

ON THE LONG-TERM DENSITY PREDICTION OF PEAK ELECTRICITY LOAD WITH DEMAND SIDE MANAGEMENT IN BUILDINGS

Youngchan Jang

Industrial and Operations Engineering
University of Michigan
Ann Arbor, MI 48105
mapsossa@umich.edu

Eunshin Byon*

Industrial and Operations Engineering
University of Michigan
Ann Arbor, MI 48105
ebyon@umich.edu

Elham Jahani

Civil, Construction and Environmental Engineering
Iowa State University
Ames, IA 50011
ejahani@iastate.edu

Kristen Cetin

Civil and Environmental Engineering
Michigan State University
East Lansing, MI 48824
cetinkri@msu.edu

ABSTRACT

Long-term daily peak demand forecast plays an important role in the effective and economic operations and planning of power systems. However, many uncertainties and building demand variability, which are associated with climate and socio-economic changes, complicate demand forecasting and expose power system operators to the risk of failing to meet electricity demand. This study presents a new approach to provide the long-term density prediction of the daily peak demand. Specifically, we make use of temperature projections from physics-based global climate models and calibrate the projections to address possible biases. In addition, the effects of population growth and demand side management efforts in buildings are taken into consideration. Finally, the daily peak demands are modeled with the nonhomogeneous generalized extreme value distribution where the parameters are allowed to vary, depending on the predicted temperature and population. A case study using actual building use data in the south-central region in Texas demonstrates that the proposed approach can quantify the uncertainties in an integrative framework and provide useful insights into the long-term evolution of peak demand density. A well-established building demand saving strategy is predicted to buffer against the growing needs of long-term peak electricity demand.

Keywords Climate change, building demand saving, uncertainty quantification

*Corresponding author

1 Introduction

Accurate electricity load forecasting is critical for reliable operations and long-term planning of the electric power grid and its infrastructure systems [1]. Such load forecasting methods help to support the cost-efficient scheduling of energy-producing resources, as well as decisions for the construction of new and upgrading of existing electric grid components. These both lead to reliable availability of power, which is critical to today's modern and highly electricity-dependent society.

In the operations and planning of electricity systems, load forecasting can be conducted on different time horizons, such as short-term [2, 3, 4, 5], medium-term [5, 6, 7], and long-term forecasts [6, 7], depending on the intended goals. Among them, short-term load forecasts, which focus on a time horizon of less than 24 hours, impact day-to-day operations of the electric grid and its transactions [8]. Medium-term, typically weeks- or months-ahead, forecasts are often used for negotiating energy and electricity-related contracts [9]. Long-term load forecasting is generally associated with a time horizon of more than a year, up to several decades [10], which plays an important role for the generation, transmission and distribution system planning. The medium- and long-term demand forecasts mainly focus on predicting the average or peak loads, while the short-term demand forecast often aims to provide hourly or sub-hourly load forecasts.

This study is concerned with the long-term daily peak demand forecasting. In many cases, utilities and grid operators consider short and medium-term load predictions for their planning purposes. However, in the face of changes in climate conditions and other socio-economic factors, it is important to consider the potential impacts of such changes on the power grid load over time. In particular, this study focuses on peak electricity demands, as peak demands often define the required capacity of generation and transmission systems [11, 12]. The purpose of this study is to evaluate the potential impacts of predicted climate change of a particular region on this peak load, rather than only consider historical weather and prior peak loads in the prediction. This provides, as opposed to the short-term prediction, a longer term evaluation to understand, over longer timescales, what potential needs for grid infrastructure planning are needed, and/or what types of peak load reduction or grid service participation development might be needed at longer time horizons, to mitigate any resource adequacy issues that may arise.

Electric loads depend on a range of factors, including weather conditions and socio-economic factors. Among the factors, it has been reported that electricity demand is significantly impacted by ambient temperature [10, 13]. A recent review on the relationship between electricity demand and weather conditions [14] also supports that outdoor temperature is a crucial factor for load forecasting. This relationship exists because during the cooling season, for a cooling-dominated climate (e.g. ASHRAE Climate Zones 1-3), close to 100% of residential and commercial buildings rely on electricity-powered cooling from the heating, ventilation and air conditioning (HVAC) systems in the U.S [15, 16]. A similar relationship exists in the heating season, however, since many buildings, particularly in colder climates, use gas or other non-electricity based heating fuels, this relationship is not typically as pronounced.

For the short-term and medium-term load forecasting, considering that the load is largely correlated with recent loads under similar weather conditions, historical data are often employed [13]. Recent literature

suggests, however, that the historical weather trends and temperature extremes change over time [17]. The Intergovernmental Panel on Climate Change (IPCC) publishes reports every several years which include projections of long-term changes in temperature extremes. The Third Assessment Report [18] concluded that there very likely had been an increase in the frequency of extremely high temperatures. The IPCC Special Report on Extremes in 2012 [19] and the IPCC Fifth and Sixth Assessment Reports in 2013 and 2018 [20, 21], respectively, have made even firmer statements. Such climate changes have strong implications for the electric grid, including peak demands, generation efficiency and availability, and transmission and distribution congestion and capacities. To address the varying weather conditions in the long term, global climate model (GCM) projections [22] can be considered. The GCMs are mathematical representations of the earth's climate components and their interactions, including the atmosphere, land, ocean, and sea ice, that are simulated over periods of time to project future weather conditions. A detailed description about GCMs is available in [17].

Along with weather conditions, electricity peak loads also largely depend on the socio-economic factors, including the population size and buildings' electricity use patterns [23]. To reduce the peak load during the extreme heat of summer when the electricity usage typically is at highest, a range of demand side management (DSM) programs have been developed, piloted and used in recent years. DSM includes the building demand reduction measures such as energy efficiency (EE) and demand response (DR) programs. While EE programs aim to reduce the electricity demand in general, DR programs mainly focus on the buildings' peak demand reduction by modifying the end-use electricity consumption patterns and changing the timing and level of instantaneous demand [24].

A broad range of DSM programs for residential and commercial buildings are currently implemented throughout the United States, typically run by utility companies and third party aggregators [25], where end-use customers receive an incentive and/or other monetary or non-monetary benefits by participating in DSM programs. Such incentives help electric utilities and power network companies to maintain a predictable level of demand adjustments that can be made to support the reliable operation of the electricity system. Although many DSM efforts for buildings are still in the pilot stages, DSM programs are projected to significantly increase moving forward, particularly as the electric grid is increasingly powered by more variable renewable energy sources [26].

The objective of this study is to develop an integrative modeling framework to estimate the long-term daily peak load with DSM efforts in buildings. We collectively use multiple data sources, including GCM projections, actual temperature measurements, population, and participation rates in building DSM programs. Considering the nature of forecasting uncertainties and demand variability associated with socio-economic and climate changes, we provide probability density predictions that allow for the quantification of how the prediction intervals, means and medians would evolve in the long-run.

Specifically, for characterizing the future daily peak temperature uncertainties, we calibrate GCM projections with actual temperature measurements. Although GCMs provide useful information, an actual trend in a specific region may deviate from the GCM projections, because physics-based climate models do not fully account for local, or regional, characteristics [17]. To address this challenge, we adjust the GCM projections

using a parametric approach. In particular, considering that the daily peak temperature represents the block maximum (i.e. a maximum value during a specific interval), we employ the extreme value distribution and assume that actual daily peak temperatures during the summer period follow the generalized extreme value (GEV) distribution. To reflect the influence of climate changes and temporal variations which possibly makes the temperature distribution nonstationary over a period of years, we allow the GEV density parameters to vary, depending on the GCM projections.

Further, to characterize the influence of the socio-economic factors on the peak demand, we analyze the population growth pattern and participation rates in DSM. To quantify the long-term effects of DSM efforts on the demand saving in buildings, we analyze the number of participants in the DSM programs using the Bass diffusion model [27]. We also characterize the population growth pattern using the logistic growth model and incorporate it into the Bass diffusion model.

In summary, the main contributions of this study are as follows. First, we characterize the future temperature uncertainties by calibrating GCM projections with actual measurements using a parametric density model. Second, we analyze the impacts of socio-economic factors (characterized by building DSM efforts and population growth) on the future long-term daily peak electricity demand reduction. Lastly, we provide a united framework for quantifying forecasting uncertainties through a probabilistic modeling approach.

A case study using actual building use data is conducted in the region of Texas that includes the city of Austin, which is located in ASHARE Climate Zone 3a. The electric grid in Texas experiences significant peak demand during the summer periods when high peak temperatures occur. Our implementation results demonstrate that the proposed approach can characterize the nonstationary characteristics of the extreme pattern of the daily peak demand and quantify the forecasting uncertainty and demand variability associated with the climate and population change, and building demand reduction in the future.

The remainder of this article is organized as follows. Section 2 reviews relevant studies in the literature. Section 3 describes the datasets used in this study. Section 4 introduces the proposed long-term daily peak demand density prediction method. Section 5 presents a case study for evaluating the density prediction performance of the proposed approach and providing long-term densities. Finally, Section 6 summarizes the findings, implications, and future research.

2 Literature review

In the literature, there has been substantial progress in developing models for short-term load forecasting. Such statistical models include linear and nonlinear regression models [28, 29], time series methods including autoregressive, autoregressive moving average, autoregressive integrated moving average models and their variations [30, 31, 32]. Different types of neural networks and their variations [4, 33, 34, 35, 36], as well as other machine learning techniques and hybrid or ensemble models [37, 38, 39], have been also studied.

In comparison with the short-term load forecasting, limited research has been conducted on the long-term prediction due to difficulties in quantifying forecasting uncertainties and demand variability. The climate change community has tended to prefer physics-based climate models, including GCMs, resisting

of statistical approaches for long-term forecasting [40, 41]. On the other hand, much of the work studying extreme heat events in the statistical field focuses on analyzing historical temperatures and detecting trends in temperature extremes without considering climate change trends. Below we summarize relevant studies on long-term temperature and load forecasting.

Chen [42] proposes a collaborative fuzzy-neural approach, utilizing multiple expert opinions about the peak or average value of annual demand forecasts. This approach mainly relies on expert opinions, aiming to minimize individual deviations and biases. However, forecasts are made based on individuals' subjective judgments. AlRashidi *et al.* [43] propose linear and quadratic models based on particle swarm optimization for five years-ahead load forecasting. In their study, the annual peak load is formulated as the linear or quadratic functions of time.

