Data Science in Energy and Environmental Systems Using Multiple Data Sources

by

Youngchan Jang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Industrial and Operations Engineering)
in The University of Michigan
2021

Doctoral Committee:

Associate Professor Eunshin Byon, chair
Assistant Professor Raed Al Kontar
Associate Professor Johanna Mathieu
Associate Professor Siqian Shen

Youngchan Jang

mapsossa@umich.edu

ORCID iD: 0000-0002-1928-3931

To my family

ACKNOWLEDGEMENTS

I am truly thankful to my Ph.D. advisor, Professor Eunshin Byon for her guidance and support during my doctoral studies. I also would like to thank my dissertation committee members: Professor Siqian Shen, Professor Raed Al Kontar, and Professor Johanna Mathieu for their suggestions from my preliminary exam to dissertation.

My deepest appreciation goes to my family. My utmost and sincere thanks go to my wife Hyonjung for her endless love and support. My lovely kids, Daniel, Arin, and Chaerin, have given me a great strength and happiness.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

The energy sector in modern society is undergoing a rapid transformation from fuel-based generation to renewable generation. Further, distributed supply and demand, grid-responsive demand management, and other complex technology increasingly rely on environmental data throughout the energy supply chain, from power production to end uses. However, data analysis in energy systems presents major technical challenges, including spatio-temporal heterogeneity, localized characteristics, and disparate data sources.

This dissertation study aims to design data science solutions that address some of these challenges and model the dynamics of ambient environmental conditions that are closely tied with energy system operations. Climate conditions are often temporally and spatially correlated and exhibit a non-stationary nature, constantly changing all the time. To fully characterize these characteristics, this study utilizes the rich data available from multiple sources, including data collected at spatially distributed locations and data generated from disparate sources, e.g., field meteorological data and physics-based numerical weather prediction data.

This dissertation initiates two major ideas: (a) making use of data collected from multiple spatially dispersed locations; and (b) integrating data generated from physics-based numerical weather prediction models with actual climate measurements through a linkage function. Specifically, the following three research topics are investigated.

The first study develops a probabilistic model for assessing wind resources at a target location by utilizing wind data collected at nearby meteorological stations. By quantifying daily and spatially correlated diurnal patterns of the wind speed at multiple locations, the developed integrative approach provides compelling capabilities for evaluating the wind variability at non-observational locations, while quantifying prediction uncertainties. The results will provide rich information for deciding suitable wind farm sites.

Next, we make use of temperature projections from physics-based global climate models for the purpose of long-term electricity load predictions. While the physics-based climate models can provide useful climate projections in the long run, they inevitably exhibit systematic discrepancies (also called 'bias'), compared to actual climate conditions, because of incomplete characterization of local or regional variations. We calibrate the climate model projections to address possible biases and provide a long-term density prediction of peak electricity load. The results provide useful insights on how the daily peak demand densities would change over time, in response to climate change and other socio-economic factors.

Finally, we present another bias correction model that quantifies the spatially and temporally correlated bias from the physics-based numerical model output for urban temperature modeling. By combining both types of data, our approach can successfully characterize localized environmental variations over space and time and greatly improve the prediction accuracy compared to that of the original physics-based numerical model. The proposed approach helps understanding temperature variations over dispersed locations, depending on urbanization intensity. Such results can be useful for predicting electricity demand and effectively managing power system operations such as demand response programs.

The advantages of all proposed approaches are demonstrated with case studies using actual data. The results validate that the proposed approaches successfully address some of the challenges discussed above that arise in energy and environmental systems.

CHAPTER I

Introduction

Modern society increasingly relies on electric power systems. The U.S. end use of electricity reached 3.99 trillion kilowatt hours (kWh) in 2019, and the total demand is expected to increase in the next decades [2–4]. The nation's renewable energy sources play an important and integral role in attaining energy independence, sustainability, and affordability. Among all of the renewable energy sources, wind and solar power are the fast-growing renewable sources in the U.S. Their market share had grown to 19% in 2019 and is expected to double over the next 30 years [5]. The U.S. Department of Energy (DOE)'s goal is that wind power will account for 20% of the total resource mix by 2030 [6]. On the demand side, smart communication capabilities facilitate demand management services, such as shedding, shifting, and modulating the electricity end use [7].

Facing such transformation, data science can play a key role in providing system-wide, integrated approaches to managing modern power systems because it can provide predictions of renewable energy generation (supply side) and energy end use (demand side). It should be noted that both wind/solar renewable generation and electricity demand highly depend on ambient weather conditions. Wind conditions, such as wind speed, turbulence intensity, and wind direction, mainly determine the

amount of wind power generation, whereas solar power depends on radiation intensity and other environmental factors. From the perspective of demand, the electricity demand highly depends on ambient weather conditions, including temperature and humidity. Therefore, environmental systems are closely tied with energy systems.

This dissertation research focuses on analyzing ambient climate data and studies three applications in energy and environmental systems. In Section 1.1, we present the characteristics of ambient climate data. Section 1.2 summarizes three applications of data science in energy and environmental systems using multiple data sources.

1.1 Ambient environmental data characteristics

Ambient weather conditions represent a few key characteristics: (a) weather conditions are temporally and spatially correlated; (b) they are non-stationary, constantly changing all the time. To fully address these characteristics, this study takes full advantage of the rich data available from multiple sources. Here, data from multiple sources include data collected at spatially distributed locations, as well as data collected from disparate sources such as field environmental data and physics-based data. Below we elaborate ambient data characteristics in more detail.

1.1.1 Data collected from disparate sources

In the operations and planning of energy and environmental systems, forecasting can be conducted on different time horizons, such as short-term [8–11], medium-term [11–13], and long-term forecasts [12, 13]. Among them, short-term forecasts typically focus on a time horizon of fewer than 6-12 hours, impacting intra-day operations [14]. Medium-term forecasts, typically hours to weeks-ahead, are often used for negotiating energy and electricity-related contracts [15]. Long-term forecasting is generally

associated with a time horizon of months or years, up to several decades [16], which plays an important role in the generation, transmission, and distribution system planning.

Depending on the purpose of the analysis and the time horizon of the prediction, the effectiveness of data collected from disparate sources can differ. First, field data collected from meteorological stations are often used for short-term forecasting. Typically, data-driven approaches for short-term forecasting identify the characteristics of the current state and predict the near future trend by analyzing temporal correlation among data points. Hence, the prediction capability of pure data-driven approaches deteriorates and prediction uncertainties rapidly grow as the forecasting horizon gets longer. This is because these approaches do not account for external influences, such as climate change and/or other meteorological changes, on the local weather conditions.

For medium- or long-term forecasting, physics-based approaches are known to be useful. Physics-based approaches use physical descriptions of the mechanisms of atmospheric conditions. One of the most popular models in this approach is the numerical weather prediction (NWP) model, which simulates atmospheric processes. However, data from physics-based models have several shortcomings. First, they typically generate outputs with coarser-level spatial and temporal resolutions, which makes it hard to capture local characteristics successfully. Next, a systematic discrepancy between the real measurements and corresponding model outputs is oftentimes observed, arising from imperfect model formulations and/or incorrect initial and boundary conditions for running the model [17]. Furthermore, in general, physics-based models provide deterministic point predictions that cannot quantify prediction uncertainties. Lastly, due to high computational cost and complexity,

running NWP models is computationally costly.

Therefore, depending on the problems and prediction time horizon, it is important to choose appropriate data sources and use them in an intelligent manner. This dissertation study investigates how to combine both field observational data and physics-based projections effectively to improve prediction performance.

1.1.2 Spatial data collected at distributed locations

Statistical temporal models for forecasting future observations at specific locations have been extensively studied in the literature. The fundamental concept is that the environmental conditions at each time are strongly correlated with their near-past conditions and thus, studying the correlation pattern between consecutive data points allows us to predict future conditions. Autoregressive moving averages and their variants have been used [18–24], and machine learning models including neural network models [10]. nonparametric approaches [21–24], and other stochastic processes [25, 26] have been proposed in the literature.

Recently, researchers have noted that more valuable information can be obtained by utilizing information collected over space [27–36]. In general, data at sites located close to each other exhibit greater similarity, compared to those at widely separated locations. Spatio-temporal modeling deals with such spatial correlation, as well as temporal correlation.

To illustrate, Figure 1.1 shows a partial layout of actual meteorological stations in West Texas Mesonet [37]. West Texas Mesonet is an integrated network of meteorological monitored stations designed to observe meteorological conditions in the West Texas region [37]. The x- and y-axes in Figure 1.1 represent the longitude and latitude, respectively. The solid circles denote the meteorological stations where

wind speed data is collected.

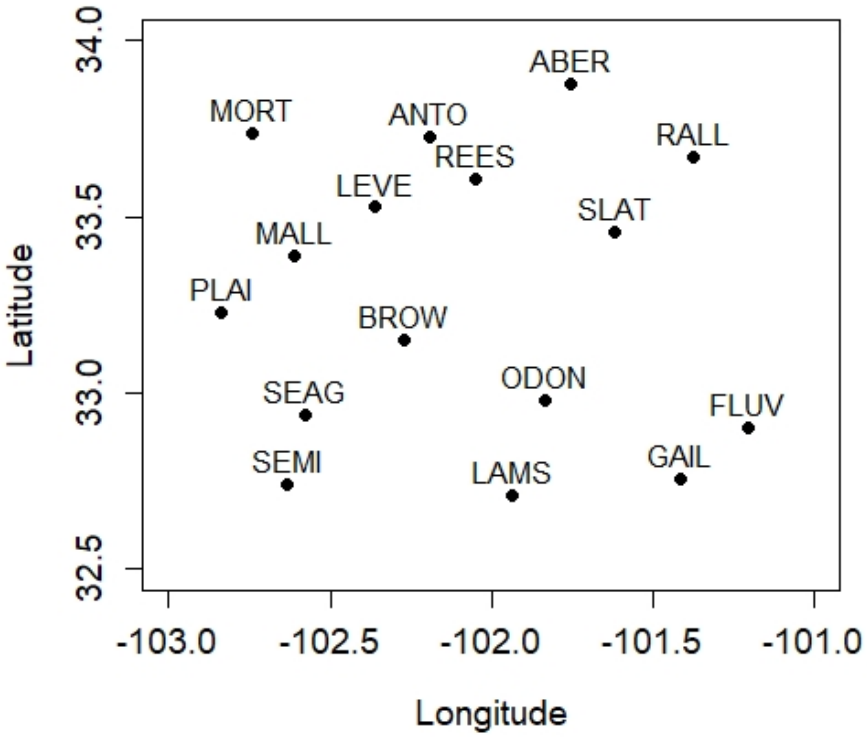


Figure 1.1: Layout of multiple stations in the West Texas region.

Figure 1.2 shows hourly average wind speeds at three stations, ANTO, REES, and SEMI, during the first week of January 2008. It is observed that the overall wind speed pattern during each day shows a diurnal cycle and the diurnal patterns among the three stations are highly correlated. In particular, the patterns in ANTO and REES are more similar to each other than they are to those in SEMI, because ANTO and REES are more closely located.

A typical approach is to directly formulate the correlation among observations. However, this has limitations with regard to integrating other information, such as NWP projections, with field observations. To address these limitations, in this study, we employ varying coefficient models and treat model coefficients as latent

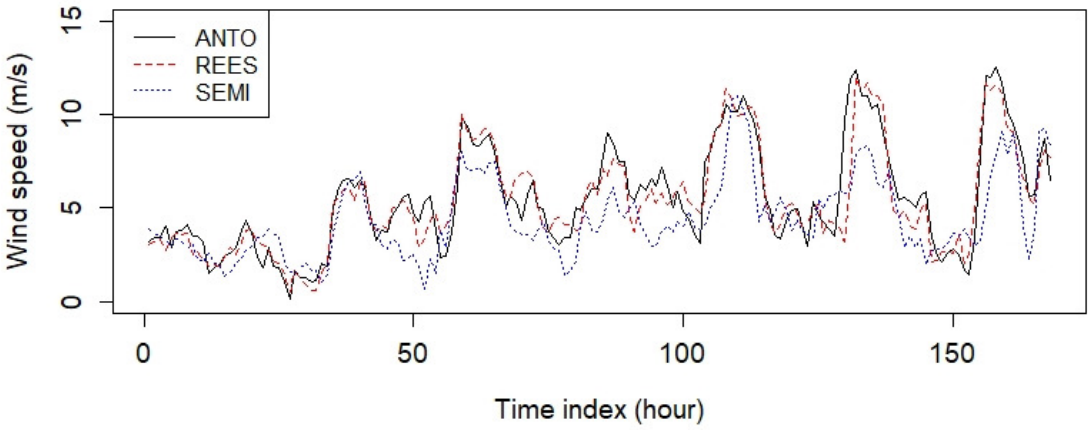


Figure 1.2: Wind speed patterns at three stations, ATNO, REES, and SEMI in West Texas in the first week of January 2008.

processes [38]. In particular, we formulate the model coefficients to vary spatially and temporally to capture the spatial and temporal correlation.

