increase of ambient temperature leads to an increase of about 5% more daily electricity usage compared with the base load in the three metropolitan areas. The COVID-19 curtailment reduced city-level electricity demand. Compared with the pre-pandemic same month in 2019, the daily electricity usages during the 2020 COVID-19 pandemic decreased by 2%–12% in the Sacramento, Los Angeles, and the New York metropolitan areas.

All the data and code used in this paper is open sourced at Github, <https://github.com/LBNL-ETA/City-Scale-Electricity-Use-Prediction>.

Acknowledgments

This research was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Building Technologies of the United States Department of Energy, under Contract No. DE-AC02-05CH11231.

Appendix

The prediction results of the three baseline machine learning models (Random Forest (Figure A1), Support Vector Machine (Figure A2), and Neural Network (Figure A3)) are included in this appendix.

All the three data-driven models under-estimated the electricity usage of New York in the summer of 2018, which means some influencing factors (in addition to the ambient temperature and time) were not captured. SVM under-estimated the summer load increase in New York more than the other two machine learning algorithms.

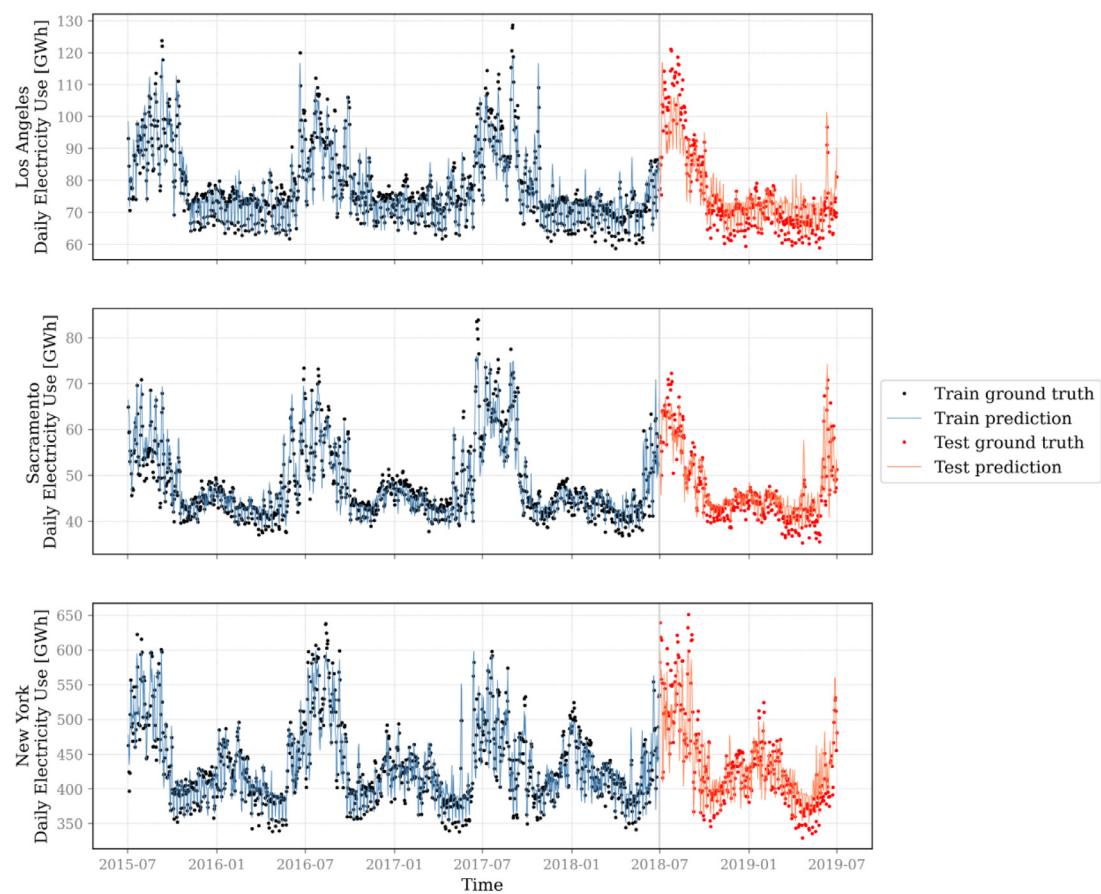


Fig. A1. Predictions from Random Forest and ground truth. From top to bottom – Los Angeles, Sacramento, and New York metropolitan areas.