Andersen *et al.* [44] identify the relationship between the aggregated hourly electricity consumption and different categories of customers such as households, agriculture, industry, and private and public services. They construct the model for each customer using the calendar effect. The future aggregated hourly consumption for each category of customers is then estimated using the weights calculated by an annual econometric model that considers the effects of the socio-economic factors. Xia *et al.* [45] employ artificial neural networks to provide short- to long-term load forecasts using historical weather data only. They do not take other socio-economic factors into account. Further, these studies [43, 44, 45] provide the point estimates of the long-term peak demands and do not fully address the uncertainty quantification.

Hyndman *et al.* [46] propose the semi-parametric additive model to forecast annual and weekly peak demand densities for the next ten years by regressing half-hour demand on the half-hour temperature, calendar effects, and the annual economic and demographic information. The future economic and demographic scenarios obtained from the Australian Energy Market Operator, as well as temperature simulated by a bootstrap method, are fed into the fitted model for the density forecast. However, they assume that the temperature is stationary for the long-term time horizon, thus they do not characterize the evolving characteristics of temperature caused by climate changes. Hong *et al.* [47] propose the use of multiple linear regression for one year-ahead load forecasting using hourly temperature and annual gross state product (GSP). They generate cross scenarios of future economic scenarios and hourly temperatures for the density prediction, however, a limited number of scenarios (ninety scenarios in total) are considered.

Some studies, based on the extreme value theory, employ nonhomogeneous GEV to characterize the nonstationary trend, where the long-term trend is quantified using historical data [48, 49, 50, 51]. Different from the aforementioned studies that use historical data only, Trotter *et al.* [41] propose the probabilistic long-term electricity demand forecasting using the multiple linear regression with GCM outputs as well as other demographic and economic factors.

In summary, most statistical approaches employ historical data for characterizing the long-term trend. On the other hand, as discussed earlier, meteorologists rely on climate models such as GCMs to predict long-term future temperatures. We believe both historical temperature data and climate models provide useful information; historical data contains information on local (or regional) characteristics, whereas GCMs are based on first principles and expected climate changes. The study by Trotter *et al.* [41] is one of the few

studies that consider both, however it does not address the possible bias of GCM projections. This study fills the research gap in the literature by connecting the GCM temperature projections with the peak load forecasting. Our approach also collectively uses other socio-economic variables for providing future density prediction.

3 Datasets

This section describes the datasets used in this study, including GCM projections, actual temperature measurements, population, and participation rate in the considered DSM program. As the focus of this study is the extreme peak temperature and its impact on electricity demand, we use the temperature and demand data collected in July and August. Therefore, the daily peak temperature and demand each year consist of 62 data points.

The actual daily peak temperature measurements are collected from the Automated Surface Observing System (ASOS) data. This dataset includes weather data from a broad network of weather stations supported by the Federal Aviation Administration and the National Weather Service [52]. We use the temperature data collected at the ATT site which is located in the urban area of Austin, TX, from 2002 to 2016. For the GCM projections, we use data from the downscaled CMIP3 and CMIP5 Climate and Hydrology Projections archive, which is publicly available in [53]. We use 1-degree bias-corrected GCM outputs for daily maximum surface air temperature in Austin, TX, from 42 climate models under the RCP 4.5 scenario. The RCP 4.5 scenario represents mild climate change [20, 21]. The GCM outputs are available from 2006 to 2100. Besides temperature, humidity also has some impacts on electric demand, as latent loads are also addressed by heating and cooling systems in buildings. However, for modeling the long-term peak density, we believe the temperature is a sufficient factor [46, 47].

Next, the daily peak demand is obtained from the Electric Reliability Council of Texas (ERCOT) [54]. Unlike the actual temperature data providing the temperature measurements in a specific location, the demand data is provided in a large spatial domain. ERCOT divides Texas into eight distinct weather zones. Each zone represents a geographic region in which climatological characteristics are similar, and ERCOT provides the total aggregated electricity demand in each region. Thus, we use the daily peak demand data in the south-central region, which includes the city of Austin.

Figure 1 depicts the box plot of daily peak demand during the summer from 2002 to 2016 in the south-central region of Texas, where each box plot shows the density of daily peak demands using 62 data points per year. There is a clear increasing trend over time. We note that the daily peak demand in 2011 is noticeably higher than that in the surrounding years. This is because of the extreme heatwave events that occurred in Texas during this year, which caused an increase in electricity demand due to the increasing usage of air conditioners and cooling appliances [55].

Electricity usage is also largely affected by socio-economic factors, such as gross domestic product (GDP), industrial production, and the population density [23, 46]. Among them, this study uses the population as one of the major factors for the daily peak demand modeling. The population data consists of the yearly

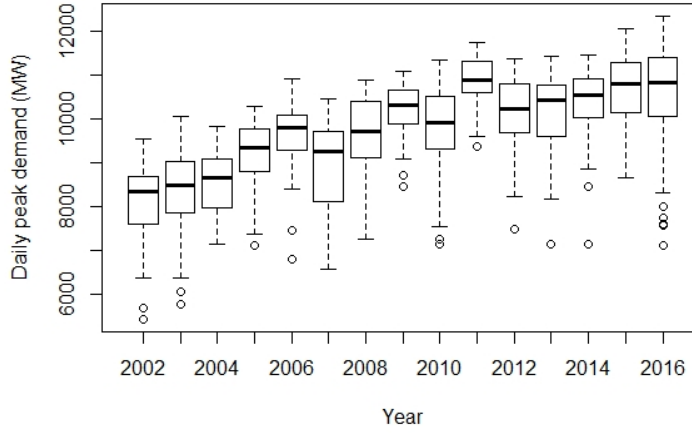


Figure 1: Box plots of actual daily peak demands during the summer from 2002 to 2016 in the south-central region of Texas

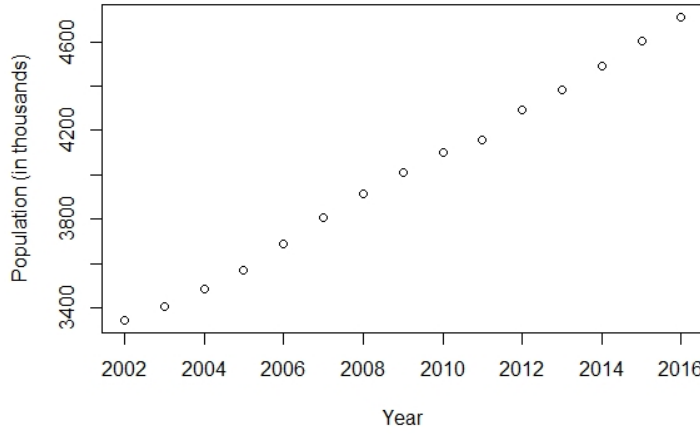


Figure 2: Total population growth in 24 counties in the south-central region of Texas from 2002 to 2016

population estimates for every county in the south-central region of Texas from 2002 to 2016, obtained from the United States Census Bureau [56]. Figure 2 shows the total number of population estimates in thousands from 24 counties in the south-central region of Texas. The x - and y -axis represent year t and the corresponding total population in thousands p_t , respectively. This shows the increasing linear trend of the population over time in this region. It should be noted that other economic factors can be additionally considered in our modeling efforts, however, they are often positively correlated with the population growth (or decay) trends. As such we consider the population to represent the economic condition of the studied area.

Further, due to the increasing interest in DSM efforts for residential and commercial buildings, the demand saving from EE/DR activities should be taken into account in the long-term demand predictions. For example, Austin Energy, the exclusive electricity provider to the city of Austin, operates the EE/DR program

titled the Custom Energy Solutions (CES) program. The participants of the program continue to grow. We use the number of participants and demand saving data reported in the Austin Energy 2017 report [57]. For example, Figure 3 shows the number of residential participants in the DR program in year t and the cumulative number of participants up to year t from 2011 to 2016. Here, the unit of the residential participants in the DR program is the number of smart thermostat devices. Typically, one residential house adopts one device. Although there are year-to-year variations, overall we observe an increasing trend. The original Austin Energy report includes demand saving data since 2007, but we note that the number of residential participants from 2007 to 2010 exhibits a decreasing trend. As the number of participants is expected to grow over years, we use the data from 2011 to 2016 for building the demand saving model. We also use the demand saving data from the community- and municipal-level DR/EE programs, obtained from the Austin Energy report [57] and the U.S. Energy Information Administration (EIA) reports [25].

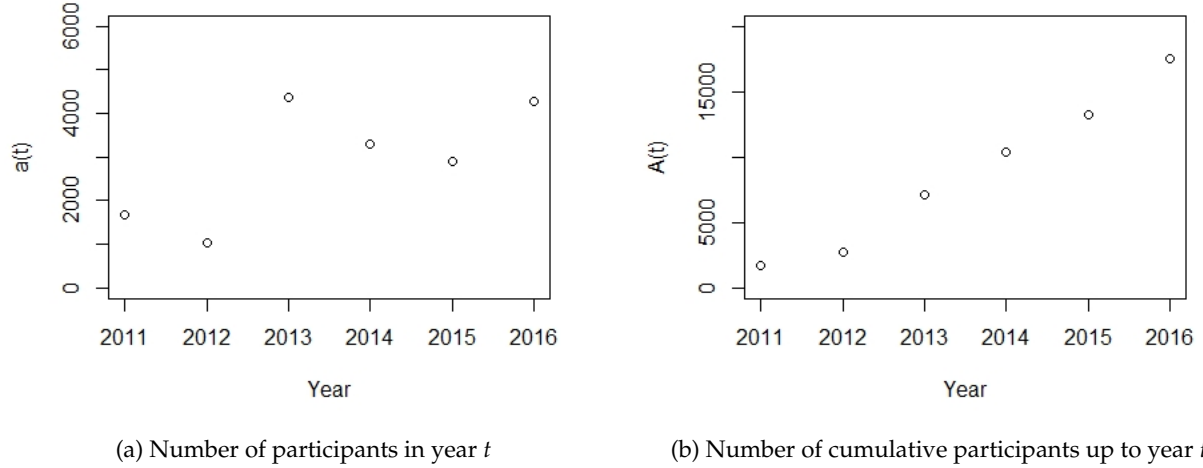


Figure 3: The number of residential participants in Austin Energy's demand saving program [57]

4 Methodology

This section presents the long-term daily peak demand density prediction method. We first formulate the daily peak temperature with the nonhomogeneous GEV model. To incorporate the influence of climate changes on the future temperature, we parameterize the GEV density parameters as functions of GCM projections. Such parameterization also enables us to calibrate the GCM projections with actual data. Next, the socio-economic factors that include the population growth and building DSM efforts are, respectively, modeled by the logistic growth model and the Bass diffusion model. Finally, the daily peak demand density is obtained by integrating the forecasts of future temperature, population, and demand saving by DSM efforts. Figure 4 shows the overall framework of the proposed approach.