We briefly explain the underlying idea here. More detailed discussions will be provided in subsequent chapters. Let $y(s,t)$ and $x(s,t)$, $s = 1, \ldots, S$ and $t = 1, \ldots, T$, denote the dependent and independent variables at the space s and time t, respectively. T and S, respectively, represent the numbers of temporal and spatial data. The dependent variable, $y(s,t)$, which can be the measured climate condition at station s and time t, can be modeled by

$$y(s,t) = \beta_0(s,t) + \beta_1(s,t)x(s,t) + \epsilon(s,t), \tag{1.1}$$

where $\beta_0(s,t)$ and $\beta_1(s,t)$, respectively, represent the intercept and slope coefficients that evolve over space and time and $\epsilon(s,t)$ is Gaussian random noise such that $\epsilon(s,t) \sim N(0,\sigma^2)$. Here, $x(s,t)$ is a basis function. Depending on the problems, there can be multiple basis functions. Here, we use one basis function for brevity.

We note that the climate conditions at stations located close to each other and at temporally adjacent points exhibit greater similarity, compared to those at widely

separated locations and/or times. This implies that the model in (1.1) should similar; in other words, $\beta_0(s,t) \approx \beta_0(s',t')$ and $\beta_1(s,t) \approx \beta_1(s',t')$, for $s \approx s'$ and $t \approx t'$. To reflect such temporal and spatial correlations, instead of using fixed coefficients, we formulate $\beta_i(s,t)$ for $i = 0, 1$ as random effects. In particular, we treat them as latent Gaussian processes (GPs). Thus, the collection of coefficients, denoted by $\mathcal{B}_i = [\beta_i(s_1,t_1), \ldots, \beta_i(s_S,t_T)]^T$ becomes

$$\mathcal{B}_i \sim GP(0, \Sigma_i), \tag{1.2}$$

where Σ_i is the spatial and temporal covariance matrix for $i = 0, 1$.

The proposed model structure has a similar model formulation as a linear linkage model, so it provides easy interpretation. It also offers wide flexibility due to its semi-parametric formulation [38–40].

1.2 Dissertation Outline

In this dissertation, we study three specific topics (Figure 1.3) arising in energy and environmental systems, which are briefly outlined below.

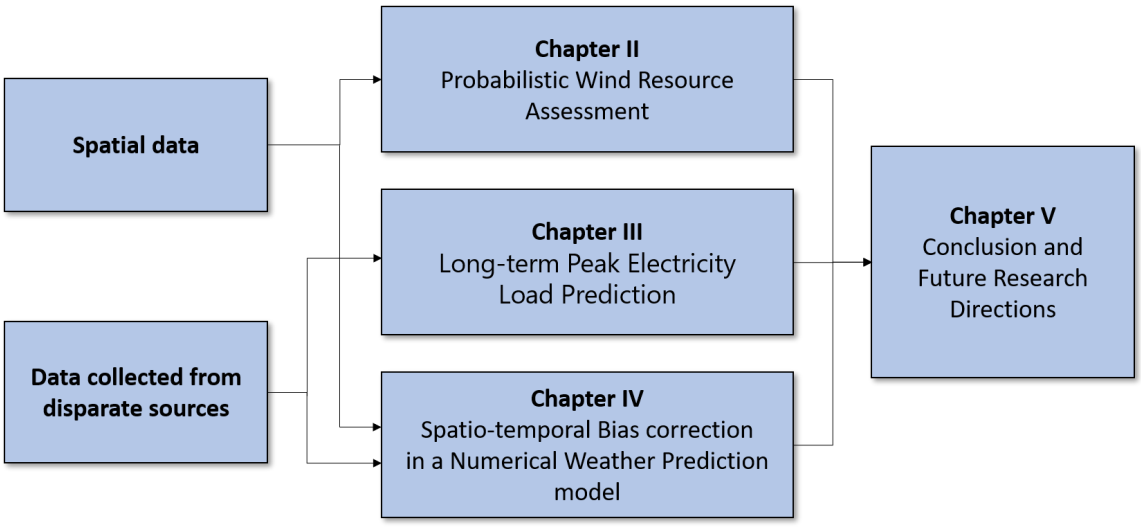


Figure 1.3: Overview of dissertation study.

1.2.1 Probabilistic Characterization of Wind Diurnal Variability for Wind Resource Assessment

In Chapter II, we present a new probabilistic modeling approach to quantifying variation in the wind diurnal pattern for assessing wind resources at unmonitored locations. As wind energy penetration is expected to grow in the future, wind resource assessment becomes important in modern power grid operations. Selecting an appropriate wind farm site can made easier by understanding the nonstationary characteristics of wind speeds. In particular, wind speed exhibits a diurnal pattern and the pattern varies, day-by-day and site-by-site.

The goal of this chapter is to develop a new probabilistic modeling approach to quantifying variation in the wind diurnal pattern for assessing wind resources at unmonitored locations. Specifically, we formulate the coefficient of the wind model as a latent random process and incorporate both day-to-day and spatial variability into the latent process. The estimation performance of the proposed approach is validated with actual data collected in West Texas.

1.2.2 On the long-term density prediction of peak electricity load with demand side management in buildings

Long-term daily peak demand forecast plays an important role in the effective and economic operation 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. Chapter III presents a new approach to providing 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 climate model projections with actual temperature

measurements. 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. We conduct a case study using actual building use data in the south-central region in Texas. The result suggests that a well-established building demand saving strategy is predicted to buffer against the growing needs of long-term peak electricity demand.

1.2.3 Spatio-temporal bias correction in a Numerical Weather Prediction model with application to urban temperature modeling during heat wave events

Characterizing localized climate conditions is becoming important in many aspects of modern society, e.g., for energy planning. The Weather Research and Forecasting (WRF) models, which are mesoscale NWP models, have been widely used to predict localized environmental variations. Further, the recently developed Urban Canopy Model (UCM), which is derived from energy balance equations, captures more detailed urban characteristics when it is coupled with the WRF model. However, as discussed earlier, such physics-based numerical models can exhibit a systematic discrepancy (also called 'bias'). The discrepancy pattern is often heterogeneous, varying over time and space.

In Chapter IV, we devise a new model that post-calibrates spatially and temporally varying discrepancies from the NWP model compared with measured climate conditions in an integrative framework. We test the proposed approach on urban temperature data collected in the Central Texas region during heat wave events.

Finally, Chapter V concludes the dissertation and discusses other possible future

directions for the data science in energy and environmental systems.

Chapters II-III include works in [41, 42]. Note that each chapter is self-contained and that the same notations or variables may be used in different chapters with different meanings. Thus, we define them in each chapter individually.

CHAPTER II

Probabilistic Characterization of Wind Diurnal Variability at Non-observational Locations For Wind Resource Assessment

2.1 Introduction

Selecting an appropriate wind farm site is vital for the success of wind energy in both financial and operational aspects. Among several factors to be considered for assessing potential wind farm site suitability, wind speed is clearly one of the most important factors. In general, windy areas are desired for installing a new wind farm. However, due to the wind's nonstationary characteristics, quantifying the wind variability is also inevitable for effective power grid operations [43].

One of the most effective ways for comprehensive wind resource assessment is to construct a wind model that characterizes its diurnal pattern. However, weather measurement at a candidate site may not be necessarily available in practice. For such a non-observational site, a meteorological tower can be installed to collect wind speed measurement. For more reliable wind resource assessment, it is essential to collect and analyze long-term wind resource (e.g., a year). However, installing new meteorological towers to collect long-term data is expensive and time-demanding for practical purposes.

Instead, one can collect wind measurement at a target site for a short-period of time. Such data collection activity is called measurement campaign [44, 45]. When a weather station, which collects long-term data, exists at a location close to the target site, the relationship of wind speeds between the two sites can be established and the speed at the target site can be estimated using the measurement at the weather station. Kwon [46] formulates the wind velocity at the target site as a linear function of the velocity at the weather station and estimates the linear function using the measurement collected at two locations. Jung *et al.* [47] further extend the approach in [46] and propose the Bayesian framework to handle various types of uncertainties due to limited data collected during the short-term measurement campaign. Similarly, Martinez-Cesena *et al.* [48] use the linear model between the annual average mean wind speeds at the target and measured locations. Even though these studies do not require long-term collection of measurement at the target location, short-term measurement is still needed. Moreover, they generally focus on quantifying the annual distribution of wind speed without considering its time-varying and nonstationary characteristics.

Another approach is to use NWP model. Zhang *et al.* [44] compare several NWP-based wind resource assessment methods using three datasets, including the Modern-Era Retrospective Analysis for Research and Applications dataset which is a low-resolution dataset, the Wind Integration National Dataset (WIND) which is a high-resolution dataset based on the WRF model, and short-term campaign measurement. It was shown that the analog ensemble method, which integrates the low-resolution NWP dataset with the short-term measurement, provides the best estimate of wind distribution in most sites considered in their study, whereas the WIND is suitable for estimating the distribution of the difference between two

consecutive hourly wind speeds. Jimenez *et al.* [49] compare two different weather prediction models at six locations, including offshore, onshore and island sites. In the study by [50], six different WRF simulations are conducted with different initial and boundary conditions and their estimation performances are compared with measured data at thirteen weather stations in Portugal. Their study shows that the new initial and boundary datasets improve the prediction accuracy over the old datasets. However, running NWP models requires considerable computational burden, and appropriate initial and boundary conditions need to be set *a priori*.

Unlike these prior studies, we consider a case where wind measurement at nearby locations are available. Recent advances in sensing technology make meteorological measurement increasingly available at many locations. This motivates us to assess wind resource when measurement at the target site does not exist, but data near the target site is available. Some recent studies propose a spatial model for predicting wind speed at a non-observational location. Lenzi *et al.* [27] apply the GP to wind measurements collected at neighbor locations at each time point. Byon *et al.* [51] spatially interpolate wind speeds at neighbor monitored stations for estimating the speed at a target unmonitored station. Although the models in [27, 51] provide estimates at non-observational sites, they only provide a snapshot estimate at each time instance, rather than providing wind models at the target site. As such, the spatial snapshot approach cannot fully describe the wind variability over time.

To illustrate, Figure 2.1 is the same layout of actual meteorological stations in West Texas Mesonet [37] shown in Figure 1.1 in Chapter I. Suppose that the red star in Figure 2.1 is a potential wind farm site where wind measurement is not available. As we seen in Figure 1.2, the patterns at closely located sites show similar diurnal patterns. Therefore, we can borrow information of time-series characteristics

at neighbor stations to assess wind resources at the target site. It is also observed that diurnal cycles change day-by-day. For example, the pattern on the first day is quite different from that on the last day in Figure 1.2.