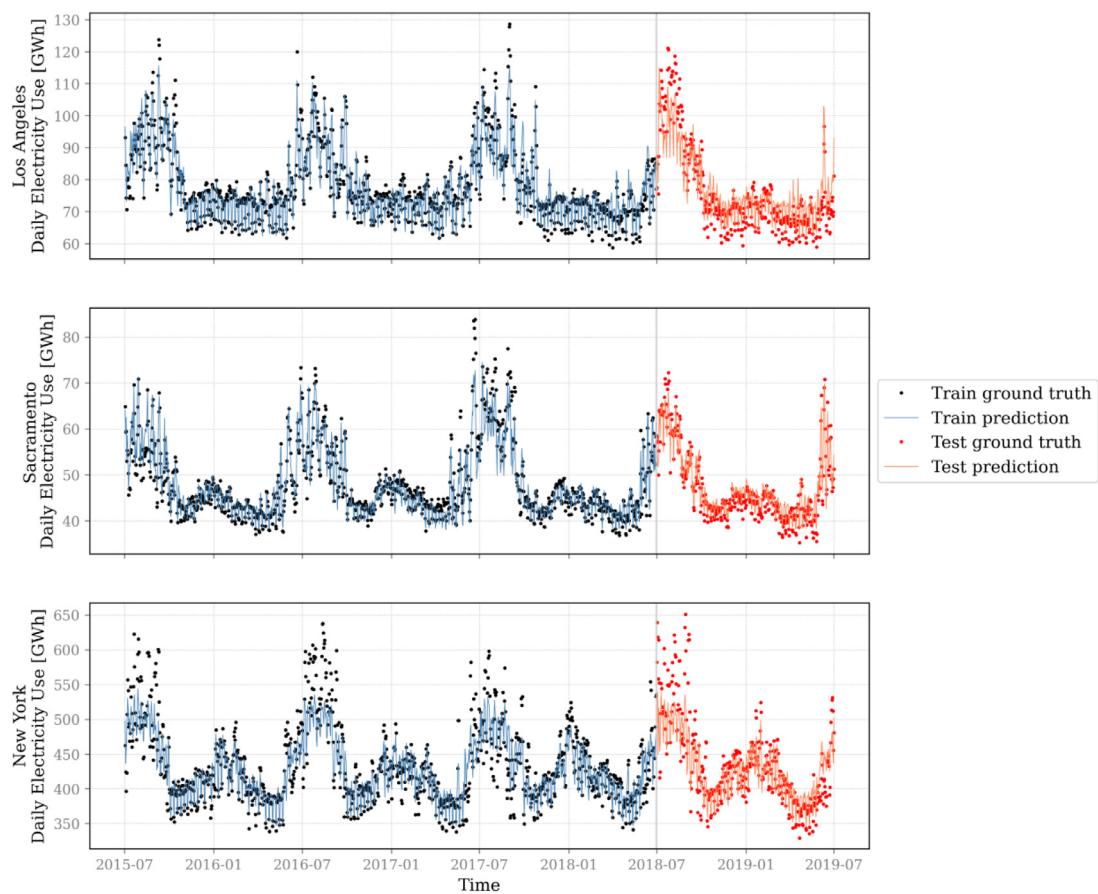


Fig. A2. Predictions from Support Vector Machine and ground truth. From top to bottom – Los Angeles, Sacramento, and New York metropolitan areas.