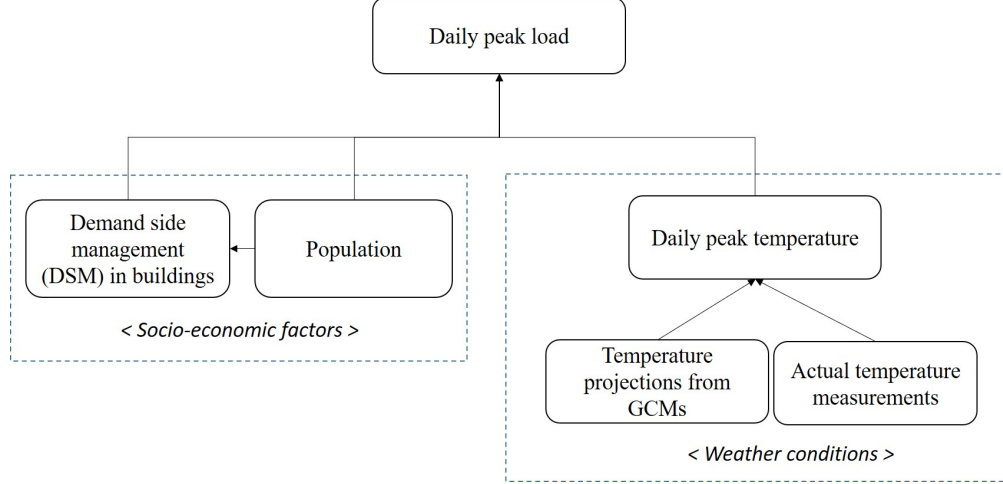


Figure 4: Overall framework of the proposed approach

4.1 Modeling long-term daily peak temperature

The daily peak temperatures represent the block maxima. Based on the limit theorem for block maxima [58, 59], we employ the GEV distributions. The GEV distribution is a family of continuous probability distributions that combines the Gumbel, Fréchet and Weibull distributions. Specifically, we formulate the density of daily peak temperature $y_{d,t}$ in year t with a GEV distribution as

$$\begin{aligned}
 y_{d,t} &\sim f_t(y|x) \\
 &= GEV(\mu_t, \sigma_t, \eta_t) \\
 &= \begin{cases} \frac{1}{\sigma_t} \left[1 + \eta_t \left(\frac{y - \mu_t}{\sigma_t} \right) \right]^{-(1/\eta_t + 1)} \exp \left\{ - \left[1 + \eta_t \left(\frac{y - \mu_t}{\sigma_t} \right) \right]^{-1/\eta_t} \right\}, & \text{for } \eta_t \neq 0 \\ \frac{1}{\sigma_t} \exp \left\{ - \left(\frac{y - \mu_t}{\sigma_t} \right) \right\} - \exp \left\{ - \left(\frac{y - \mu_t}{\sigma_t} \right) \right\}, & \text{for } \eta_t = 0 \end{cases} \quad (1)
 \end{aligned}$$

for $\{y : 1 + \eta_t(y - \mu_t)/\sigma_t > 0\}$, where $\mu_t \in \mathbb{R}$, $\sigma_t > 0$, $\eta_t \in \mathbb{R}$ are the location, scale, and shape parameters, respectively [60]. The location parameter μ_t affects the central value of the density. The scale parameter σ_t quantifies the spread of the distribution, and the shape parameter η_t controls the weight of the distribution tail. Depending on the shape parameter η_t , the GEV distribution is categorized into Gumbel, Fréchet and Weibull distributions [60].

To capture the long-term yearly varying temperature characteristics affected by climate changes, we parameterize the temperature distribution using the GCM outputs. We use 42 different GCM models. Let $g_{d,t}^{(j)}$ denote the daily peak temperature projection on day d in year t from the j^{th} GCM model. Because the GCM model focuses on the long-term projection, the daily variability in each GCM model is incoherent. Rather, an average across 62 days during July and August would represent the overall climate change influence on the summer temperature. As such, we take the grand ensemble average of $g_{d,t}^{(j)}$'s to quantify the forced temperature change caused by climate change. We define the grand ensemble average x_t as

$$x_t = \frac{1}{62 \times 42} \sum_{d=1}^{62} \sum_{j=1}^{42} g_{d,t}^{(j)}. \quad (2)$$

Then we parameterize the location and scale parameters as linear functions of x_t and use a constant value for the shape parameter in order to avoid an overly complicated model [60], as follows.

$$\begin{aligned}\mu_t &= \alpha_0 + \alpha_1 x_t, \\ \sigma_t &= \beta_0 + \beta_1 x_t, \\ \eta_t &= \eta_0,\end{aligned}\tag{3}$$

where $\alpha_0, \alpha_1, \beta_0, \beta_1$, and η_0 becomes the density parameters that need to be estimated with actual temperature measurements.

The nonhomogeneous GEV temperature formulation with the GCM outputs has several important implications. First, as discussed earlier, it incorporates the possible temperature changes due to climate change into the long-term forecasts. Second, even the downscaled 1-degree bias-corrected GCM models do not account for local/regional characteristics. The formulation in (3) enables us to correct the inherent systematic bias and discrepancy with actual data. Third, the GEV density function quantifies prediction uncertainties, whereas original GCM projections provide deterministic forecasting.

We estimate the model parameters using the maximum likelihood estimation (MLE). Specifically, we maximize the log-likelihood function $\ell(\theta; \mathcal{D}_y)$ where $\mathcal{D}_y$ implies a dataset with the measured daily peak temperature, i.e., $\mathcal{D}_y = \{y_{d,t}, d = 1, 2, \dots, D, t = 1, 2, \dots, T\}$. Let $\theta = [\alpha_0, \alpha_1, \beta_0, \beta_1, \eta_0]$ denote the parameter vector. Assuming $y_{d,t}$'s are independently distributed, we obtain its MLE estimates $\hat{\theta}_{MLE}$ as follows.

$$\begin{aligned}\hat{\theta}_{MLE} &= \arg \max_{\theta} \ell(\theta; \mathcal{D}_y) \\ &= \arg \max_{\theta} \sum_{t=1}^T \sum_{d=1}^D \log(f_t(y_{d,t}; \theta)) \\ &= \arg \max_{\theta} \left\{ -D \sum_{t=1}^T \log \sigma_t - \sum_{t=1}^T \sum_{d=1}^D \left(1 + 1/\eta_t \right) \log \left[1 + \eta_t \left(\frac{y_{d,t} - \mu_t}{\sigma_t} \right) \right] - \sum_{t=1}^T \sum_{d=1}^D \left[1 + \eta_t \left(\frac{y_{d,t} - \mu_t}{\sigma_t} \right) \right]^{-1/\eta_t} \right\},\end{aligned}$$

for $\eta_t \neq 0$, where μ_t, σ_t , and η_t are formulated in (3), $D (= 62)$ is the number of summer days each year, and T is the number of years in the data used for getting the MLE estimates. Since there is no analytical solution, we obtain $\hat{\theta}_{MLE}$ numerically. In our analysis, we use 'ismev' package in the statistical software, R, for solving the optimization problem [59]. Similarly, we also obtain the MLE estimates for $\eta_t = 0$ and between two, we choose the estimates that provides higher log-likelihood values.

4.2 Modeling long-term socio-economic pattern

This section discusses the modeling of socio-economic factors, including population growth and building DSM efforts. We note that the population is the representative factor among many possible factors and it is often positively correlated with other economic conditions in the developed countries [61, 62]. Moreover, it is relatively easier to predict the population growth (or decay) than other factors at the local or regional level.

In the studied south-central region of Texas, clearly the population has been linearly increasing over years, as we observed in Fig. 2. However, it is unrealistic to assume the same growth rate for the long-term future.

Furthermore, the U.S. Census Bureau predicts the population growth rate will likely decrease over longer time scale [63]. Thus, we expect that the population would continue to grow, but at a slower rate farther in the future. To represent such growth pattern, we adopt the logistic growth model that formulates an increasing trend yet at a slower rate until it reaches to the certain limit [64] as

$$p_t \sim g_t(p) = N\left(\frac{a}{1 + \exp\{-(t-b)/c\}}, \sigma_p^2\right), \quad (4)$$

where a is the maximum population, b is the point where the growth rate turns from increase to decrease, and c is the logistic growth rate, which controls the steepness of the curve. The error is assumed to be normally distributed with zero mean and constant variance σ_p^2 . Although the logistic growth model is typically used for characterizing growth pattern, it is flexible enough to represent the decreasing pattern as well. When c in (4) is positive, the logistic growth model shows an increasing pattern, whereas it exhibits a decay pattern with the negative c value.

To estimate the parameters in (4), we use Levenberg-Marquardt nonlinear least-squares algorithm [65, 66]. However, the resulting a is around 10,000, implying the maximum population is 10 million in south-central Texas, which appears too large, considering that the U.S. Census Bureau predicts the population growth rate will likely decrease over longer time scale [63]. This unduly large value was obtained because the sample size is limited to 2002 to 2016 and the south-central Texas population linearly grew during these years. To adjust, we consider the maximum population in the south-Central Texas region would be no more than 6 million and set $a = 6000$ (note that the unit of p_t is in 1000). Then we estimate b and c using the Levenberg-Marquardt algorithm and use the sample variance from the residuals to estimate σ_p^2 .

The predicted population is depicted in Figure 5 where the black circle represents the historical population from 2002 to 2016 and red solid and dashed lines, respectively, represent the point prediction and 90% prediction interval (PI). The result suggests that the population would continue to grow, but the growth rate would decrease gradually over time.