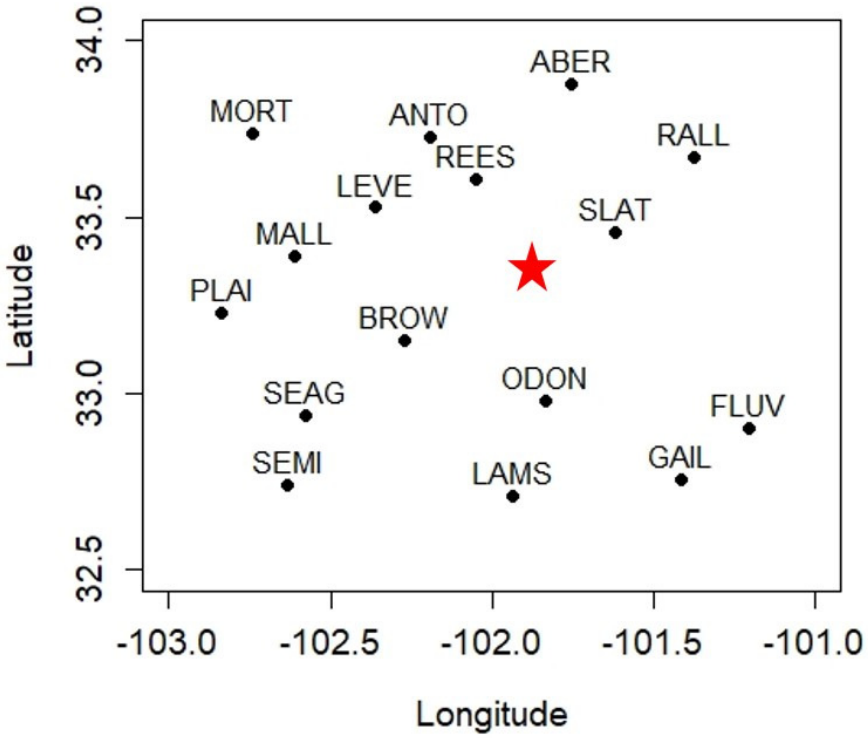


Figure 2.1: Layout of multiple stations in west Texas region.

Figure 2.2 further shows daily patterns at BROW during January 2008, where the thick curve represents the average diurnal pattern during January. Although there is commonality, wind patterns substantially differ day-by-day. As such, one cannot fully characterize wind variability with the average pattern only. Therefore, the wind resource assessment requires thorough understanding of the spatially- and daily-varying nonstationary characteristics.

This chapter develops a systematic approach to estimate diurnal patterns of wind speed and to quantify estimation uncertainties by using measurements collected at

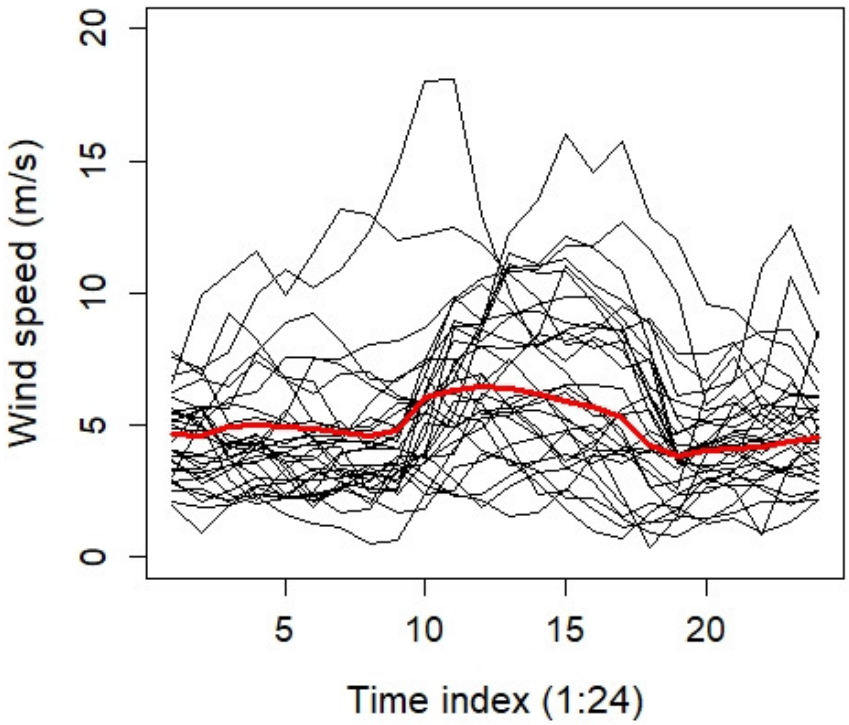


Figure 2.2: Day-to-day variations at BROW on January 2008.

spatially dispersed nearby stations. We present a new modeling approach that formulates the time-varying pattern with daily- and spatially-varying coefficients. The parameters in the proposed model are estimated in a Bayesian hierarchical framework.

The main contribution of this chapter is two-fold: (1) Unlike the aforementioned studies that use the short-term measurement campaign data and/or NWP data, our approach uses wind measurement collected at nearby locations; (2) In contrast to the studies in [27, 51], the proposed approach provides a probabilistic wind model, which enables us to fully characterize the time-varying pattern of wind speed and quantify the uncertainties. The resulting model can generate scenarios of wind speed trajectories, which can be used for investment decision-making in wind power projects [48].

A case study is carried out using actual data collected in West Texas Mesonet. The implementation results demonstrate that the proposed approach is capable of successfully characterizing the wind variability at unmonitored sites, which provides useful insights for wind resource assessment.

The remainder of this chapter is organized as follows. Section 2.2 discusses the proposed modeling approach and parameter estimation procedure. Section 2.3 presents a case study and Section 2.4 concludes the chapter.

2.2 Methodology

This section develops an integrative framework for quantifying the day-to-day and spatial variability in wind's diurnal pattern at non-observational locations.

2.2.1 Integrative Modeling Approach

We formulate the wind model using trigonometric functions to characterize a non-stationary pattern. Considering that diurnal patterns at neighbor locations should exhibit similarity and could change over different days, we make the model coefficients spatially correlated and daily-varying.

Let $Y(s,d,t)$ denote the wind speed at a location s at time t on day d. In this chapter, an hourly average measurements are considered, but the proposed approach can be applied to data with different temporal resolutions. To address the cyclic diurnal pattern, the wind speed, $Y(s,d,t)$, is formulated using L pairs of trigonometric functions [32, 33] as follows.

$$Y(s,d,t) = \mu(s,d,t) + \epsilon(s,d,t) \tag{2.1}$$

with

$$\mu(s,d,t) = \beta_0(s,d) + \sum_{\ell=1}^{L} \Big[\beta_{1,\ell}(s,d)\sin\frac{2\ell\pi t}{24} + \beta_{2,\ell}(s,d)\cos\frac{2\ell\pi t}{24}\Big], \tag{2.2}$$

where $\epsilon(s,d,t) \sim N(0,\sigma^2)$ is a Gaussian random noise. As a remark, an additional seasonal cycles can be included in $\mu(s,d,t)$ [51]. But such model assumes that the diurnal pattern remains the same in different seasons, which may not hold in practice. Instead, we suggest building monthly models with the formulation in (2.2), so heterogeneous diurnal patterns which could vary, depending on seasons, can be captured.

To capture day-to-day and location-to-location variations, we formulate each model coefficient as a latent process and decompose it into day-specific and site-specific random effects (Figure 2.3). The day-specific latent process captures the day-to-day variation, whereas the site-specific random effects quantify the spatial correlation among multiple sites. Specifically, let $\beta(s,d)$ denote a vector of model coefficients, i.e.,

$$\beta(s,d) = [\beta_0(s,d), \beta_{1,1}(s,d), \beta_{2,1}(s,d), \ldots, \beta_{1,L}(s,d), \beta_{2,L}(s,d)]^T. \tag{2.3}$$

and $\beta^i(s,d)$ denote the i^{th} coefficient of $\beta(s,d)$. Depending on the flexibility to specify $\beta^i(s,d)$, three different models, referred to as integrative model 1, 2 and 3 (shortly, IM1, IM2 and IM3) are proposed as follows.

- IM1 assumes that the day-specific parameter on day d are dependent on that on day $d-1$, implying that wind diurnal patterns on the two consecutive days are similar each other.

- IM2 allows more flexibility in describing day-to-day variations. It lets the day-specific parameter randomly vary day-by-day. Thus, IM2 is appropriate when the daily pattern changes significantly.

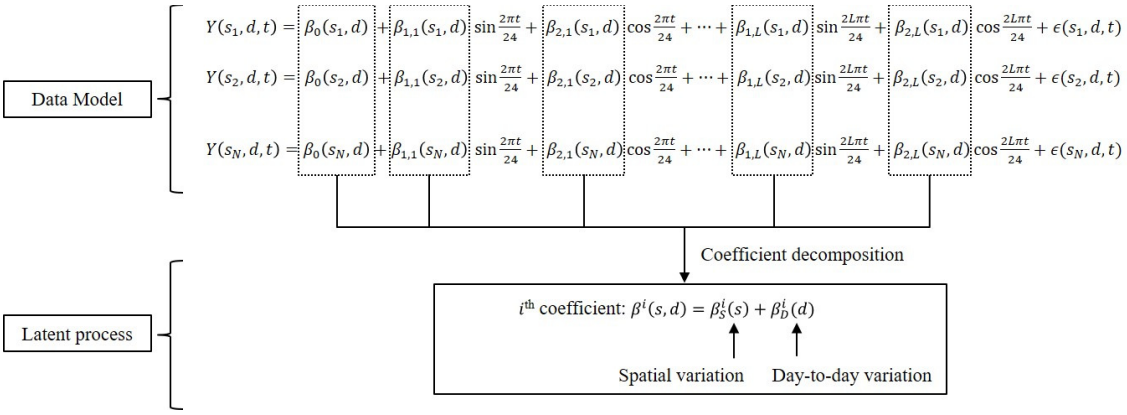


Figure 2.3: Overall framework of the integrated model 1 (IM1).

- IM3 further allows the spatial correlation structure to be heterogeneous on different days, in contrast to IM1 and IM2 which implicitly assume the homogeneous spatial correlation structure.

Below we describe each model in more detail.

2.2.2 Integrative Model 1 (IM1)

To quantify spatial and daily variations, $\beta^i(s,d)$ is decomposed into two components as

$$\beta^i(s,d) = \beta^i_S(s) + \beta^i_D(d) \tag{2.4}$$

where $\beta^i_D(d)$ is a day-specific coefficient and $\beta^i_S(s)$ is a location-specific coefficient. Figure 2.3 shows the overall framework of IM1 model.

First, to capture the spatial correlation, the site-specific coefficient $\beta^i_S(s)$ in (2.4) is formulated as a spatially-varying parameter [52, 53]. It should be noted that the pattern at closely located sites is similar to one another, as observed in Figure 1.2. Accordingly, $\beta^i_S(s_j)$ should be similar to $\beta^i_S(s_k)$ for closely located sites, s_j and s_k. To characterize such spatial dependency, $\beta^i_S(s)$ is modeled with the latent GP [54]

as

$$\beta_S^i(s) \sim \mathcal{GF}(\mu_i(s), C_i), \tag{2.5}$$

for $i = 1, 2, \cdots, 2L+1$, where μ_i and C_i denote the mean and covariance functions for $\beta_S^i(s)$, respectively.

By the consistency property of GP, a collection of $\beta_S^i(s)$'s jointly follow multivariate normal distribution [54]. Suppose that there are N monitored stations, $s = s_1, s_2, \cdots, s_N$. Let $\beta_{S,obs}^i$ denote an $N \times 1$ vector of $\beta_S^i(s)$'s, i.e., $\beta_{S,obs}^i = [\beta_S^i(s_1), \beta_S^i(s_2), \cdots, \beta_S^i(s_N)]^T$. Then, we have

$$\beta_{S,obs}^i \sim MVN(0, \Sigma_i), \tag{2.6}$$

where Σ_i denotes an $N \times N$ covariance matrix whose $(j,k)^{th}$ component, $c_i(s_j, s_k)$, is the covariance function between stations s_j and s_k. Here $c_i(s_j, s_k)$ is a positive definite kernel function. Among several choices for modeling $c_i(s_j, s_k)$, one of the commonly used covariance functions is the Matérn covariance function, defined as

$$c_i(s_j, s_k) = \frac{\tau_i^2}{2^{\nu-1}\Gamma(\nu)}(\kappa_i||x_j - x_k||)^\nu K_\nu(\kappa_i||x_j - x_k||), \tag{2.7}$$

where x_j is the location of station s_j, τ_i^2 is the marginal variance, K_ν is the modified Bessel function of second kind of order $\nu > 0$, $\Gamma(\cdot)$ is the Gamma function, $||\cdot||$ is the Euclidean distance, and κ_i is the decay parameter [55]. The parameter, ν, is a smoothness parameter, affecting differentiability of the underlying process. In general, ν is fixed to 1 for computational convenience [56]. Due to its flexibility and computational advantage, the Matérn covariance function is employed in our implementation, however, other covariance functions can be employed in the proposed framework [57].