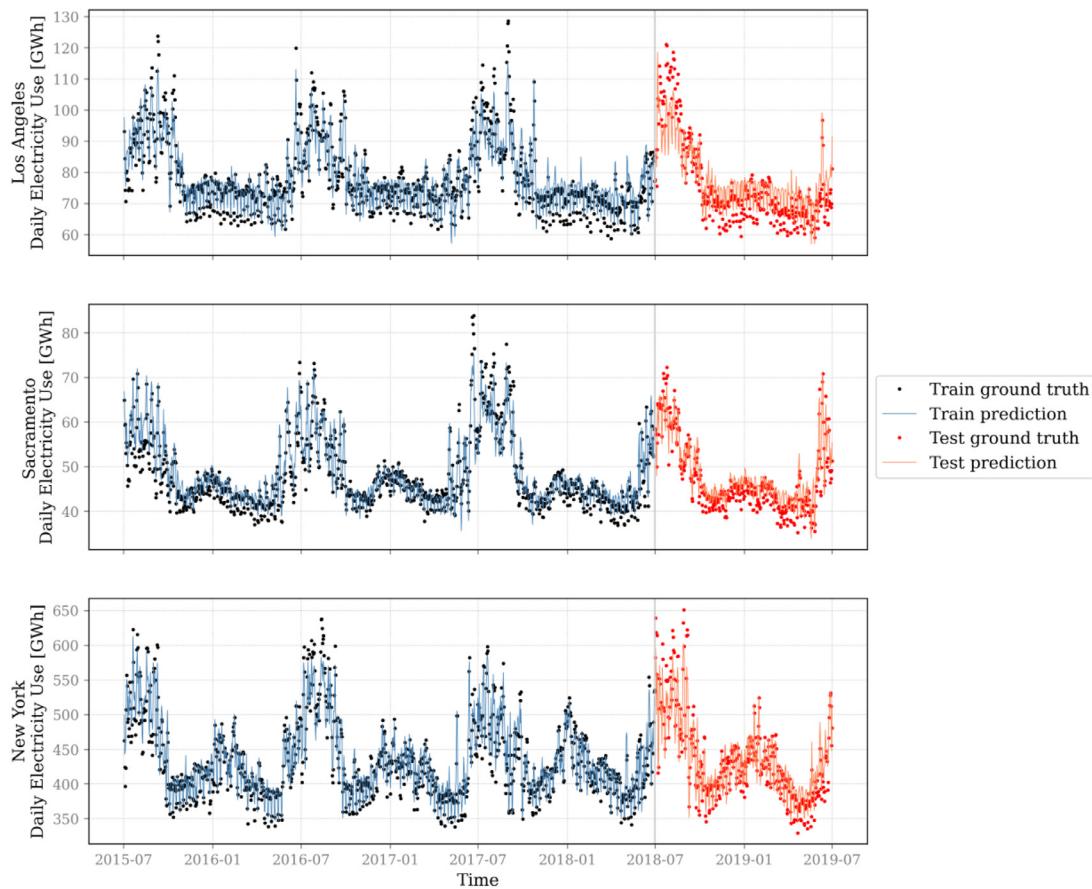


Fig. A3. Predictions from Neural Network and ground truth. From top to bottom – Los Angeles, Sacramento, and New York metropolitan areas.