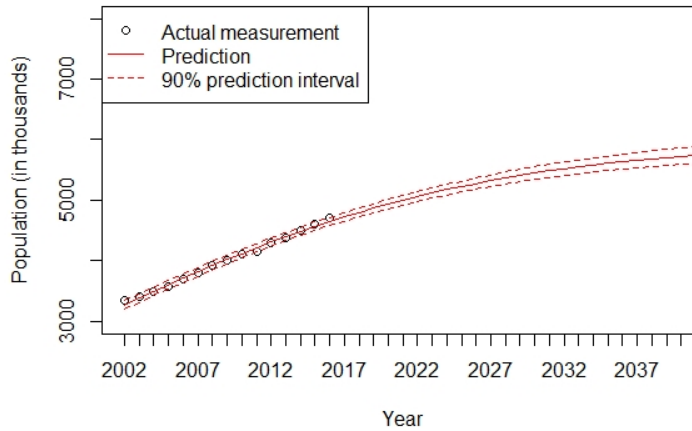


Figure 5: Population prediction in the south-central region of Texas from 2002 to 2040

Next, for predicting future demand saving in buildings, we formulate the participation rates in DSM programs. The challenge is that data is scarce, as many programs are relatively young, with limited historical data. Thus, statistical models, e.g., regression and time series models, are not appropriate. To address this, we employ the Bass diffusion model, which has been widely used for forecasting the sales of new products or adoption of new technologies [27]. The Bass diffusion model describes how the new products or technology can be adopted by investigating a relationship between current adopters and the potential adopters [27]. The Bass diffusion model typically assumes a constant market potential, that is, it is assumed that the maximum number of potential adopters remains the same over time. However, in our case, the maximum number of potential adopters would change over time, as populations change. Thus, we modify the original Bass diffusion model and consider a varying potential market [67].

The DSM participants consist of residential households, commercials, communities and municipals in EE and DR programs, among which we first model the household participant growth pattern in the DR program. Let $M(t)$ be the maximum number of potential participants in year t . As the number of households changes over time, we let $M(t)$ be proportional to the households each year as

$$M(t) = d_s \cdot h_t, \quad (5)$$

where h_t is the number of households in year t and d_s is the maximum portion of total household size that potentially adopts the DR program.

Individual households may participate in the DR program in year t , or they may wait. As the program operates, $M(t)$ potential entities eventually join the DR program. Because it is unrealistic to expect all households joins the program, d_s is typically less than 1. With very limited DR participation data accumulated to date, it is difficult to estimate an appropriate value for d_s . Therefore, in our study, we consider 20% – 50% of the total households would be the maximum number of potential participants, i.e., d_s is assumed to range from 0.2 to 0.5. We perform the sensitivity analysis with different values of d_s from 0.2 to 0.5 with an increment of 0.05 in Section 5.

Let $\ell(t)$ be the portion of the potential entities that participate in the DR program in year t , i.e., $\ell(t) = \frac{a(t)}{M(t)}$, where $a(t)$ is the number of entities that newly participate in year t , and let $L(t)$ be the portion of the entities that have participated up to year t , i.e., $L(t) = \frac{A(t)}{M(t)}$, where $A(t)$ is the cumulative number of participants up to year t . The Bass diffusion model formulates the portion of new participates to non-participants in year t , i.e., $\frac{\ell(t)}{1-L(t)}$, as a linear function of those who had participated [27] as

$$\frac{\ell(t)}{1-L(t)} = m + nL(t), \quad (6)$$

where m is the coefficient of the innovation (or external influence) and n is the imitation among participants (or internal influence).

By multiplying $1 - L(t)$ in both sides in (6), the portion of the participants in year t becomes

$$\ell(t) = \frac{dL(t)}{dt} = (m + nL(t))(1 - L(t)) = m + (n - m)L(t) - nL(t)^2. \quad (7)$$

We solve the nonlinear differential equation with the initial value of $L(0) = 0$ with the fixed value of d_s to estimate the parameters m and n [27] and get

$$L(t) = \frac{1 - \exp\{-(m+n)t\}}{1 + \frac{n}{m} \exp\{-(m+n)t\}}, \quad (8)$$

$$\ell(t) = \frac{\frac{(m+n)^2}{m} \exp\{-(m+n)t\}}{\left(1 + \frac{n}{m} \exp\{-(m+n)t\}\right)^2}, \quad (9)$$

Using the relationship of $\ell(t) = \frac{a(t)}{M(t)}$ and $L(t) = \frac{A(t)}{M(t)}$, we obtain $a(t)$ and $A(t)$ as

$$a(t) = \ell(t)M(t) = \frac{\frac{(m+n)^2}{m} \exp\{-(m+n)t\}}{\left(1 + \frac{n}{m} \exp\{-(m+n)t\}\right)^2} \cdot d_s h_t, \quad (10)$$

and

$$A(t) = L(t)M(t) = \frac{1 - \exp\{-(m+n)t\}}{1 + \frac{n}{m} \exp\{-(m+n)t\}} \cdot d_s h_t, \quad (11)$$

With the resulting $A(t)$, the demand saving in year t , denoted by $s(t)$, can be calculated as

$$s(t) = C \cdot A(t), \quad (12)$$

where C denotes the demand saving per participant (unit: kW).

The parameters m and n in the Bass diffusion model define the participation pattern. To estimate them, we use the Austin Energy's demand saving data from 2011 to 2016 [57]. For C , we employ the average saving per residential participant and get $C = 0.8765 \text{ kW}$. In obtaining the number of households, we collect the population information from the U.S. Census Bureau and use the fact that each household consists of 2.84 people on the average [56].

Ideally we should use the peak demand saving data in other utility companies to cover the entire south-central Texas, as Austin is a part of the area. However, no detailed information from other utility companies is available to us. As a remedy, the residential DR pattern in south-central Texas is assumed to be similar to that in Austin. Specifically, once we estimate m and n with Austin data, we use the same m and n values, but plug the number of households in the south-central area into h_t in (10) and (11). To predict a number of future residential participants, we employ the estimated population p_t in year t , obtained from the logistic growth model in (4), and use $h_t = p_t / 2.84$.

Similar approaches can be applied to the community- and municipal-level DR demand saving projections. However, their potential markets $M(t)$ are not easily quantified and estimated. Therefore, we assume the proportion of community- and municipal-level demand saving to the residential-level saving remains similar in the future. Noting the residential-level demand reduction has been about 35% of the total demand saving from the DR program on average in south-central Texas [57], we obtain the total demand saving from DR programs by multiplying $1/0.35$ to the residential-level demand saving. Likewise, we can obtain the peak

demand saving from EE programs using the Bass diffusion model. However, there are many categories in EE programs and data in each category is scarce. We note that the ratio of the total peak demand saving from EE programs to that from DR programs is about 50% on average in south-central Texas [25]. Based on this fact, we approximate the demand saving from EE programs by multiplying 0.5 to the demand saving from DR programs.

4.3 Daily peak load density prediction

Let $z_{d,t}$ denote the daily peak demand on day d in year t without considering the demand saving. Similar to the daily peak temperature, the daily peak demand in year t is assumed to follow nonhomogeneous GEV distribution and its location and scale parameters are parameterized by the linear function of the population p_t , the daily peak temperature y_t , and year t . Let $z_{d,t}$ denote the daily peak demand on day d in year t without considering the demand saving. Then the conditional density of z_t , given the daily peak temperature and population, is given by

$$z_{d,t}|y_{d,t}, p_t \sim h_t(z|y_{d,t}, p_t) = \text{GEV}(\mu'_{d,t}, \sigma'_{d,t}, \eta'_t), \quad (13)$$

with

$$\begin{aligned} \mu'_{d,t} &= \alpha'_0 + \alpha'_1 y_{d,t} + \alpha'_2 p_t + \alpha'_3 t, \\ \sigma'_{d,t} &= \beta'_0 + \beta'_1 y_{d,t} + \beta'_2 p_t + \beta'_3 t, \\ \eta'_t &= \eta'_0. \end{aligned} \quad (14)$$

An alternative formulation for μ'_t and σ'_t would be to exclude the year term t , while keeping y_t and p_t in (14). However, with the studied datasets from the south-central Texas area, we notice that excluding the year term result in underestimations of the daily peak load. This result indicates that each individual's energy consumption has been increasing over years, possibly due to the escalated dependence on electricity in modern society and increasing trend in the conditioned area per home [68].

Similar to the procedure in (4), the parameters in (14) can be estimated using MLE by maximizing the loglikelihood with the daily peak load measurements $z_{d,t}$. Recall that $z_{d,t}$ is the daily peak load without considering the demand saving. However, the daily peak load data from ERCOT represents the realized peak load with DSM efforts. Thus, we need to recover the daily peak loads without the demand saving by adding the DSM demand saving to ERCOT's reported peak demand data. As no exact daily demand saving data in south-central Texas is available to us, we approximate it using demand saving information provided by the Austin Energy's report [57] and EIA reports [25]. Figure 6 shows the adjusted box plots of daily peak demand, assuming no DSM efforts. In the beginning years, there were no, or very limited, DSM efforts in buildings, so the box plots in those years are similar to the corresponding box plots in Figure 1. As the DSM efforts become more active, demand saving became more significant. So, in later years, the box plots in Figure 6 became more shifted upward, compared to those in Figure 1. We use the adjusted daily peak demand data to estimate the parameters in (14).

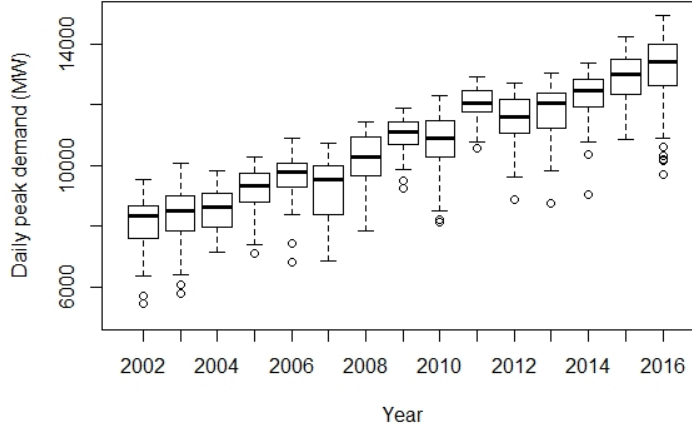


Figure 6: Adjusted Box plots of daily peak demands without DSM efforts in buildings during the summer in the south-central region of Texas

Next, the density function of the daily peak demand in year t , denoted by $u_t(z)$, can be obtained by

$$u_t(z) = \int h_t(z|y_t, p_t) g_t(p) f_t(y|x_t) dp dy, \quad (15)$$

where $h_t(z|y_t, p_t)$ is the conditional probability density function (pdf) of the daily peak demand in year t in (13), $g_t(p)$ is the pdf of the population in year t in (4), and $f_t(y|x_t)$ is the pdf of the daily peak temperature in year t in (1), given the grand ensemble temperature projection x_t from GCMs in (2).