Using the coefficients, $\beta_S^i(s)$'s ($s = s_1, s_2, \cdots, s_N$), the coefficient at an unmonitored site is obtained. Let s_0 denote a non-observational location. Given the parameter vector, $\beta_{S,obs}^i$, the coefficient, $\beta_S^i(s_0)$, at the non-observational site becomes normally distributed as

$$\beta_S^i(s_0)|\beta_{S,obs}^i \sim N(\mu_i(s_0), \tau_i^2(s_0)), \tag{2.8}$$

with

$$\mu_i(s_0) = c_i(s_0)^T \cdot \Sigma_i^{-1} \cdot \beta_{S,obs}^i, \tag{2.9}$$

$$\tau_i^2(s_0) = \tau_i^2 - c_i(s_0)^T \Sigma_i^{-1} c_i(s_0), \tag{2.10}$$

where $c_i(s_0) = [c_i(s_0, s_1), c_i(s_0, s_2), \cdots, c_i(s_0, s_N)]^T$ is an $N \times 1$ vector for $i = 1, 2, \cdots, 2L+1$. The results implies that once the spatial parameters at monitored sites are estimated, the parameter at an unmonitored site can be estimated accordingly. While we present the estimation procedure at a single unmonitored site in (2.8)-(2.10), the results can be extended for simultaneously estimating parameters at multiple sites [54].

Next, in describing the day-to-day variability, we observe that the daily pattern in one day tends to be similar to the pattern in the next day (Figure 1.2). Therefore, the day-specific parameter, $\beta_D^i(d)$, could be highly correlated with $\beta_D^i(d-1)$. To characterize such temporal correlation, we apply the autoregressive (AR) process to $\beta_D^i(d)$ as

$$\beta_D^i(d) = \beta_{D,0}^i + \rho_i \beta_D^i(d-1) + \epsilon_D^i, \tag{2.11}$$

where ϵ_D^i denotes the random noise, $\epsilon_D^i \sim N(0, \delta_{1i}^2)$. Here, we present AR1 for simplicity, but a higher order AR process can be employed.

2.2.3 Integrative Model 2 (IM2)

The second model, IM2, uses the same decomposition structure in (2.4) with the same spatial term. However, unlike IM1 that restricts the day-specific parameter to be temporally correlated, IM2 lets $\beta_D^i(d)$ be fully random day-by-day.

Specifically, $\beta_D^i(d)$ is formulated as random effects. Let β_D^i denote $D \times 1$ vector of $\beta_D^i(d)$'s, i.e., $\beta_D^i = [\beta_D^i(1), \beta_D^i(2), \ldots, \beta_D^i(D)]^T$. Then we have

$$\beta_D^i \sim MVN(0, \delta_{2i}^2 I), \tag{2.12}$$

for $i = 1, 2, \ldots, 2L+1$, where I is a $D \times D$ identity matrix and δ_{2i}^2 is the corresponding variance term.

2.2.4 Integrative Model 3 (IM3)

In IM1 and IM2, the spatial correlation is assumed to be homogeneous in different days (note that $\beta_S^i(s)$ is the same for all d's in (2.4)). To allow the heterogeneous spatial correlation structure on different days, IM3 breaks down $\beta^i(s, d)$ into two components as follows.

$$\beta^i(s, d) = \beta_S^i(s, d) + \beta_D^i(d). \tag{2.13}$$

Note that the spatial effect, $\beta_S^i(s, d)$, also depends on d, unlike $\beta_S^i(s)$ that depends on s only in IM1 and IM2 (see (2.4)).

As in IM1 and IM2, $\beta_S^i(s, d)$ is modeled as a latent GP. Let $\beta_{S,obs}^i(d)$ denote an $N \times 1$ vector of $\beta_S^i(s, d)$'s at monitored stations on day d, i.e., $\beta_{S,obs}^i(d) = [\beta_S^i(s_1, d), \beta_S^i(s_2, d), \cdots, \beta_S^i(s_N, d)]^T$. Then,

$$\beta_{S,obs}^i(d) \sim MVN(0, \Sigma_i(d)), \tag{2.14}$$

Here, $\Sigma_i(d)$ is an $N \times N$ covariance matrix on day d. It is assumed that $\beta_{S,obs}^i(d)$ is independent of $\beta_{S,obs}^i(d')$ for $d \neq d'$.

For the day-specific parameter $\beta_D^i(d)$, it is modeled as random effects in (2.12) as in IM2. Table 2.1 summarizes the decomposition structure in three integrative models.

Table 2.1: Decomposition structure in latent process

	IM1	IM2	IM3
Decomposition Structure	$\beta_S^i(s) + \beta_D^i(d)$	$\beta_S^i(s) + \beta_D^i(d)$	$\beta_S^i(s,d) + \beta_D^i(d)$
Spatial parameter	$\beta_S^i(s) \sim GP$	$\beta_S^i(s) \sim GP$	$\beta_S^i(s,d) \sim GP(d)$
Temporal parameter	$\beta_D^i(d) \sim AR$	$\beta_D^i(d) \sim$ *random effects*	$\beta_D^i(d) \sim$ *random effects*

2.2.5 Implication

Before discussing the parameter estimation procedure, it is worthwhile to discuss the primary difference between the proposed approach and the snapshot approach in [27, 51]. The snapshot approach directly formulates the correlation among $Y(s,d,t)$'s through interpolation techniques such as GP and kriging. For example, GP is applied to the measurement at monitored locations to estimate wind speed at the unmonitored location at each time instant. The salient feature of the proposed approach is that we characterize the correlation structure through the latent process, $\beta^i(s,d)$'s, instead of $Y(s,d,t)$'s. Assuming that the day-specific effect and spatial random effect, are independent, the covariance of model coefficients in IM1 is given by

$$\begin{aligned} & Cov(\beta^i(s_j,d), \beta^i(s_k,d')) \\ & = Cov(\beta_D^i(d), \beta_D^i(d')) + Cov(\beta_S^i(s_j), \beta_S^i(s_k)) \end{aligned} \tag{2.15}$$

where $Cov(\beta_S^i(s_j), \beta_S^i(s_k))$ is given in (2.7) and $Cov(\beta_D^i(d), \beta_D^i(d'))$ in AR1 [58] is

$$Cov(\beta_D^i(d), \beta_D^i(d')) = \delta_{1i}^2 \frac{\rho_i^{|d-d'|}}{1-\rho_i^2}. \tag{2.16}$$

Therefore, we get

$$Cov(\beta^i(s_j, d), \beta^i(s_k, d')) = \begin{cases} \tau_i^2 + \delta_{1i}^2 \frac{1}{1-\rho_i^2}, & \text{for } j = k, d = d' \\ c_i(s_j, s_k) + \delta_{1i}^2 \frac{1}{1-\rho_i^2}, & \text{for } j \neq k, d = d' \\ \tau_i^2 + \delta_{1i}^2 \frac{\rho_i^{|d-d'|}}{1-\rho_i^2}, & \text{for } j = k, d \neq d' \\ c_i(s_j, s_k) + \delta_{1i}^2 \frac{\rho_i^{|d-d'|}}{1-\rho_i^2}, & \text{for } j \neq k, d \neq d' \end{cases} \tag{2.17}$$

where $c_i(s_j, s_k)$ is the $(j,k)^{th}$ component of the covariance matrix, Σ_i, in (2.14). Note that $\beta_D^i(d)$ and $\beta_D^i(d')$ are strongly correlated when d and d' are closer, thereby making $\beta^i(s_j, d)$ and $\beta^i(s_j, d')$ similar to each other. Likewise, $\beta_S^i(s_j)$ and $\beta_S^i(s_k)$ at closely located s_j and s_k have larger covariance, $c_i(s_j, s_k)$. The covariance structures in IM2 and IM3 can be similarly specified. We omit them to save space.

2.2.6 Parameter Estimation

This section discusses the parameter estimation procedure. We focus our discussion on estimating parameters in IM1. The parameter estimation in IM2 and IM3 can be performed in a straightforward way. We use wind measurement at N monitored stations at time $t = 1, 2, \ldots, T$ (e.g., $T = 24$ for hourly collected data) during D days ($d = 1, 2, \cdots, D$). Because the day-specific and site-specific parameters are formulated as latent processes, the proposed model has a multi-level hierarchical structure. The first level formulates the data model in (2.1). The second level specifies the latent processes for the spatially- and daily-varying parameters. The last level provides a prior density for hyperparameters.

We estimate the parameters in the Bayesian inference framework [59]. Let $\mathcal{D}$ denote the dataset used for inference, and let Θ denote a set of all parameters in the model. The joint posterior density of Θ is given by

$$
\begin{aligned}
p(\Theta|\mathcal{D}) \propto & \left[\Pi_{s,d,t} f(Y(s,d,t)|\mu(s,d,t),\sigma^2)\right] p(\sigma^2) \\
& \times \Pi_i \left[MVN(0,\mathbf{\Sigma}_i) p(\rho_i) p(\delta_{1i}^2) p(\tau_i^2) p(\kappa_i)\right],
\end{aligned} \tag{2.18}
$$

where $f(Y(s,d,t)|\mu(s,d,t),\sigma^2)$ represents the likelihood of wind speed with $Y(s,d,t) \sim N(\mu(s,d,t),\sigma^2)$, $p(\rho^i)$ and $p(\delta_{1i}^2)$ imply the priors in the latent AR process for the daily-varying coefficient in (2.12), and $MVN(0,\mathbf{\Sigma}_i)$ denotes the latent Gaussian process in (2.14) for spatially-varying coefficients. Lastly, $p(\rho_i)$, $p(\delta_{1i}^2)$, $p(\tau_i^2)$, $p(\kappa_i)$, and $p(\sigma^2)$ denote prior densities for their corresponding parameters and hyperparameters.

The posterior mean from the posterior density in (2.18) is used for estimating parameters. Obtaining the posterior density requires multi-dimensional integration, and it is not derived in a closed form. Therefore, simulation-based methods such as Markov chain Monte Carlo (MCMC) can be used to approximate the posterior density. However, implementing MCMC demands expensive computational cost, so we use an approximation method. In particular, we employ the integrated nested Laplace approximations (INLA) in our analysis [57]. For more details on the INLA approximation procedure, please refer to [57]. In our analysis, 'R-INLA' package in the statistical software, R [56] is used. In our implementation, priors are specified as suggested in INLA. When parameters are estimated in the Bayesian hierarchical framework, it has been known that the deviance information criterion (DIC) is useful for choosing a model order, L, in our model [60]. For more details on DIC, please refer to [60].

Once the parameters are estimated, the estimated distribution of the wind speed at the target site s_0 is provided by

$$Y(s_0, d, t) \sim N(\hat{\mu}(s_0, d, t), \hat{\sigma}^2(s_0, d, t)) \tag{2.19}$$

with

$$\hat{\mu}(s_0, d, t) = \hat{\beta}_0(s_0, d) + \sum_{\ell=1}^{L} \Big[\hat{\beta}_{1,\ell}(s_0, d) \sin \frac{2\ell\pi t}{24} + \hat{\beta}_{2,\ell}(s_0, d) \cos \frac{2\ell\pi t}{24}\Big], \tag{2.20}$$

where $\hat{\beta}(\cdot)$'s in (2.20) denote the posterior means for the corresponding parameters and $\hat{\sigma}^2(s_0, d, t)$ is the posterior variance of $Y(s_0, d, t)$.

2.3 Case Study

We use wind measurement collected at 16 stations in West Texas Mesonet in this chapter. The shortest and average distances between two adjacent stations are 18.6 km and 36.3 km, respectively. The location information of the stations can be found in [37].

As discussed earlier, we suggest building monthly diurnal models to account for heterogeneous diurnal patterns in different seasons. Wind resource assessment requires quantification of year-long wind pattern. Due to the time limitation, we were not able to estimate year-long pattern, but we implement our method using four months data, including January, April, July and November in 2008. The original data contains 5-minute average wind speeds at a height of 10 meter above the surface. In this chapter, hourly-averaged wind speeds are used.