References

- [1] Kontokosta CE, Tull C. A data-driven predictive model of city-scale energy use in buildings. *Appl Energy* 2017;197:303–17. doi:[10.1016/j.apenergy.2017.04.005](https://doi.org/10.1016/j.apenergy.2017.04.005).
- [2] Deschênes O, Greenstone M. Climate Change, Mortality, and Adaptation: evidence from Annual Fluctuations in Weather in the US. *Am Econ J Appl Econ* 2011;3(4):152–85. doi:[10.1257/app.3.4.152](https://doi.org/10.1257/app.3.4.152).
- [3] Vázquez-Canteli JR, Ulyanov S, Kämpf J, Nagy Z. Fusing TensorFlow with building energy simulation for intelligent energy management in smart cities. *Sustain Cities Soc* 2019;45:243–57. doi:[10.1016/j.scs.2018.11.021](https://doi.org/10.1016/j.scs.2018.11.021).
- [4] Davoudi S, Brooks E, Mehmood A. Evolutionary resilience and strategies for climate adaptation. *Plan Pract Res* 2013;28(3):307–22. doi:[10.1080/02697459.2013.787695](https://doi.org/10.1080/02697459.2013.787695).
- [5] Yi-Ling H, Hai-Zhen M, Guang-Tao D, Jun S. Influences of urban temperature on the electricity consumption of Shanghai. *Adv Clim Change Res* 2014;5(2):74–80. doi:[10.3724/SP.J.1248.2014.074](https://doi.org/10.3724/SP.J.1248.2014.074).
- [6] Sathaye JA, Dale LL, Larsen PH, Fitts GA. Estimating impacts of warming temperatures on California's electricity system. *Glob Environ Change* 2013;23(2):499–511. doi:[10.1016/j.gloenvcha.2012.12.005](https://doi.org/10.1016/j.gloenvcha.2012.12.005).
- [7] "Joint Response to Governor Newsom Letter August192020.pdf." Accessed: 2020. [Online]. Available: https://www.cpuc.ca.gov/uploadedFiles/CPUCWebsite/Content/News_Room/NewsUpdates/2020/Joint%20Response%20to%20Governor%20Newsom%20Letter%20August192020.pdf.
- [8] Carvalho M, de M Delgado DB, de Lima KM, de C Cancela M, dos Siqueira CA, de Souza DLB. Effects of the COVID-19 pandemic on the Brazilian electricity consumption patterns. *Int J Energy Res* 2021;45(2):3358–64. doi:[10.1002/er.5877](https://doi.org/10.1002/er.5877).
- [9] Bahmanyar A, Estebsari A, Ernst D. The impact of different COVID-19 containment measures on electricity consumption in Europe. *Energy Res Soc Sci* 2020;68:101683. doi:[10.1016/j.erss.2020.101683](https://doi.org/10.1016/j.erss.2020.101683).
- [10] Janzen B, Radulescu D. Electricity use as a real-time indicator of the economic Burden of the COVID-19-Related Lockdown: evidence from Switzerland. *CESifo Econ. Stud.* 2020;66(4):303–21. doi:[10.1093/cesifo/ifa010](https://doi.org/10.1093/cesifo/ifa010).
- [11] Lim H, Zhai ZJ. Review on stochastic modeling methods for building stock energy prediction. *Build Simul* 2017;10(5):607–24. doi:[10.1007/s12273-017-0383-y](https://doi.org/10.1007/s12273-017-0383-y).
- [12] Swan LG, Ugursal VI. Modeling of end-use energy consumption in the residential sector: a review of modeling techniques. *Renew Sustain Energy Rev* 2009;13(8):1819–35. doi:[10.1016/j.rser.2008.09.033](https://doi.org/10.1016/j.rser.2008.09.033).
- [13] Bianco V, Manca O, Nardini S. Electricity consumption forecasting in Italy using linear regression models. *Energy* 2009;34(9):1413–21. doi:[10.1016/j.energy.2009.06.034](https://doi.org/10.1016/j.energy.2009.06.034).