The predictive density $u_t(z)$ in (15) does not take a closed-form. Thus, we determine the density using Monte Carlo sampling [69]. Specifically, n_1 and n_2 realizations of the daily peak temperature and population are sampled from the corresponding nonhomogeneous GEV $f_t(y|x_t)$, and logistic growth model, $g_t(p)$, respectively. Then, given each sampled y_t and p_t , we draw n_3 samples from the conditional daily peak load density, $h_t(z|y_t, p_t)$. With the total $n = n_1 \times n_2 \times n_3$ realizations, we obtain the unconditional daily peak load density, $u_t(z)$ in each year. Note that these random samples can be treated as potential scenarios of the daily peak temperature, population, and electricity demand. The sampling distribution from n realizations converges to the theoretical distribution $u_t(z)$, when n is sufficiently large. We use $n_1 = n_2 = n_3 = 1,000$ to get 10^9 samples each year in our implementation. Lastly, let z_t^s denote the the daily peak demand that accounts for the demand saving. We get z_t^s by subtracting the peak demand saving, discussed in Section 4.2, from z_t .

5 Case study

5.1 Validation

We use temperature and electricity usage data collected in south-central Texas in 2002-2016, as discussed in Section 3. For evaluating the density prediction performance, we divide the 15 years (2002-2016) of data into the two sets: training and test sets. The training set includes the data from 2002 to 2010 for the parameter

estimation, whereas the testing set contains data from 2011 to 2016. The density estimation performance is validated by comparing the predicted densities of daily peak temperature and demand with their real histograms in the testing set. Since the explanatory variables have different units, which can possibly contribute unequally in the analysis, we first standardize them by subtracting the sample mean and scaling to unit variance.

First, Figure 7 compares the histogram of actual standardized daily peak temperature with its density estimation (red curve) $f_t(y)$ in the testing set. Here, the histogram is obtained from 62 daily peak temperatures during July and August each year. Overall, the estimated density successfully characterizes the yearly varying nonstationary temperature pattern. In 2011, the actual histogram somewhat deviates from its predicted one. There were extraordinary extreme heatwaves in 2011 [70]. The daily peak temperature in the summer of 2011 ranged from 36.7°C to 43.0°C and its mean was 39.2°C . However, the GCM projections did not properly capture such record-breaking extreme events. In other years, we observe good agreements between the actual and predicted densities.

Next, we compare the histogram of the actual daily peak demand and its density estimation (red curve) $h_t(z|y_t, p_t)$ in Figure 8. The density estimation for the daily peak demand successfully matches the histogram of actual data across all years.

Both comparisons in Figures 7 and 8 suggest that the nonhomogeneous GEV distributions provide reasonably good fits for modeling the future daily peak temperature and demand densities. In both densities, the estimated shape parameters η_t and η'_t are negative, implying that the GEV distributions become the Weibull distributions. Therefore, the projected densities are negatively skewed (left-heavy tailed).

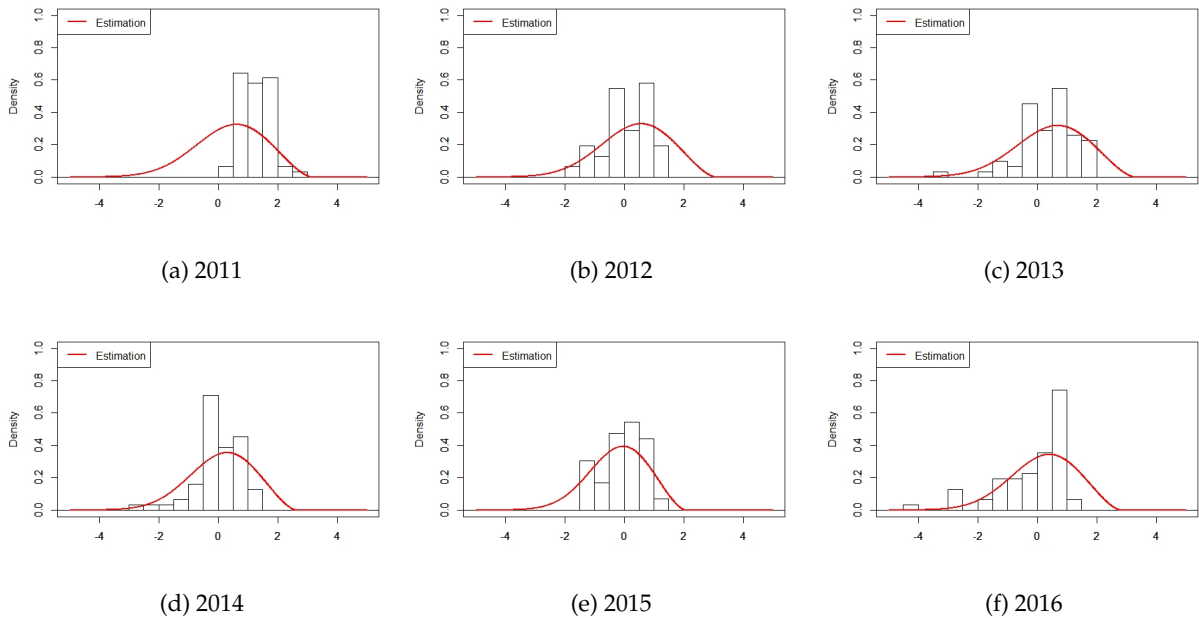


Figure 7: Comparison between empirical and estimated densities of daily peak temperature in the testing set

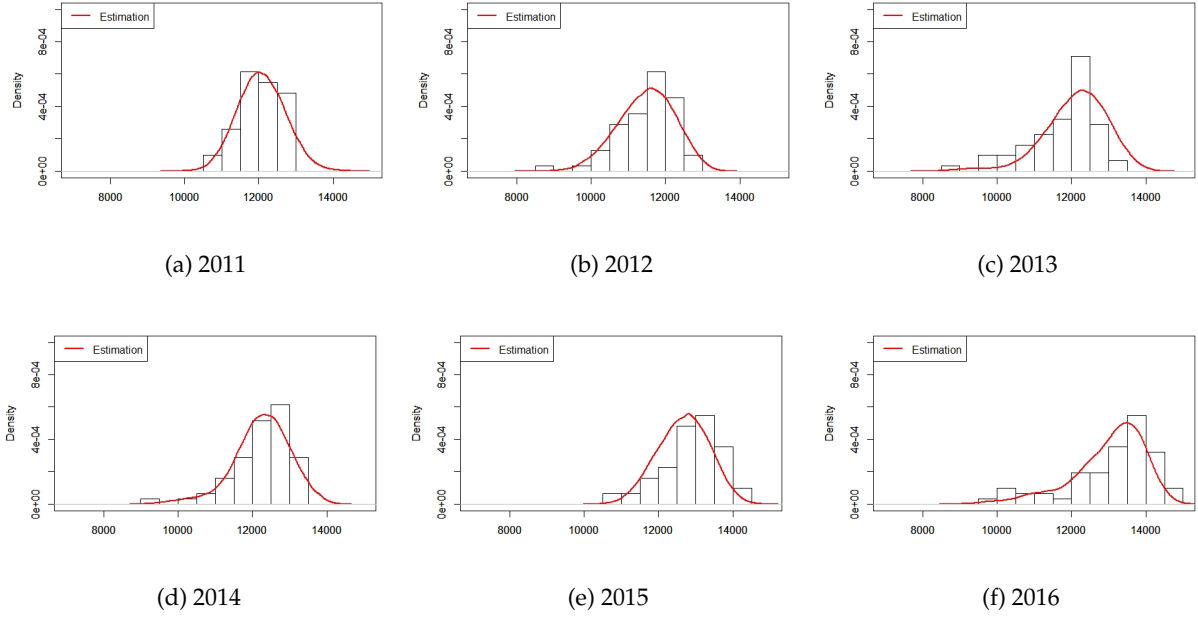


Figure 8: Comparison between empirical and estimated densities of daily peak demand in the testing set

We further compare the peak load density estimation performance of the proposed approach with the alternative the trend-based approach discussed in [43]. AlRashidi *et al.* [43] consider the historical trend to make predictions, using the linear and quadratic functions of year, unlike our approach that incorporates the effect of climate change and socio-economic factors on the daily peak loads. Their approach provides the point prediction only, so it cannot be directly compared with our approach. Alternatively, we formulate a similar structure by modeling the peak load as a function of year under the Normality assumption and estimate the parameters using the maximum likelihood estimation.

Figure 9 shows the density prediction results, where the red solid and blue dashed lines, respectively, represent the estimated densities from the trend-based model and proposed approach. While both prediction results are comparable in most cases, the trend-based model cannot capture the sudden changes in the daily peak loads in 2011 when heat wave events occurred. This result demonstrates the advantage of the proposed approach that uses GCM to capture the yearly-varying climate conditions.

5.2 Density prediction of daily peak demand

This section presents the predictive densities of daily peak demands from 2021 to 2040 with and without considering the effect of the demand saving from the DSM programs. Long-term grid asset planning typically studies 5 years to a few decades [71, 72]. In this study, 20 years was chosen to illustrate the demand change pattern during the typical planning horizon. First, Figure 10 shows the grand ensemble average of the daily peak temperature projections from GCMs. Overall GCMs suggest an increasing trend, which represents the influence of climate changes on the temperature.