2.3.1 Alternative Two-step Approach

This section presents an alternative approach that extends the spatial snapshot approach [51]. First, the snapshot estimate, $\tilde{Y}(s_0, d, t)$, for the target station, s_0,

is estimated by spatially interpolating wind speeds at neighbor stations, $Y(s,d,t)$ $(s = s_1, s_2, \cdots, s_N)$, through the ordinary kriging as

$$\tilde{Y}(s_0,d,t) = \mathbf{w}^T \mathbf{Y}_{d,t}, \tag{2.21}$$

for each d and t, $d = 1, 2, \ldots, D$, and $t = 1, 2, \cdots, T$. Here $\mathbf{Y}_{d,t}$ is a vector whose i^{th} component is $Y(s_i,d,t)$ and $\mathbf{w} = [w_1, \ldots, w_N]^T$ is the weight matrix, defined as

$$\mathbf{w} = \mathbf{C}^{-1}\mathbf{c} - \frac{\mathbf{C}^{-1}\mathbf{1}\mathbf{1}^T\mathbf{C}^{-1}\mathbf{c}}{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}} + \frac{\mathbf{C}^{-1}\mathbf{1}}{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}}, \tag{2.22}$$

where $\mathbf{C}$ is an $N \times N$ covariance matrix among $\mathbf{Y}_{d,t}$, $\mathbf{c}$ is an $N \times 1$ dimensional covariance between $\mathbf{Y}_{d,t}$ and $\tilde{Y}(s_0,d,t)$ and $\mathbf{1}$ is an $N \times 1$ dimensional vector with 1 elements. Although we present the ordinary kriging, other kriging models (e.g., universal kriging), or GP, can be employed for performing the spatial interpolation.

Once the snapshot estimate is estimated, the time series model at the unmonitored site, s_0, is fitted as follows.

$$\begin{aligned}\tilde{Y}(s_0,d,t) &= \beta_0(s_0,d) \\ &+ \sum_{\ell=1}^{L}\left[\beta_{1,\ell}(s_0,d)\sin\frac{2\ell\pi t}{24} + \beta_{2,\ell}(s_0,d)\cos\frac{2\ell\pi t}{24}\right] \\ &+ \sum_{i=1}^{p}\gamma_h(s_0,d)\tilde{Y}(s_0,d,t-i) + \epsilon(s_0,d,t),\end{aligned} \tag{2.23}$$

where p denotes the model order in the AR process, which is decided based on the Akaike information criterion (AIC). The noise term, $\epsilon(s_0,d,t)$, is assumed to be an independent Gaussian random variable. We estimate parameters using maximum likelihood estimation.

Below this two-step alternative approach is summarized.

- Step 1: At each time point, obtain a snapshot estimate by spatially interpolating the wind measurements collected at neighbor stations using (2.21)-(2.22).
- Step 2: Fit the linear model with the snapshot estimates, using (2.23).

2.3.2 Implementation Results

For evaluating the estimation performance, we divide the 16 stations into two sets: training set (in-sample) and testing set (out-of-sample). The training set includes measurements at 15 stations, representing observational sites, whereas the testing set contains data collected at the remaining station which represents a non-observational site. Using data from the 15 stations in the training set, we estimate the wind speed at the testing station and evaluate the prediction performance by comparing its estimated and measured wind speeds. This procedure is repeated 16 times to get all estimation results for 16 testing stations. Therefore, with four months data, our case study includes 16 stations $\times$ 4 months = 64 testing cases.

We measure the estimation performance with several criteria. First, for evaluating the point estimation capability, root mean square error (RMSE) is used. We also employ the continuous ranked probability score (CRPS) [61, 62], defined as follows, when parameters are estimated in the Bayesian framework [61, 62].

$$CRPS = \frac{1}{DT}\sum_{d=1}^{D}\sum_{t=1}^{T}\Big[\frac{1}{m}\sum_{j=1}^{m}|\hat{Y}^{(j)}(s_0,d,t) - Y(s_0,d,t)| - \frac{1}{2m^2}\sum_{j=1}^{m}\sum_{k=1}^{m}|\hat{Y}^{(j)}(s_0,d,t) - \hat{Y}^{(k)}(s_0,d,t)|\Big] \tag{2.24}$$

where m is the number of posterior samples in the posterior predictive density and $\hat{Y}^{(j)}(s_0,d,t)$ denotes the j^{th} samples. For the alternative approach, CRPS measure presented in [61] is used. The smaller CRPS indicates better performance.

Tables 2.2 and 2.3 summarize the RMSE and CRPS results for four months, respectively, for 16 testing stations where each testing station is considered as a non-observational site. Overall, the estimation performance at testing stations located in the center of monitored stations (e.g., REES and BROW) is generally better than that at boundary stations (e.g., ABER, GAIL and MORT). This is understandable

because the central stations have more informative spatial information from their neighbor stations than boundary stations. Overall, IM3 consistently provides the smallest values in most testing sites in both criteria. For example, on average it generates 14% lower RMSE and 18% lower CRPS, compared to the alternative approach on January.

It is worthwhile to mention that, although IM1 and IM2 generate performance comparable to the alternative approach, they do so with much lower model complexity. In the alternative approach, three parameters need to be estimated to get a snapshot estimate using the ordinary kriging at each time instant. Therefore, it requires $3DT+2L+p+2$ parameters in total for each month. With $D = 31$, $T = 24$, $L = 5$, and $p = 2$, it uses 2,246 parameters. On the contrary, IM1 uses $4(2L+1)+1$ parameters (4 parameters, τ_i, κ_i, $\beta^i_{D,0}$, and ρ_i, for each i and variance parameter, σ^2), whereas IM2 uses $3(2L+1)+1$ parameters (3 parameters, τ_i, κ_i, and δ^2_{2i}, for each i and variance parameter). Thus, IM1 and IM2, respectively, employ 45 and 34 parameters only, which account for about 1.5% and 2% of the alternative's. With such remarkably smaller number of parameters, they lead to the estimation performance similar to the alternative approach. The number of parameters required in IM3 is larger than those in IM1 and IM2, because its spatial parameters differ day-by-day, however, IM3 still reduces the model complexity over the alternative approach by about 69%.

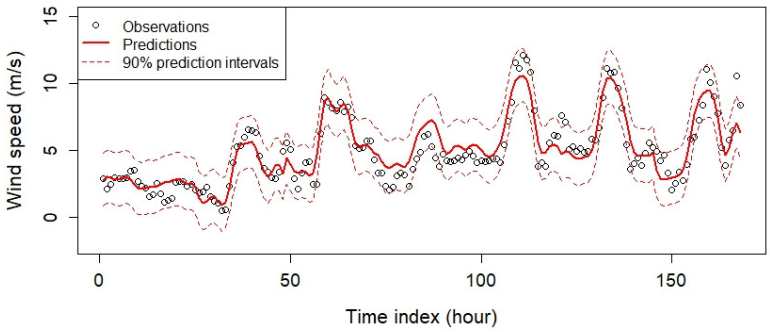
Another advantage of the proposed approach is that it can better quantify the estimation uncertainty. The proposed approach can obtain the posterior predictive density and prediction interval (PI) in the Bayesian framework. Figure 2.4 presents the measured and estimated speeds at a testing station, BROW, for the first week of January, along with PI. The bold central lines denote the predicted values and

Table 2.2: Comparison of RMSEs at 16 testing stations (unit: m/s, SD in the last row represents standard deviation)

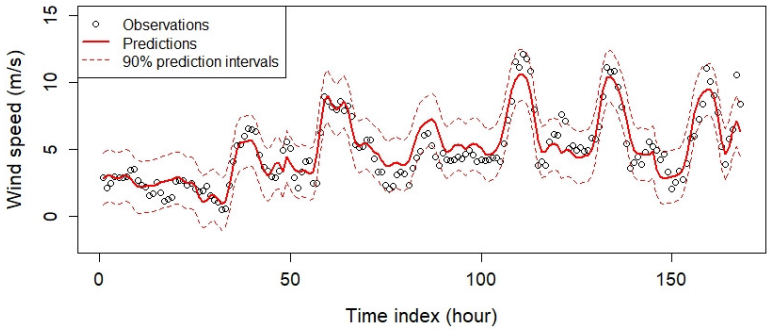
	January				April			
	IM1	IM2	IM3	Two-step	IM1	IM2	IM3	Two-step
ABER	1.28	1.28	1.09	1.33	1.41	1.36	1.08	1.35
REES	0.96	0.96	0.78	0.96	1.07	1.07	0.88	1.08
RALL	1.21	1.21	1.10	1.10	1.23	1.25	1.10	1.23
ANTO	1.20	1.20	0.97	1.18	1.20	1.20	0.96	1.16
SLAT	1.33	1.34	1.23	1.37	1.44	1.46	1.31	1.45
LEVE	1.12	1.11	0.91	1.14	1.12	1.12	0.93	1.13
MORT	1.36	1.37	1.12	1.45	1.42	1.42	1.25	1.52
BROW	0.92	0.92	0.83	0.94	0.97	0.97	0.90	0.98
MALL	0.97	0.97	0.76	0.98	1.01	1.01	0.84	0.99
ODON	1.12	1.12	1.00	1.05	1.44	1.44	1.32	1.35
FLUV	1.48	1.49	1.19	1.15	1.55	1.56	1.30	1.26
PLAI	1.19	1.18	0.94	1.18	1.29	1.31	1.04	1.20
GAIL	1.48	1.47	1.31	1.38	1.63	1.61	1.40	1.52
SEAG	1.21	1.20	1.00	1.19	1.33	1.34	1.09	1.29
LAMS	1.16	1.14	0.96	1.10	1.31	1.31	1.10	1.25
SEMI	1.30	1.31	1.06	1.18	1.37	1.35	1.09	1.22
Average	1.20	1.20	1.01	1.17	1.30	1.30	1.10	1.25
(SD)	(0.17)	(0.17)	(0.15)	(0.15)	(0.19)	(0.19)	(0.17)	(0.16)
	July				November			
	IM1	IM2	IM3	Two-step	IM1	IM2	IM3	Two-step
ABER	0.97	0.98	0.87	0.95	0.91	0.90	0.85	0.97
REES	1.07	1.08	0.92	1.04	0.77	0.78	0.69	0.80
RALL	0.99	0.99	0.87	0.91	1.02	1.04	0.97	1.02
ANTO	0.94	0.94	0.83	0.93	0.96	0.97	0.83	0.96
SLAT	1.20	1.20	1.08	1.17	1.18	1.18	1.16	1.20
LEVE	0.83	0.82	0.70	0.83	0.75	0.75	0.67	0.76
MORT	1.03	1.01	0.92	1.00	1.13	1.13	0.95	1.16
BROW	0.95	0.95	0.90	1.01	0.85	0.84	0.79	0.85
MALL	0.83	0.83	0.71	0.80	0.85	0.86	0.71	0.86
ODON	0.80	0.80	0.76	0.78	0.95	0.98	0.86	0.90
FLUV	1.09	1.10	1.00	1.06	1.35	1.36	1.22	1.20
PLAI	0.93	0.93	0.79	0.88	0.99	0.99	0.79	0.96
GAIL	1.16	1.17	1.10	1.07	1.21	1.21	1.17	1.11
SEAG	0.90	0.89	0.77	0.86	1.01	1.01	0.90	0.99
LAMS	0.96	0.96	0.88	0.97	1.06	1.02	0.86	0.91
SEMI	1.01	1.00	0.85	0.95	1.03	1.03	0.90	0.96
Average	0.98	0.98	0.87	0.95	1.00	1.00	0.90	0.98
(SD)	(0.11)	(0.11)	(0.12)	(0.11)	(0.16)	(0.11)	(0.17)	(0.13)

Table 2.3: Comparison of CRPSs at 16 testing stations (unit: m/s, SD in the last row represents standard deviation)