- [14] Shorrocks L, Dunster J. The physically-based model BREHOMES and its use in deriving scenarios for the energy use and carbon dioxide emissions of the UK housing stock. *Energy Policy* 1997;25(12):1027–37. doi:[10.1016/S0301-4215\(97\)00130-4](https://doi.org/10.1016/S0301-4215(97)00130-4).
- [15] Wang Z, Zhao Z, Lin B, Zhu Y, Ouyang Q. Residential heating energy consumption modeling through a bottom-up approach for China's Hot Summer–Cold Winter climatic region. *Energy Build* 2015;109:65–74. doi:[10.1016/j.enbuild.2015.09.057](https://doi.org/10.1016/j.enbuild.2015.09.057).
- [16] Chen Y, Hong T, Piette MA. Automatic generation and simulation of urban building energy models based on city datasets for city-scale building retrofit analysis. *Appl Energy* 2017;205:323–35. doi:[10.1016/j.apenergy.2017.07.128](https://doi.org/10.1016/j.apenergy.2017.07.128).
- [17] Abbasabadi N, Ashayeri M. Urban energy use modeling methods and tools: a review and an outlook. *Build Environ* 2019;161:106270. doi:[10.1016/j.buildenv.2019.106270](https://doi.org/10.1016/j.buildenv.2019.106270).
- [18] Lindsey M, Schofer JL, Durango-Cohen P, Gray KA. The effect of residential location on vehicle miles of travel, energy consumption and greenhouse gas emissions: Chicago case study. *Transp Res Part Transp Environ* 2011;16(1):1–9. doi:[10.1016/j.trd.2010.08.004](https://doi.org/10.1016/j.trd.2010.08.004).
- [19] Kuusela P, Norros I, Weiss R, Sorasalmi T. Practical lognormal framework for household energy consumption modeling. *Energy Build* 2015;108:223–35. doi:[10.1016/j.enbuild.2015.09.008](https://doi.org/10.1016/j.enbuild.2015.09.008).
- [20] Park SK, Moon HJ, Min KC, Hwang C, Kim S. Application of a multiple linear regression and an artificial neural network model for the heating performance analysis and hourly prediction of a large-scale ground source heat pump system. *Energy Build* 2018;165:206–15. doi:[10.1016/j.enbuild.2018.01.029](https://doi.org/10.1016/j.enbuild.2018.01.029).
- [21] Olden JD, Joy MK, DeAngelis RG. An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecol Model* 2004;178(3):389–97. doi:[10.1016/j.ecolmodel.2004.03.013](https://doi.org/10.1016/j.ecolmodel.2004.03.013).
- [22] J.K. Kissack, J.S. Haberl, and D.E. Claridge, "Development of a toolkit for calculating linear, change-point linear and multiple-linear inverse building energy analysis models, ASHRAE Research Project 1050-RP, final report," Energy Systems Laboratory, Texas A&M University, Technical Report, Nov. 2002. Accessed: Jan. 19, 2021. [Online]. Available: <https://oaktrust.library.tamu.edu/handle/1969.1/2847>.
- [23] Zhang Y, O'Neill Z, Dong B, Augenbroe G. Comparisons of inverse modeling approaches for predicting building energy performance. *Build Environ* 2015;86:177–90. doi:[10.1016/j.enbuild.2014.12.023](https://doi.org/10.1016/j.enbuild.2014.12.023).
- [24] Geng Y, Ji W, Lin B, Hong J, Zhu Y. Building energy performance diagnosis using energy bills and weather data. *Energy Build* 2018;172:181–91. doi:[10.1016/j.enbuild.2018.04.047](https://doi.org/10.1016/j.enbuild.2018.04.047).
- [25] Li H, Bekhit A, Szum C, Nesler C, Lisauskas S, Snyder SC. Targeting building energy