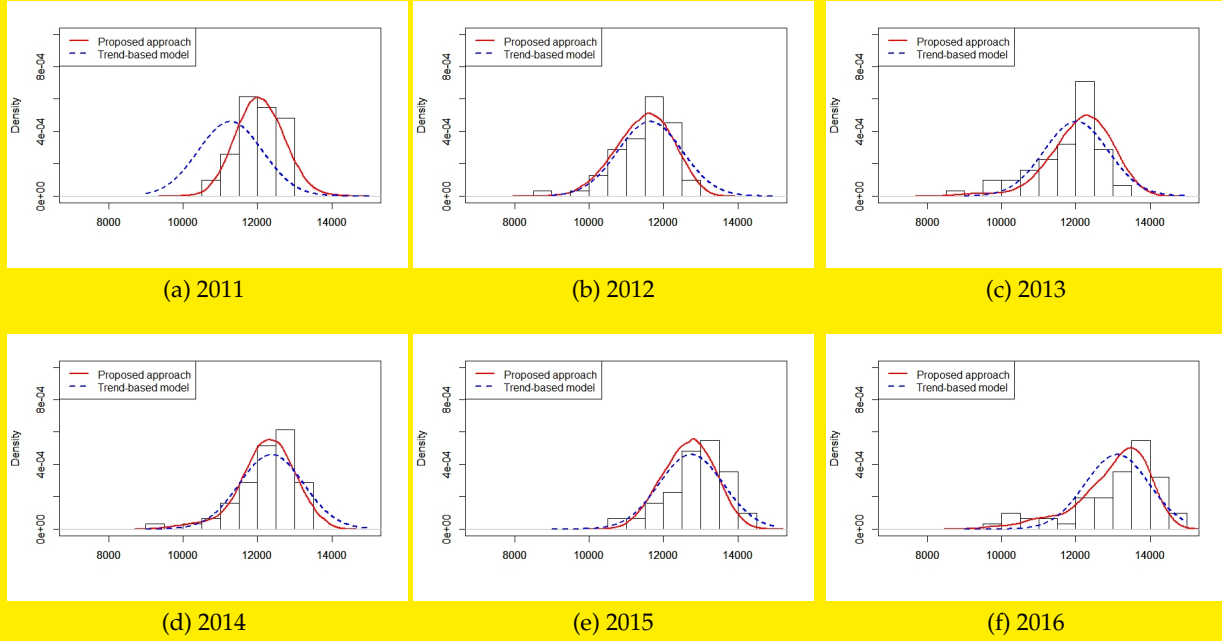


Figure 9: Comparison of the estimated densities of daily peak load from the trend-based and proposed approach in the testing set

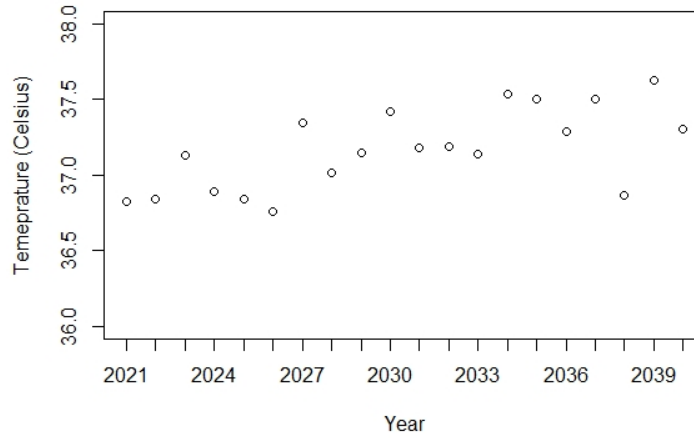


Figure 10: Grand ensemble average of daily peak temperature projections from GCMs

Figure 11 depicts the box and density plots of the predicted daily peak demand from 2021 to 2040 without considering the buildings' demand saving. In general, we can observe an increasing pattern over time in Figure 11a. The predictive mean and median of daily peak demands in 2040 are about 22.4 GW, if there would be no DSM efforts. Figure 11b further shows the predicted density in several selected years. As the prediction uncertainty increases in the long-term future, the predictive density becomes more flattened in later years.

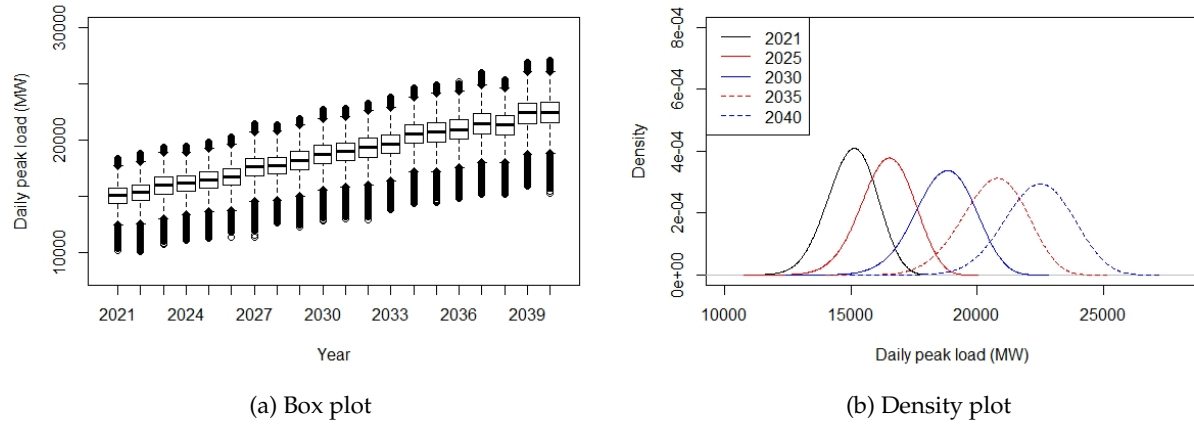


Figure 11: Box plot (left) and density plot (right) of predicted daily peak demands in 2021-2040 without buildings' demand saving

To account for the effect of the demand saving from DSM programs, we consider different values of d_s in (5). Figure 12 shows the expected demand saving in the south-central region of Texas from 2021 to 2040 with multiple values of d_s from 0.2 to 0.5 with an increment of 0.05. The demand saving increases and the range of the demand saving with different values of d_s becomes wider over time. In 2040, the demand saving is expected to range between 2.1GW to 4.1GW with $d_s \in [0.2, 0.5]$.

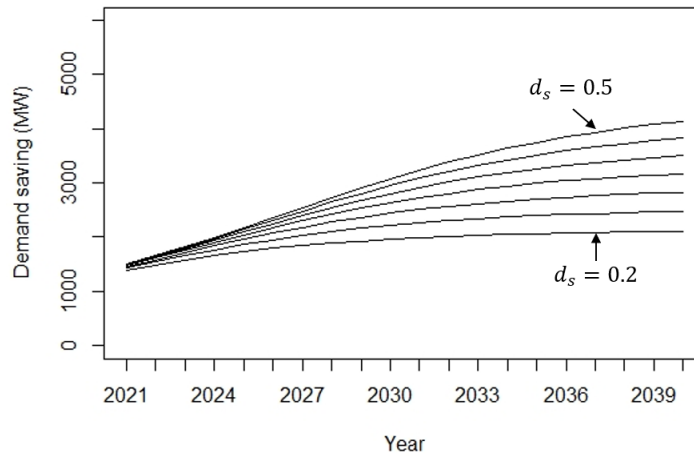


Figure 12: Predicted demand saving from DR programs from 2021 to 2040 in the south-central region of Texas

Figure 13 presents the predictive density of the daily peak demand after accommodating the effect of demand saving. Each sub-figure shows the densities with $d_s = 0.2, 0.3, 0.4$, and 0.5 . In the near future, the demand savings are not significantly different with different values of d_s , so the corresponding densities are highly overlapped. This is due to the fact that a small number of participants join the demand saving program in early years. As more people join DSM programs, the effect of the demand saving becomes more substantial

so the predicted densities become more shifted to the left in later years. Moreover, the difference in demand savings with $d_s = 0.2$ and $d_s = 0.5$ become more clear in later years.

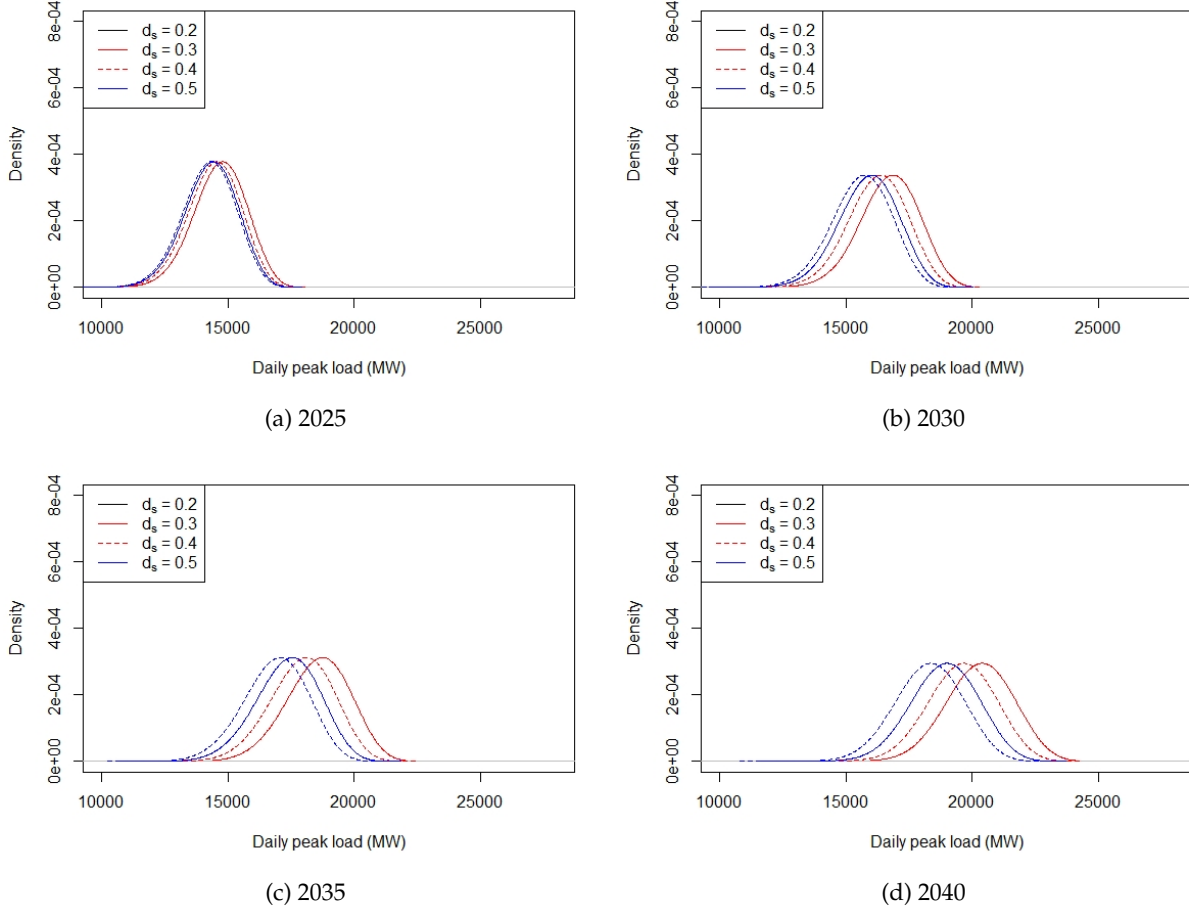


Figure 13: Daily peak density prediction with and without demand saving from DR programs

Table 1 summarizes the predicted mean and median and 90% PI of the daily peak demand every 5 years from 2025 to 2040. It is expected that the average peak demand would be reduced by 9.43%, 12.62%, 15.64%, and 18.46%, respectively, in 2040 with $d_s = 0.2, 0.3, 0.4$, and 0.5 , compared to the case with no demand saving (i.e., $d_s = 0$). We also observe that the 90% PI becomes wider over time due to the increasing uncertainty in the long-term future.