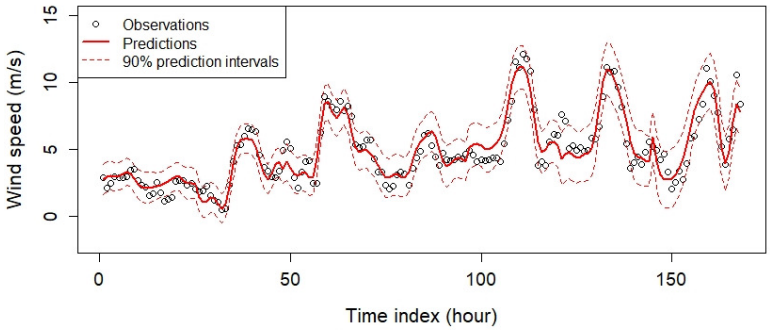
	January				April			
	IM1	IM2	IM3	Two-step	IM1	IM2	IM3	Two-step
ABER	0.71	0.71	0.58	0.77	0.78	0.74	0.59	0.75
REES	0.54	0.54	0.42	0.53	0.60	0.60	0.48	0.59
RALL	0.68	0.68	0.61	0.53	0.68	0.69	0.60	0.68
ANTO	0.66	0.66	0.54	0.68	0.67	0.67	0.54	0.66
SLAT	0.74	0.75	0.68	0.84	0.80	0.81	0.74	0.87
LEVE	0.61	0.61	0.49	0.65	0.62	0.62	0.51	0.64
MORT	0.77	0.76	0.62	0.86	0.80	0.80	0.70	0.90
BROW	0.52	0.52	0.46	0.54	0.55	0.55	0.49	0.56
MALL	0.55	0.55	0.42	0.56	0.57	0.57	0.46	0.56
ODON	0.61	0.61	0.54	0.66	0.79	0.79	0.73	0.79
FLUV	0.86	0.85	0.67	0.66	0.88	0.89	0.74	0.72
PLAI	0.65	0.66	0.52	0.68	0.70	0.71	0.56	0.68
GAIL	0.82	0.82	0.74	0.86	0.93	0.92	0.82	0.96
SEAG	0.68	0.68	0.56	0.73	0.73	0.73	0.59	0.78
LAMS	0.63	0.64	0.54	0.65	0.72	0.73	0.61	0.73
SEMI	0.73	0.72	0.59	0.69	0.75	0.74	0.60	0.72
Average	0.67	0.67	0.56	0.68	0.72	0.72	0.61	0.72
(SD)	(0.10)	(0.10)	(0.09)	(0.11)	(0.11)	(0.11)	(0.11)	(0.12)
	July				November			
	IM1	IM2	IM3	Two-step	IM1	IM2	IM3	Two-step
ABER	0.53	0.54	0.48	0.53	0.51	0.50	0.46	0.55
REES	0.60	0.60	0.52	0.62	0.44	0.43	0.38	0.46
RALL	0.55	0.55	0.48	0.54	0.57	0.58	0.53	0.61
ANTO	0.52	0.52	0.45	0.52	0.53	0.54	0.46	0.57
SLAT	0.68	0.67	0.61	0.71	0.67	0.67	0.66	0.74
LEVE	0.45	0.46	0.38	0.47	0.43	0.43	0.38	0.44
MORT	0.57	0.55	0.51	0.59	0.63	0.63	0.52	0.69
BROW	0.53	0.53	0.50	0.60	0.48	0.47	0.43	0.50
MALL	0.46	0.47	0.39	0.44	0.48	0.48	0.40	0.50
ODON	0.45	0.45	0.43	0.45	0.53	0.53	0.48	0.54
FLUV	0.61	0.61	0.57	0.64	0.77	0.77	0.69	0.71
PLAI	0.52	0.53	0.44	0.51	0.55	0.55	0.44	0.57
GAIL	0.66	0.66	0.64	0.66	0.68	0.68	0.67	0.66
SEAG	0.49	0.49	0.43	0.51	0.57	0.57	0.51	0.61
LAMS	0.53	0.52	0.48	0.58	0.60	0.58	0.49	0.55
SEMI	0.56	0.57	0.47	0.56	0.57	0.58	0.57	0.57
Average	0.54	0.54	0.49	0.56	0.56	0.56	0.50	0.58
(SD)	(0.07)	(0.07)	(0.07)	(0.08)	(0.09)	(0.09)	(0.10)	(0.09)



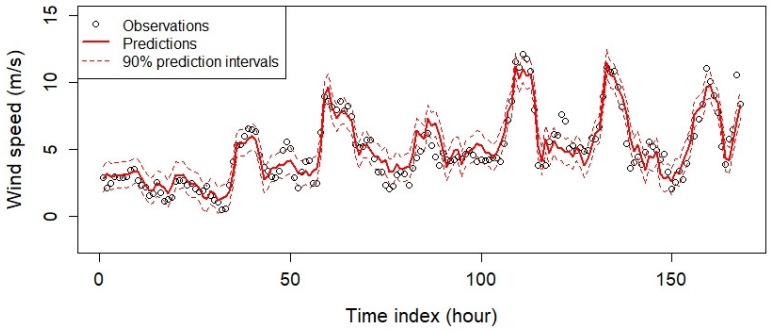
(a) IM1



(b) IM2



(c) IM3



(d) Two-step approach

Figure 2.4: Comparison of observed and predicted wind speeds and prediction intervals at the testing station, BROW, in January

the dotted upper and lower lines denote the 90% PIs. It shows that most estimated speeds in IM1, IM2, and IM3 belong to the PIs.

However, it is difficult, if not possible, to accommodate all uncertainties in the alternative approach, because it characterizes the spatial and temporal correlation separately through the two-step procedure. Therefore, we alternatively treat the snapshot estimates as real values and build the PI with the model in (2.23). Figure 2.4(d) shows that the resulting PIs are unduly narrow and thus, several data points are located outside the intervals, indicating underestimated uncertainties. As a result, the coverage rate of the alternative model is much lower than ours. Here, the coverage rate implies the ratio of the number of estimates within PIs to the total number of estimates. Ideally, the coverage rate should be close to the nominal rate. For example, in January, the average coverage rate of the alternative approach remains at 67.6% for the 90% PI, whereas the coverage rates from IM1, IM2, and IM3 are 89.3%, 89.5%, and 86.7%, respectively, which are close to the nominal rate. It is also worthwhile to mention that the PIs from our approach are wider, because it fully quantifies uncertainties for estimating spatial and day-to-day variability.

To further assess probabilistic estimation performance, reliability diagram [27] is employed. To construct the reliability diagram, an indicator variable that compares an actual speed, $Y(s_0, d, t)$ with its α-quantile forecast, $\hat{Y}^{(\alpha)}(s_0, d, t)$, for $0 \leq \alpha \leq 1$ is obtained as

$$I^{(\alpha)}_{s_0,d,t} = \begin{cases} 1, & \text{if } Y(s_0, d, t) \leq \hat{Y}^{(\alpha)}(s_0, d, t) \\ 0, & \text{if } Y(s_0, d, t) > \hat{Y}^{(\alpha)}(s_0, d, t), \end{cases} \tag{2.25}$$

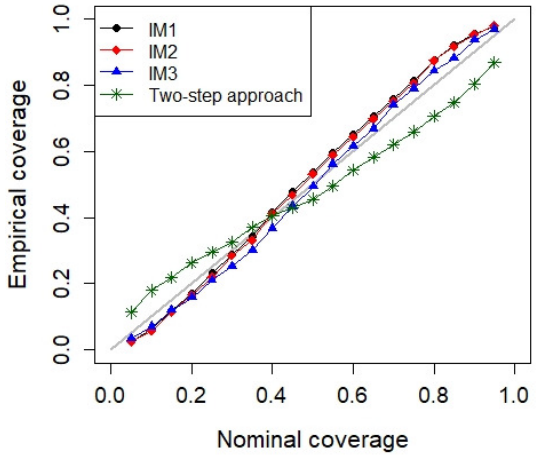
Then, similar to the PI coverage, the empirical coverage in the reliability diagram is obtained as

$$\hat{a}^{(\alpha)}_{s_0} = \frac{1}{DT} \sum_{d=1}^{D} \sum_{t=1}^{T} I^{(\alpha)}_{s_0,d,t}. \tag{2.26}$$

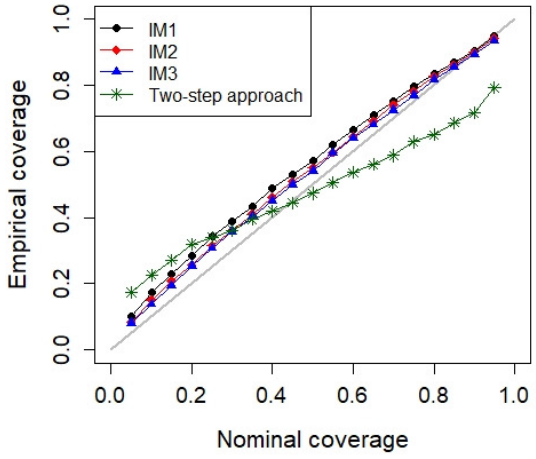
The reliability diagram compares the empirical coverage with the nominal coverage. The empirical coverage becomes close to the nominal coverage, α, when the probabilistic estimation is performed appropriately. We compare the empirical coverage at nominal levels from 5% to 95% on increments of 5% in Figure 2.5. The average reliability diagram in Figure 2.5(c) is constructed by taking the average of empirical coverage from all 16 testing stations. The empirical coverage from IM1, IM2, and IM3 align with the diagonal line in all cases, while those from the two-step approach deviate from the diagonal line. This result demonstrates a stronger probabilistic assessment capability of the proposed approach over the alternative one.

In summary, although the proposed integrative and alternative two-step approaches provide comparable point estimation capability, our approach quantifies wind variability better so its estimated density is more accurate. Our strong probabilistic assessment performance is mainly due to the fact that our approach can capture different types of uncertainties arising from spatial and diurnal variations in an integrative way.

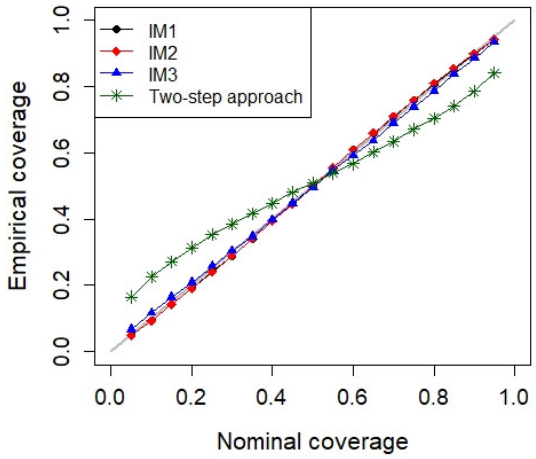
Among the studied models, the performance of IM1 and IM2 were comparable in most cases. This result coincides with our observation in Figure 2 where the daily pattern is similar on consecutive days. Therefore, when the diurnal patterns do not rapidly change, either the AR formulation in (2.11) or random effect formulation in (2.12) would provide similar results. When the diurnal pattern changes considerably, we suggest the random effect formulation used in IM2. Regarding the spatial correlation, it has been known that the dominating wind direction substantially affects the correlation structure [59]. Therefore, IM3 would perform better than IM1 and IM2 when wind direction varies substantially even during the same month.



(a) Testing station: BROW



(b) Testing station: MORT



(c) Average

Figure 2.5: Reliability diagram in January

2.4 Summary

This chapter develops a probabilistic model for assessing wind resource by characterizing the spatial and temporal correlations through the model parameters. Specifically, the model parameters are treated as latent spatially- and daily-varying random processes. Such collective treatment enables the proposed integrative approach to provide compelling capabilities for evaluating the wind variability at non-observational locations.

A case study with west Texas Mesonet data demonstrates that the proposed approach is capable of fully quantifying wind variations, which provides insights for selecting wind farm locations. The proposed approach can be applied to the wind resource assessment in other regions where time-series of wind measurements at spatially dispersed locations are available.

The proposed approach can be also extended to other environmental factors, such as temperature and air pollutant level, which exhibit spatially and daily varying patterns. Finally, the proposed framework can be extended to a large-scale analysis. In this case, some approximation methods (e.g., variational inference) may be needed to reduce the computational time for estimating model parameters.

CHAPTER III

On the Long-term Density Prediction of Peak Electricity Load with Demand Side Management in Buildings

3.1 Introduction

Accurate electricity load forecasting is critical for reliable operations and long-term planning of the electric power grid and its infrastructure systems [63]. 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.

This chapter 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 chapter focuses on peak electricity demands, as peak demands often define the required capacity of generation and transmission systems [64, 65]. The purpose of this chapter 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 [16, 66, 67]. A recent review on the relationship between electricity demand and weather conditions [68] 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 [69, 70]. 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 [66]. Recent literature suggests, however, that the historical weather trends and temperature extremes change over time [71]. 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 [72] concluded that there very likely had been an increase in the

frequency of extremely high temperatures. The IPCC Special Report on Extremes in 2012 [73] and the IPCC Fifth and Sixth Assessment Reports in 2013 and 2018 [74, 75], 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 [76] 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 [71].

Along with weather conditions, electricity peak loads also largely depend on the socio-economic factors, including the population size and buildings‘ electricity use patterns [77]. 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 [78].

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 [79], 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 [80].