- efficiency opportunities: an open-source analytical & benchmarking tool. *ASHRAE Trans* 2019;125:470–8.
- [26] Büyükalaca O, Bulut H, Yilmaz T. Analysis of variable-base heating and cooling degree-days for Turkey. *Appl Energy* 2001;69(4):269–83. doi:[10.1016/S0306-2619\(01\)00017-4](https://doi.org/10.1016/S0306-2619(01)00017-4).
- [27] Chandramouli SN, Felder FA. Impact of climate change on electricity systems and markets – A review of models and forecasts. *Sustain Energy Technol Assess* 2014;5:62–74. doi:[10.1016/j.seta.2013.11.003](https://doi.org/10.1016/j.seta.2013.11.003).
- [28] D'Amico A, Ciulla G, Panno D, Ferrari S. Building energy demand assessment through heating degree days: the importance of a climatic dataset. *Appl Energy* 2019;242:1285–306. doi:[10.1016/j.apenergy.2019.03.167](https://doi.org/10.1016/j.apenergy.2019.03.167).
- [29] Bolattürk A. Optimum insulation thicknesses for building walls with respect to cooling and heating degree-hours in the warmest zone of Turkey. *Build Environ* 2008;43(6):1055–64. doi:[10.1016/j.buildenv.2007.02.014](https://doi.org/10.1016/j.buildenv.2007.02.014).
- [30] Roshan GhR, Ghangerneh AA, Attia S. Determining new threshold temperatures for cooling and heating degree day index of different climatic zones of Iran. *Renew Energy* 2017;101:156–67. doi:[10.1016/j.renene.2016.08.053](https://doi.org/10.1016/j.renene.2016.08.053).
- [31] “scipy.optimize.curve_fit” — SciPy v1.5.2 Reference Guide.” https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.curve_fit.html (accessed 2020).
- [32] De Gooijer JG, Hyndman RJ. 25 years of time series forecasting. *Int J Forecast* 2006;22(3):443–73. doi:[10.1016/j.ijforecast.2006.01.001](https://doi.org/10.1016/j.ijforecast.2006.01.001).
- [33] Erdogdu E. Natural gas demand in Turkey. *Appl Energy* 2010;87(1):211–19. doi:[10.1016/j.apenergy.2009.07.006](https://doi.org/10.1016/j.apenergy.2009.07.006).
- [34] Saab S, Badr E, Nasr G. Univariate modeling and forecasting of energy consumption: the case of electricity in Lebanon. *Energy* 2001;26(1):1–14. doi:[10.1016/S0360-5442\(00\)00049-9](https://doi.org/10.1016/S0360-5442(00)00049-9).
- [35] Sp Pappas S, Economou I, Ch Karamousantas D, Chatzarakis GE, Katsikas SK, Liatsis P. Electricity demand loads modeling using AutoRegressive Moving Average (ARMA) models. *Energy* 2008;33(9):1353–60. doi:[10.1016/j.energy.2008.05.008](https://doi.org/10.1016/j.energy.2008.05.008).
- [36] Conejo AJ, Plazas MA, Espinola R, Molina AB. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. *IEEE Trans Power Syst* 2005;20(2):1035–42. doi:[10.1109/TPWRS.2005.846054](https://doi.org/10.1109/TPWRS.2005.846054).
- [37] Suganthi L, Samuel AA. Energy models for demand forecasting—A review. *Renew Sustain Energy Rev* 2012;16(2):1223–40. doi:[10.1016/j.rser.2011.08.014](https://doi.org/10.1016/j.rser.2011.08.014).
- [38] Rahman A, Srikumar V, Smith AD. Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. *Appl Energy* 2018;212:372–85. doi:[10.1016/j.apenergy.2017.12.051](https://doi.org/10.1016/j.apenergy.2017.12.051).
- [39] Wang W, Hong T, Xu X, Chen J, Liu Z, Xu N. Forecasting district-scale energy dynamics through integrating building network and long short-term memory learning algorithm. *Appl Energy* 2019;248:217–30. doi:[10.1016/j.apenergy.2019.04.085](https://doi.org/10.1016/j.apenergy.2019.04.085).
- [40] Wang Z, Hong T, Piette MA. Predicting plug loads with occupant count data through a deep learning approach. *Energy* 2019;181:29–42. doi:[10.1016/j.energy.2019.05.138](https://doi.org/10.1016/j.energy.2019.05.138).
- [41] “Prophet,” *Prophet*. <http://facebook.github.io/prophet/> (accessed 2020).
- [42] Taylor SJ, Letham B. Forecasting at scale. *Am. Stat.* 2018;72(1):37–45. doi:[10.1080/00031305.2017.1380080](https://doi.org/10.1080/00031305.2017.1380080).
- [43] Asha J, Rishidas S, SanthoshKumar S, Reena P. Analysis of temperature prediction using random forest and Facebook Prophet Algorithms.”. *Innov Data Commun Technol Application* 2020;432–9. doi:[10.1007/978-3-030-38040-3_49](https://doi.org/10.1007/978-3-030-38040-3_49).