To further explore the impact of the DSM participation rate on the long-term peak demand, Figure 14 shows the predictive mean trajectories in a range of d_s value. It echos our previous observation: as the DSM program proceeds, its effectiveness becomes more clear in the long-run. In particular, an aggressive adoption of demand saving that can possibly lead to higher d_s can substantially alleviate the burden on the electric power grid to meet the increasing demand.

Table 1: Predicted mean and median and 90% PI of daily peak demand with DR programs (unit: GW)

Year	Demand saving rate (d_s)	Predicted mean	Predicted median	90% PI
2025	0	16.4	16.4	(14.6, 18.0)
	0.2	14.6	14.7	(12.8, 16.3)
	0.3	14.4	14.4	(12.6, 16.0)
	0.4	14.3	14.3	(12.5, 15.9)
	0.5	14.2	14.2	(12.4, 15.8)
2030	0	18.6	18.7	(16.6, 20.5)
	0.2	16.2	16.2	(14.2, 18.0)
	0.3	16.2	16.2	(14.2, 18.0)
	0.4	15.8	15.9	(13.8, 17.7)
	0.5	15.6	15.6	(13.6, 17.4)
2035	0	20.6	20.6	(18.4, 22.6)
	0.2	18.5	18.6	(16.3, 20.5)
	0.3	17.9	17.9	(15.7, 19.9)
	0.4	17.3	17.4	(15.1, 19.3)
	0.5	16.8	16.9	(14.6, 18.8)
2040	0	22.4	22.4	(20.1, 24.5)
	0.2	20.2	20.3	(18.0, 22.4)
	0.3	19.5	19.6	(17.3, 21.7)
	0.4	18.9	18.9	(16.6, 21.0)
	0.5	18.2	18.3	(16.0, 20.4)

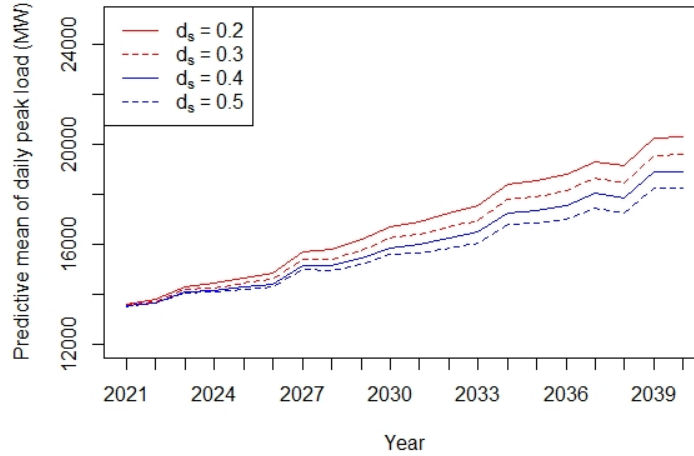


Figure 14: Predictive mean of the daily peak demand with and without demand savings participation

6 Conclusions and future plans

This study projects the long-term density of the daily peak demand with the goal of understanding its growing pattern and ultimately reducing the resulting burden on the power grid. Our approach accounts for the changes both in temperature due to climate change and in the socio-economic variables. Specifically, the proposed daily peak temperature model with the nonhomogeneous GEV framework allows us to adjust the possible biases in the GCM projections while keeping the temporal variation suggested by the GCMs. In addition, the expected population growth (or decay) pattern and the building demand saving from

DSM programs are formulated with the logistic growth model and Bass diffusion model, respectively. The presented approach is validated in the case study with actual data collected in the south-central region of Texas. The results provide useful insights on how the daily peak demand densities would change over time, in response to climate change, population growth and participation in DSM activities.

While this study provides a generic framework for characterizing the progression of peak loads in the long-term, it has limitations, because the results were obtained with limited data. In particular, as we observe in Section 5, DSM activities for buildings could substantially affect the peak load. Although we use actual building use data collected in south-central Texas, DSM programs and their effectiveness will evolve over time. At the same time, electricity usage patterns, as well as the willingness to participate in DSM programs could change in the future, as a result of technology advance such as growing popularity of smart appliances and electric vehicles, distributed renewable energy and improved internet connectivity. As such, our demand reduction model due to DSM needs to be updated and refined, as we get more data in the future. Therefore, the resulting prediction that incorporates the current and past DSM interventions, presented in this study, will be more useful in a shorter time horizon.

We plan to accommodate changes in DSM programs into the proposed modeling framework in our future study. Moreover, the datasets employed in this study have been collected in different spatial resolutions. For example, we use the temperature measurement in Austin and assume they represent the overall temperature pattern in south-central Texas. To address such limitations, we plan to extend our analysis to consider the spatial and temporal characteristics of the temperature, electricity demand and demand saving in multiple areas. Other socio-economic variables, such as gross domestic product, industrial production, will be additionally considered to enhance the prediction capability in our future study.

References

- [1] Haeran Cho, Yannig Goude, Xavier Brossat, and Qiwei Yao. Modeling and forecasting daily electricity load curves: a hybrid approach. *Journal of the American Statistical Association*, 108(501):7–21, 2013.
- [2] AK Srivastava, Ajay Shekhar Pandey, and Devender Singh. Short-term load forecasting methods: A review. In *2016 International Conference on Emerging Trends in Electrical Electronics & Sustainable Energy Systems (ICETEESES)*, 130–138. IEEE, 2016.
- [3] Fatima Amara, Kodjo Agbossou, Yves Dubé, Soussou Kelouwani, Alben Cardenas, and Sayed Saeed Hosseini. A residual load modeling approach for household short-term load forecasting application. *Energy and Buildings*, 187:132–143, 2019.
- [4] Hanane Dagdougui, Fatemeh Bagheri, Hieu Le, and Louis Dessaint. Neural network model for short-term and very-short-term load forecasting in district buildings. *Energy and Buildings*, 203:109408, 2019.
- [5] Tanveer Ahmad and Huanxin Chen. Short and medium-term forecasting of cooling and heating load demand in building environment with data-mining based approaches. *Energy and Buildings*, 166:460–476, 2018.

- [6] Wanlei Xue, Chenyang Li, Xuejiao Mao, Xuan Li, Long Zhao, and Xin Zhao. Medium and long term load forecasting of regional power grid in the context of economic transition. In *2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2)*, 1–4, 2018.
- [7] Tanveer Ahmad and Huanxin Chen. Potential of three variant machine-learning models for forecasting district level medium-term and long-term energy demand in smart grid environment. *Energy*, 160:1008–1020, 2018.
- [8] Elias Kyriakides and Marios Polycarpou. Short term electric load forecasting: A tutorial. In *Trends in Neural Computation*, 391–418. Springer, 2007.
- [9] Eva Gonzalez-Romera, Miguel A. Jaramillo-Moran, and Diego Carmona-Fernandez. Monthly electric energy demand forecasting based on trend extraction. *IEEE Transactions on power systems*, 21(4):1946–1953, 2006.
- [10] Eugene A. Feinberg and Dora Genethliou. Load forecasting. In *Applied Mathematics for Restructured Electric Power Systems*, 269–285. Springer, 2005.
- [11] Spencer Abraham. National transmission grid study. Technical report, USDOE Office of the Secretary of Energy, Washington, DC (United States), 2003.
- [12] Eric Hirst. US transmission capacity: present status and future prospects. *Edison Electric Institute*, 2004.
- [13] Henrique Steinherz Hippert, Carlos Eduardo Pedreira, and Reinaldo Castro Souza. Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on power systems*, 16(1):44–55, 2001.
- [14] Iain Staffell and Stefan Pfenninger. The increasing impact of weather on electricity supply and demand. *Energy*, 145:65–78, 2018.
- [15] 2015 Residential Energy Consumption Survey Data. <https://www.eia.gov/consumption/residential/data/2015/>, accessed in July, 2019.
- [16] 2018 Commercial Buildings Energy Consumption Survey Data. <https://www.eia.gov/consumption/commercial/data/2018/index.php?view=methodology>, accessed in July, 2019.
- [17] David E. Jahn, Jr. Gallus, William A., Phong T. T. Nguyen, Qiyun Pan, Kristen Cetin, Eunshin Byon, Lance Manuel, Yuyu Zhou, and Elham Jahani. Projecting central U.S. most likely annual urban heat extremes through mid-century. *Atmosphere*, 10:727, 2019.
- [18] Bert Metz, Ogunlade Davidson, Rob Swart, and Jiahua Pan. *Climate change 2001: mitigation: contribution of Working Group III to the third assessment report of the Intergovernmental Panel on Climate Change*, volume 3. Cambridge University Press, 2001.
- [19] Christopher B. Field, Vicente Barros, Thomas F. Stocker, and Qin Dahe. *Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change*. Cambridge University Press, 2012.
- [20] Thomas Stocker. *Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2014.