The objective of this chapter 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 [71]. 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 [81]. 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 chapter 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 chapter is organized as follows. Section 3.2 reviews relevant studies in the literature. Section 3.3 describes the datasets used in this chapter. Section 3.4 introduces the proposed long-term daily peak demand density prediction

method. Section 3.5 presents a case study for evaluating the density prediction performance of the proposed approach and providing long-term densities. Finally, Section 3.6 summarizes the findings, implications, and future research.

3.2 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 [82, 83], time series methods including autoregressive, autoregressive moving average, autoregressive integrated moving average models and their variations [18–20]. Different types of neural networks and their variations [10, 21–24], as well as other machine learning techniques and hybrid or ensemble models [84–86], 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 [87, 88]. 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 [89] 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.* [90] 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.* [91] 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.* [92] 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 [90–92] provide the point estimates of the long-term peak demands and do not fully address the uncertainty quantification.

Hyndman *et al.* [93] 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.* [94] 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 [95–98]. Different from the aforementioned studies that use historical data only, Trotter *et al.* [88] propose the probabilistic long-term electricity demand forecasting using the multiple linear regression with GCM outputs as wells 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.* [88] is one of the few studies that consider both, however it does not address the possible bias of GCM projections. This chapter 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.

3.3 Datasets

This section describes the datasets used in this chapter, including GCM projections, actual temperature measurements, population, and participation rate in the considered DSM program. As the focus of this chapter 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 [99]. 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 [100]. 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 [74, 75]. 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 [93, 94].

Next, the daily peak demand is obtained from the Electric Reliability Council of Texas (ERCOT) [101]. 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 3.1 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 [102].

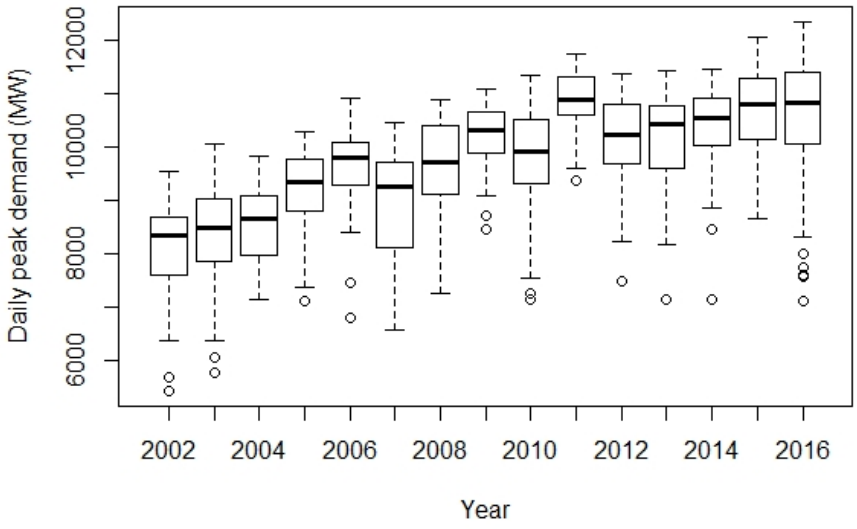


Figure 3.1: Box plots of actual daily peak demands during the summer from 2002 to 2016 in the south-central region of Texas

Electricity usage is also largely affected by socio-economic factors, such as gross domestic product (GDP), industrial production, and the population density [77, 93]. Among them, this chapter uses the population as one of the major factors for the daily peak demand modeling. The population data consists of the yearly population estimates for every county in the south-central region of Texas from 2002 to 2016, obtained from the United States Census Bureau [103]. Figure 3.2 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

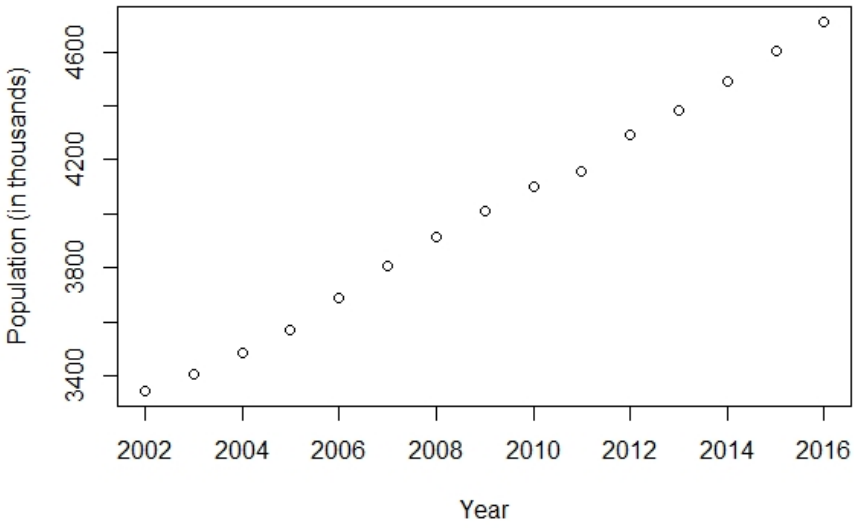


Figure 3.2: Total population growth in 24 counties in the south-central region of Texas from 2002 to 2016

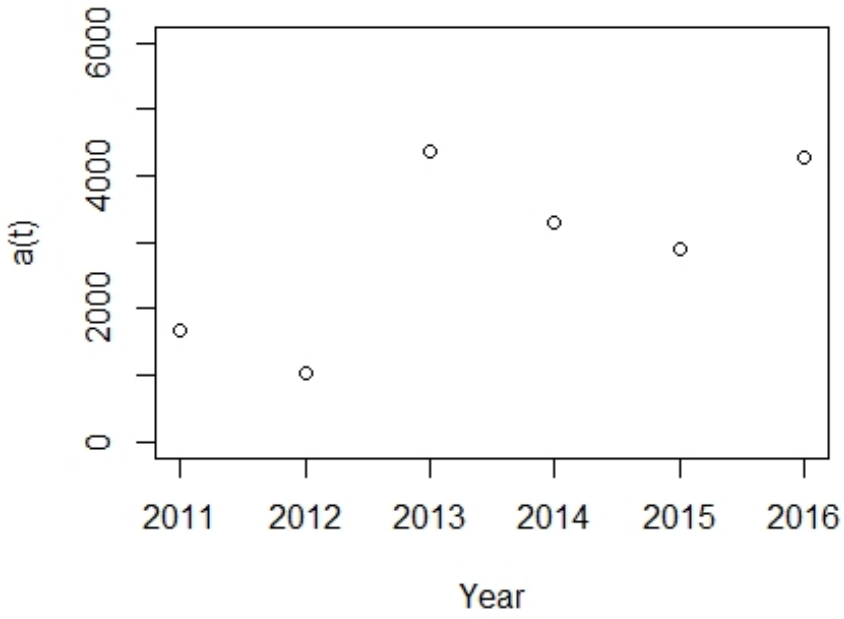
the population over time in this region. It should be noted that other economic factors can be additionally considered in our modeling efforts, however, they are often positively correlated with the population growth (or decay) trends. As such we consider the population to represent the economic condition 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 [1]. For example, Figure 3.3 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,

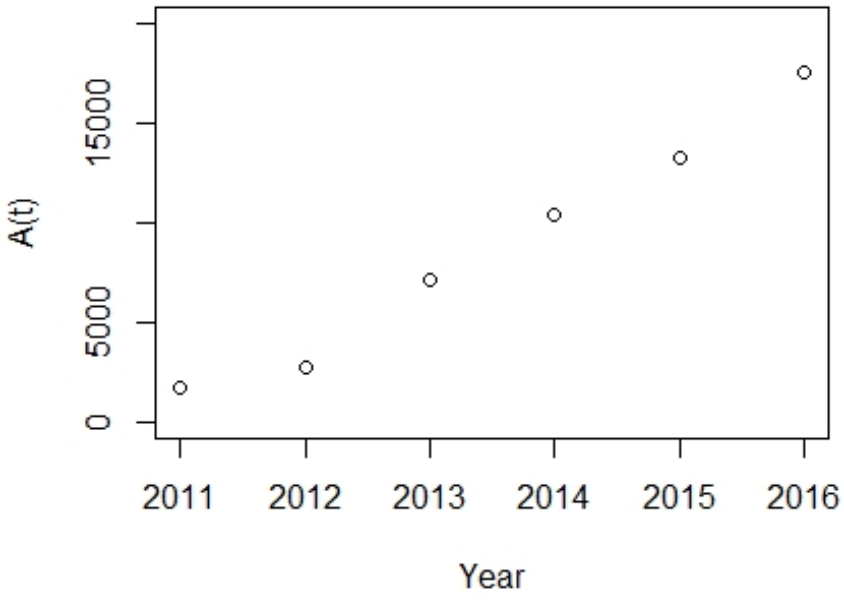
one residential house adopts one device. Although there are year-to-year variations, overall we observe an increasing trend. The original Austin Energy report includes demand saving data since 2007, but we note that the number of residential participants from 2007 to 2010 exhibits a decreasing trend. 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 [1] and the U.S. Energy Information Administration (EIA) reports [79].

3.4 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 3.4 shows the overall framework of the proposed approach.



(a) Number of participants in year t



(b) Number of cumulative participants up to year t

Figure 3.3: The number of residential participants in Austin Energy's demand saving program [1]

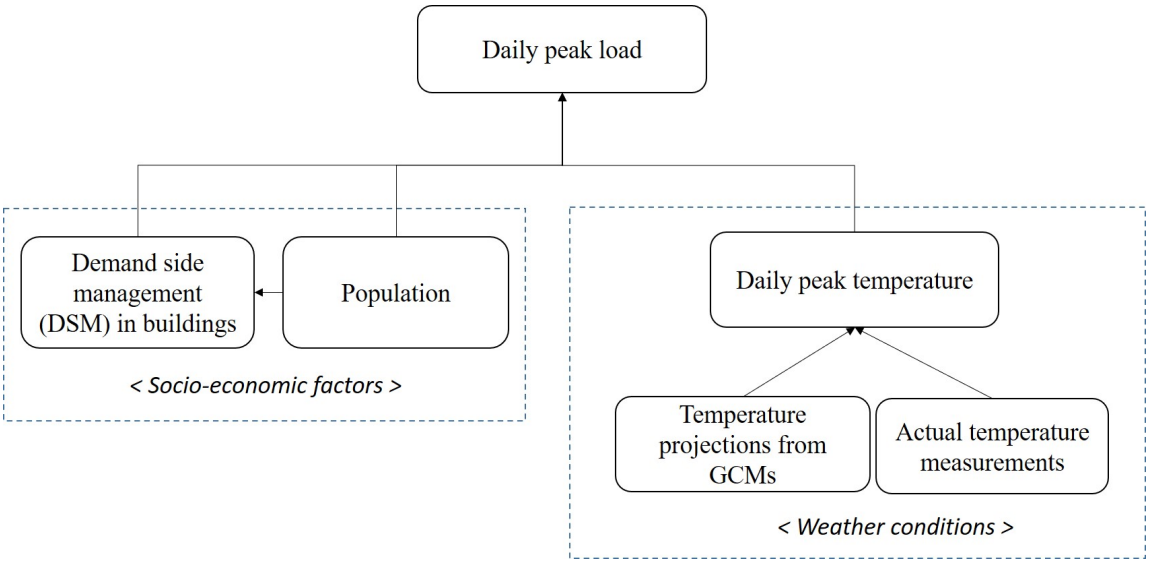


Figure 3.4: Overall framework of the proposed approach

3.4.1 Modeling Long-term Daily Peak Temperature

The daily peak temperatures represent the block maxima. Based on the limit theorem for block maxima [104, 105], 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

$$
\begin{aligned}
y_{d,t} &\sim f_t(y|x) \\
&= GEV(\mu_t, \sigma_t, \eta_t) \\
&= \begin{cases} \frac{1}{\sigma_t}\Big[1+\eta_t\Big(\frac{y-\mu_t}{\sigma_t}\Big)\Big]^{-(1/\eta_t+1)} \exp\Big\{-\Big[1+\eta_t\Big(\frac{y-\mu_t}{\sigma_t}\Big)\Big]^{-1/\eta_t}\Big\}, & \text{for } \eta_t \neq 0 \\ \frac{1}{\sigma_t}\exp\Big\{-\Big(\frac{y-\mu_t}{\sigma_t}\Big) - \exp\Big\{-\Big(\frac{y-\mu_t}{\sigma_t}\Big)\Big\}\Big\}, & \text{for } \eta_t = 0 \end{cases} \qquad (3.1)
\end{aligned}
$$

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 [45]. 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 [45].