- [44] Fang W-X, Lan P-C, Lin W-R, Chang H-C, Chang H-Y, Wang Y-H. Combine Facebook Prophet and LSTM with BPNN Forecasting financial markets: the Morgan Taiwan Index. In: Proceedings of the 2019 International symposium on intelligent signal processing and communication systems (ISPACS); 2019. p. 1–2. doi:[10.1109/IS-PACS48206.2019.8986377](https://doi.org/10.1109/IS-PACS48206.2019.8986377).
- [45] Mahmud S. Bangladesh COVID-19 daily cases time series analysis using Facebook Prophet Model. Social science research network. Rochester, NY: SSRN Scholarly Paper ID 3660368; 2020. doi:[10.2139/ssrn.3660368](https://doi.org/10.2139/ssrn.3660368).
- [46] Fernández I, Borges CE, Penya YK. Efficient building load forecasting. in *ETFA2011* 2011:1–8. doi:[10.1109/ETFA.2011.6059103](https://doi.org/10.1109/ETFA.2011.6059103).
- [47] Tso GKF, Yau KKW. Predicting electricity energy consumption: a comparison of regression analysis, decision tree and neural networks. *Energy* 2007;32(9):1761–8. doi:[10.1016/j.energy.2006.11.010](https://doi.org/10.1016/j.energy.2006.11.010).
- [48] J. Roth, A. Bailey, S. Choudhary, and R.K. Jain, “Spatial and temporal modeling of urban building energy consumption using machine learning and open data” pp. 459–67, 2019, doi: [10.1061/9780784482445.059](https://doi.org/10.1061/9780784482445.059).
- [49] Li Y, Wen Z, Cao Y, Tan Y, Sidorov D, Panasetsky D. A combined forecasting approach with model self-adjustment for renewable generations and energy loads in smart community. *Energy* 2017;129:216–27. doi:[10.1016/j.energy.2017.04.032](https://doi.org/10.1016/j.energy.2017.04.032).
- [50] Al-Qahtani FH, Crone SF. Multivariate k-nearest neighbour regression for time series data — A novel algorithm for forecasting UK electricity demand. In: Proceedings of the 2013 international joint conference on neural networks (IJCNN); 2013. p. 1–8. doi:[10.1109/IJCNN.2013.6706742](https://doi.org/10.1109/IJCNN.2013.6706742).
- [51] Fonseca JA, Schlüter A. Integrated model for characterization of spatiotemporal building energy consumption patterns in neighborhoods and city districts. *Appl Energy* 2015;142:247–65. doi:[10.1016/j.apenergy.2014.12.068](https://doi.org/10.1016/j.apenergy.2014.12.068).
- [52] Kontokosta CE, Tull C. A data-driven predictive model of city-scale energy use in buildings. *Appl Energy* 2017;197:303–17. doi:[10.1016/j.apenergy.2017.04.005](https://doi.org/10.1016/j.apenergy.2017.04.005).
- [53] Miller C, et al. The ASHRAE Great Energy Predictor III competition: overview and results. *Sci Technol Built Environ* 2020;0(0):1–21. doi:[10.1080/23744731.2020.1795514](https://doi.org/10.1080/23744731.2020.1795514).
- [54] Wang Z, Hong T, Piette MA. Building thermal load prediction through shallow machine learning and deep learning. *Appl Energy* 2020;263:114683. doi:[10.1016/j.apenergy.2020.114683](https://doi.org/10.1016/j.apenergy.2020.114683).
- [55] “Welcome to LightGBM’s documentation! — LightGBM 3.0.0.99 documentation.” <https://lightgbm.readthedocs.io/en/latest/> (accessed 2020).
- [56] Ruggles TH, Farnham DJ, Tong D, Caldeira K. Developing reliable hourly electricity demand data through screening and imputation. *Sci. Data* 2020;7(1) Art. no. 1. doi:[10.1038/s41597-020-0483-x](https://doi.org/10.1038/s41597-020-0483-x).
- [57] ASHRAE, “ANSI/ASHRAE/IES Standard 90.1-2019 – Energy Standard for Buildings Except Low-Rise Residential Buildings” 2019.
- [58] U.S. Energy Information Administration, “About the EIA-930 data.” 2019, [Online]. Available: <https://www.eia.gov/beta/electricity/gridmonitor/about>.
- [59] ANSI/ASHRAE/IES Standard 90.1, “Energy standard for building except low-rise residential buildings” 2019, Accessed: Oct. 04, 2020. [Online]. Available: <https://www.ashrae.org/technical-resources/bookstore/standard-90-1>.
- [60] Meehl GA, Tebaldi C. More intense, more frequent, and longer lasting heat waves in the 21st century. *Science* 2004;305(5686):994–7. doi:[10.1126/science.1098704](https://doi.org/10.1126/science.1098704).
- [61] Auffhammer M, Mansur ET. Measuring climatic impacts on energy consumption: a review of the empirical literature. *Energy Econ* 2014;46:522–30. doi:[10.1016/j.eneco.2014.04.017](https://doi.org/10.1016/j.eneco.2014.04.017).