- [21] IPCC Sixth Assessment Report. <https://www.ipcc.ch/assessment-report/ar6/>, accessed in July, 2019.
- [22] National Oceanic and Atmospheric Administration. Climate modeling. https://www.gfdl.noaa.gov/wp-content/uploads/files/model_development/climate_modeling.pdf, accessed in July, 2019.
- [23] Juan Pablo Carvallo, Peter H. Larsen, Alan H. Sanstad, and Charles A. Goldman. Load forecasting in electric utility integrated resource planning. Technical report, Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States), 2017.
- [24] Mohamed H. Albadi and Ehab F. El-Saadany. A summary of demand response in electricity markets. *Electric Power Systems Research*, 78(11):1989–1996, 2008.
- [25] U.S. Energy Information Administration. <https://www.eia.gov/electricity/data/eia861/>, accessed in April, 2020.
- [26] Smart Electric Power Alliance. 2018 Utility Demand Response Market Snapshot. <https://sepapower.org/resource/2018-demand-response-market-snapshot/>, accessed in April, 2020.
- [27] Frank M. Bass. A new product growth for model consumer durables. *Management Science*, 15(5):215–227, 1969.
- [28] Shu Fan and Rob J. Hyndman. Short-term load forecasting based on a semi-parametric additive model. *IEEE Transactions on Power Systems*, 27(1):134–141, 2011.
- [29] Patrick E. McSharry, Sonja Bouwman, and Gabriël Bloemhof. Probabilistic forecasts of the magnitude and timing of peak electricity demand. *IEEE Transactions on Power Systems*, 20(2):1166–1172, 2005.
- [30] Shyh-Jier Huang and Kuang-Rong Shih. Short-term load forecasting via ARMA model identification including non-Gaussian process considerations. *IEEE Transactions on power systems*, 18(2):673–679, 2003.
- [31] James W. Taylor and Patrick E. McSharry. Short-term load forecasting methods: An evaluation based on European data. *IEEE Transactions on Power Systems*, 22(4):2213–2219, 2007.
- [32] Cheng-Ming Lee and Chia-Nan Ko. Short-term load forecasting using lifting scheme and ARIMA models. *Expert Systems with Applications*, 38(5):5902–5911, 2011.
- [33] Kunjin Chen, Kunlong Chen, Qin Wang, Ziyu He, Jun Hu, and Jinliang He. Short-term load forecasting with deep residual networks. *IEEE Transactions on Smart Grid*, 2018.
- [34] Song Li, Peng Wang, and Lalit Goel. A novel wavelet-based ensemble method for short-term load forecasting with hybrid neural networks and feature selection. *IEEE Transactions on power systems*, 31(3):1788–1798, 2015.
- [35] Che Guan, Peter B. Luh, Laurent D. Michel, Yuting Wang, and Peter B. Friedland. Very short-term load forecasting: wavelet neural networks with data pre-filtering. *IEEE Transactions on Power Systems*, 28(1):30–41, 2012.
- [36] Junhong Kim, Jihoon Moon, Eenjun Hwang, and Pilsung Kang. Recurrent inception convolution neural network for multi short-term load forecasting. *Energy and Buildings*, 194:328–341, 2019.

- [37] Ervin Ceperic, Vladimir Ceperic, and Adrijan Baric. A strategy for short-term load forecasting by support vector regression machines. *IEEE Transactions on Power Systems*, 28(4):4356–4364, 2013.
- [38] Ehab E. Elattar, John Goulermas, and Q. Henry Wu. Electric load forecasting based on locally weighted support vector regression. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(4):438–447, 2010.
- [39] Jorjeta G Jetcheva, Mostafa Majidpour, and Wei-Peng Chen. Neural network model ensembles for building-level electricity load forecasts. *Energy and Buildings*, 84:214–223, 2014.
- [40] Matthew Collins, Reto Knutti, Julie Arblaster, Jean-Louis Dufresne, Thierry Fichet, Pierre Friedlingstein, Xuejie Gao, William J. Gutowski, Tim Johns, Gerhard Krinner, Mxolisi Shongwe, Claudia Tebaldi, Andrew J. Weaver, Michael F. Wehner, Myles R. Allen, Tim Andrews, Urs Beyerle, Cecilia M. Bitz, Sandrine Bony, and Ben B. Booth. Long-term climate change: projections, commitments and irreversibility. In *Climate Change 2013-The Physical Science Basis: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 1029–1136. Cambridge University Press, 2013.
- [41] Ian M. Trotter, Torjus Folsland Bolkesjø, José Gustavo Féres, and Lavinia Hollanda. Climate change and electricity demand in Brazil: A stochastic approach. *Energy*, 102:596–604, 2016.
- [42] Toly Chen. A collaborative fuzzy-neural approach for long-term load forecasting in Taiwan. *Computers & Industrial Engineering*, 63(3):663–670, 2012.
- [43] M. R. AlRashidi and K. M. El-Naggar. Long term electric load forecasting based on particle swarm optimization. *Applied Energy*, 87(1):320–326, 2010.
- [44] Frits Møller Andersen, Helge V. Larsen, and Trine Krogh Boomsma. Long-term forecasting of hourly electricity load: Identification of consumption profiles and segmentation of customers. *Energy conversion and Management*, 68:244–252, 2013.
- [45] Changhao Xia, Jian Wang, and Karen McMenemy. Short, medium and long term load forecasting model and virtual load forecaster based on radial basis function neural networks. *International Journal of Electrical Power & Energy Systems*, 32(7):743–750, 2010.
- [46] Rob J. Hyndman and Shu Fan. Density forecasting for long-term peak electricity demand. *IEEE Transactions on Power Systems*, 25(2):1142–1153, 2010.
- [47] Tao Hong, Jason Wilson, and Jingrui Xie. Long term probabilistic load forecasting and normalization with hourly information. *IEEE Transactions on Smart Grid*, 5(1):456–462, 2013.
- [48] Viatcheslav V. Kharin, Francis W. Zwiers, Xuebin Zhang, and Gabriele C. Hegerl. Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *Journal of Climate*, 20(8):1419–1444, 2007.
- [49] Daniel Cooley. Extreme value analysis and the study of climate change. *Climatic change*, 97(1-2):77, 2009.
- [50] Richard W. Katz. Statistical methods for nonstationary extremes. In *Extremes in a Changing Climate*, 15–37. Springer, 2013.

- [51] Linyin Cheng, Amir AghaKouchak, Eric Gilleland, and Richard W. Katz. Non-stationary extreme value analysis in a changing climate. *Climatic change*, 127(2):353–369, 2014.
- [52] Iowa Environmental Mesonet, Iowa State University. https://mesonet.agron.iastate.edu/request/download.phtml?network=TX_ASOS, accessed in July, 2019.
- [53] Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections Archive. <https://gdo-dcp.uc11nl.org/downscaled/cmip/projections>, accessed in July, 2019.
- [54] The Electric Reliability Council of Texas (ERCOT). <http://www.ercot.com/gridinfo/load>, accessed in July, 2019.
- [55] William Naggaga Lubega and Ashlynn S. Stillwell. Maintaining electric grid reliability under hydrologic drought and heat wave conditions. *Applied Energy*, 210:538–549, 2018.
- [56] The United States Census Bureau. <https://www.census.gov>, accessed in July, 2019.
- [57] Austin Energy. Annual report fiscal year 2017. <https://data.austintexas.gov/Utilities-and-City-Services/Energy-Efficiency-Peak-Demand-Reduction/3d4a-wzcg>, accessed in July, 2019.
- [58] Richard W. Katz. Statistics of extremes in climate change. *Climatic change*, 100(1):71–76, 2010.
- [59] Stuart Coles, Joanna Bawa, Lesley Trenner, and Pat Dorazio. *An introduction to statistical modeling of extreme values*, volume 208. Springer, 2001.
- [60] G. Lee, E. Byon, L. Ntamo, and Y. Ding. Bayesian spline method for assessing extreme loads on wind turbines. *The Annals of Applied Statistics*, 7(4):2034–2061, 2013.
- [61] Derek D. Headey and Andrew Hodge. The effect of population growth on economic growth: A meta-regression analysis of the macroeconomic literature. *Population and Development Review*, 35(2):221–248, 2009.
- [62] Dean Baker, J. Bradford De Long, and Paul R. Krugman. Asset returns and economic growth. *Brookings Papers on Economic Activity*, 2005(1):289–330, 2005.
- [63] The United States Census Bureau. International programs: international data base. <https://www.census.gov/programs-surveys/international-programs/about/idb.html>, accessed in July, 2019.
- [64] Cesare Marchetti, Perrin S. Meyer, and Jesse H. Ausubel. Human population dynamics revisited with the logistic model: how much can be modeled and predicted? *Technological forecasting and social change*, 52(1):1–30, 1996.
- [65] Timur V. Elzhov, Katharine M. Mullen, A. Spiess, and B. Bolker. R interface to the Levenberg-Marquardt nonlinear least-squares algorithm found in MINPACK. *Plus Support for Bounds*, 1–2, 2010.
- [66] Jorge J. Moré. The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical analysis*, 105–116. Springer, 1978.
- [67] Gustav Feichtinger. Optimal pricing in a diffusion model with concave price-dependent market potential. *Operations Research Letters*, 1(6):236–240, 1982.

- [68] United States Census Bureau. New Residential Construction. <https://www.census.gov/construction/nrc/index.html>, accessed in April, 2020.
- [69] Christopher Z Mooney. *Monte Carlo simulation*, volume 116. Sage publications, 1997.
- [70] Frank Heinz. Hottest summer smackdown: 2011 vs. 1980. <https://www.nbcdfw.com/news/local/heat-streak-scorches-on-126138273/2121542/>, 2011.
- [71] Ankita Singh Gaur, Partha Das, Anjali Jain, Rohit Bhakar, and Jyotirmay Mathur. Long-term energy system planning considering short-term operational constraints. *Energy Strategy Reviews*, 26:100383, 2019.
- [72] S. R. Khuntia, B. W. Tuinema, J. L. Rueda, and M. A. M. M. van der Meijden. Time-horizons in the planning and operation of transmission networks: an overview. *IET Generation, Transmission Distribution*, 10(4):841–848, 2016.

Acknowledgement

We thank Mr. Wesley Graham for helping us collect the Austin Energy’s demand saving program data and model the demand saving pattern. This work was supported by the National Science Foundation under Grants CMMI-1662553 and CMMI-1662691.