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

$$x_t = \frac{1}{62 \times 42} \sum_{d=1}^{62} \sum_{j=1}^{42} g_{d,t}^{(j)}. \tag{3.2}$$

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 [45], as follows.

$$\begin{aligned} \mu_t &= \alpha_0 + \alpha_1 x_t, \\ \sigma_t &= \beta_0 + \beta_1 x_t, \\ \eta_t &= \eta_0, \end{aligned} \tag{3.3}$$

where α_0, α_1, β_0, β_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 (3.3) 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., $\mathcal{D}_y = \{y_{d,t}, d = 1, 2, \cdots, D, t = 1, 2, \cdots, T\}$. 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.

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

for $\eta_t \neq 0$, where $\mu_t, \sigma_t,$ and $, \eta_t$ are formulated in (3.3), $D(= 62)$ 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 [105]. Similarly, we also obtain the MLE estimates for $\eta_t = 0$ and between two, we choose the estimates that provides higher log-likelihood values.

3.4.2 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 [106, 107]. 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. 3.2. 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 [108]. 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 [109] as

$$p_t \sim g_t(p) = N\Big(\frac{a}{1+\exp\{-(t-b)/c\}}, \sigma_p^2\Big), \tag{3.4}$$

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 (3.4) 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 (3.4), we use Levenberg-Marquardt nonlinear least-squares algorithm [110, 111]. 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 [108]. This unduly large value was ob-

tained 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 3.5 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.

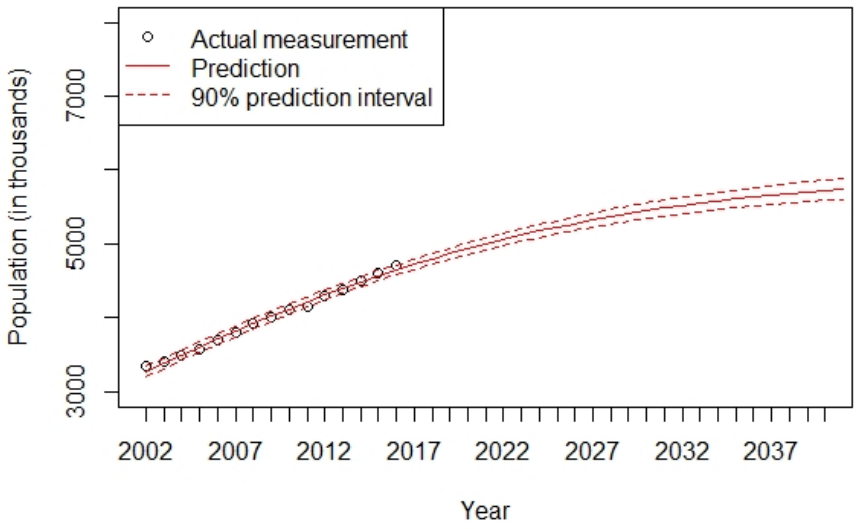


Figure 3.5: Population prediction in the south-central region of Texas from 2002 to 2040

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 mod-

els, 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 [81]. 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 [81]. 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 [112].

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

$$M(t) = d_s \cdot h_t, \tag{3.5}$$

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 this chapter, 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 3.5.

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 [81] as

$$\frac{\ell(t)}{1 - L(t)} = m + nL(t), \tag{3.6}$$

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 (3.6), the portion of the participants in year t becomes

$$\ell(t) = \frac{dL(t)}{dt} = (m + nL(t))(1 - L(t)) = m + (n - m)L(t) - nL(t)^2. \tag{3.7}$$

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 [81] and get

$$L(t) = \frac{1 - \exp\{-(m + n)t\}}{1 + \frac{n}{m}\exp\{-(m + n)t\}}, \tag{3.8}$$

$$\ell(t) = \frac{\frac{(m+n)^2}{m}\exp\{-(m + n)t\}}{\left(1 + \frac{n}{m}\exp\{-(m + n)t\}\right)^2}, \tag{3.9}$$

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

$$a(t) = \ell(t)M(t) = \frac{\frac{(m+n)^2}{m}\exp\{-(m + n)t\}}{\left(1 + \frac{n}{m}\exp\{-(m + n)t\}\right)^2} \cdot d_s h_t, \tag{3.10}$$

and

$$A(t) = L(t)M(t) = \frac{1 - \exp\{-(m+n)t\}}{1 + \frac{n}{m}\exp\{-(m+n)t\}} \cdot d_s h_t, \tag{3.11}$$

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

$$s(t) = C \cdot A(t), \tag{3.12}$$

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 [1]. For C, we employ the average saving per residential participant and get $C = 0.8765kW$. 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 [103].

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 (3.10) and (3.11). 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 (3.4), 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 [1], 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 [79]. Based on this fact, we approximate the demand saving from EE programs by multiplying 0.5 to the demand saving from DR programs.

3.4.3 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

$$z_{d,t}|y_{d,t}, p_t \sim h_t(z|y_{d,t}, p_t) = GEV(\mu'_{d,t}, \sigma'_{d,t}, \eta'_t), \tag{3.13}$$

with

$$\begin{aligned} \mu'_{d,t} &= \alpha'_0 + \alpha'_1 y_{d,t} + \alpha'_2 p_t + \alpha'_3 t, \\ \sigma'_{d,t} &= \beta'_0 + \beta'_1 y_{d,t} + \beta'_2 p_t + \beta'_3 t, \\ \eta'_t &= \eta'_0. \end{aligned} \tag{3.14}$$

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 (3.14). 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 [113].

Similar to the procedure in (3.4), the parameters in (3.14) 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 [1] and EIA reports [79]. Figure 3.6 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 3.1. As the DSM efforts become more active, demand saving became more significant. So, in later years, the box plots in Figure 3.6 became more shifted upward, compared to those in Figure 3.1. We use the adjusted daily peak demand data to estimate the parameters in (3.14).

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

$$u_t(z) = \int h_t(z|y_t, p_t) g_t(p) f_t(y|x_t) dp dy, \tag{3.15}$$

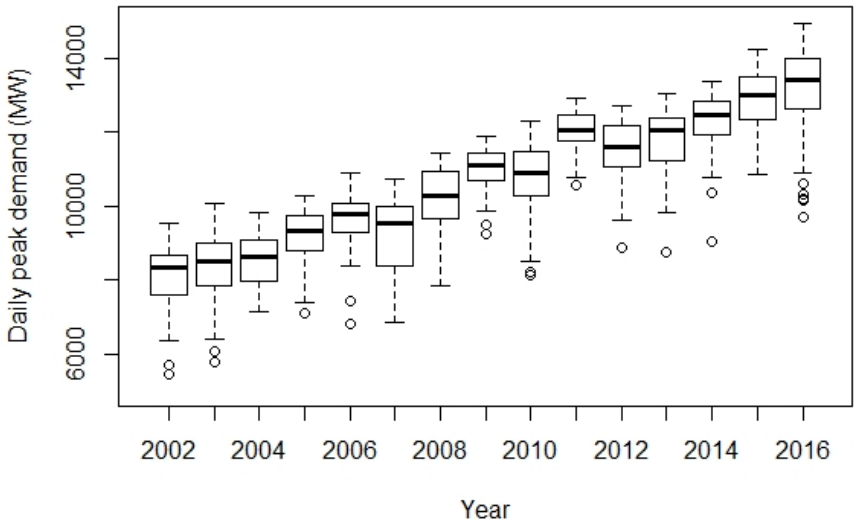


Figure 3.6: Adjusted Box plots of daily peak demands without DSM efforts in buildings during the summer in the south-central region of Texas

where $h_t(z|y_t, p_t)$ is the conditional probability density function (pdf) of the daily peak demand in year t in (3.13), $g_t(p)$ is the pdf of the population in year t in (3.4), and $f_t(y|x)$ is the pdf of the daily peak temperature in year t in (3.1), given the grand ensemble temperature projection x_t from GCMs in (3.2).

The predictive density $u_t(z)$ in (3.15) does not take a closed-form. Thus, we determine the density using Monte Carlo sampling [114]. 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 3.4.2, from z_t.

3.5 Case Study

3.5.1 Validation

We use temperature and electricity usage data collected in south-central Texas in 2002-2016, as discussed in Section 3.3. 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 3.7 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 non-stationary temperature pattern. In 2011, the actual histogram somewhat deviates from its predicted one. There were extraordinary extreme heatwaves in 2011 [115]. 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.

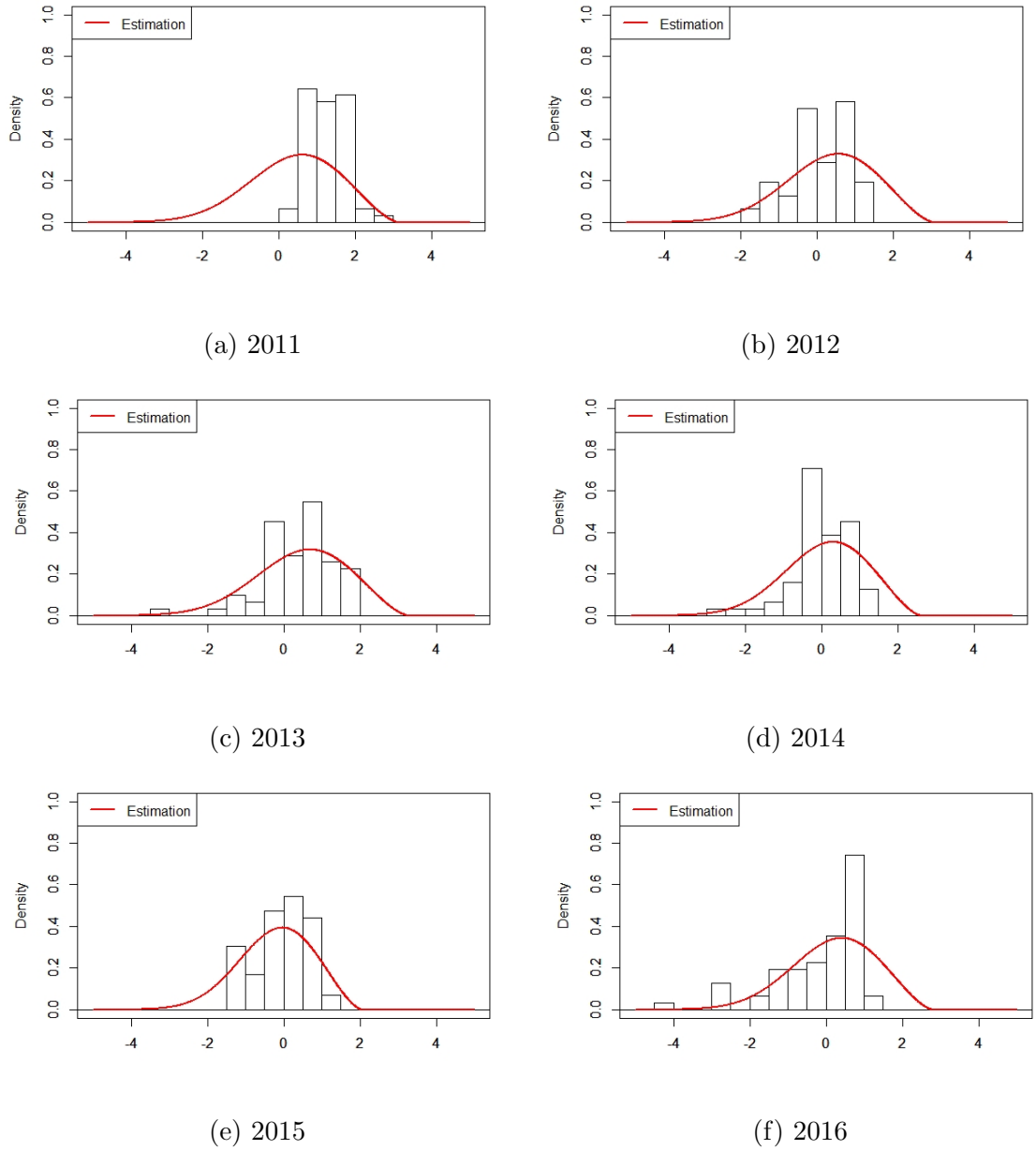
Next, we compare the histogram of the actual daily peak demand and its density estimation (red curve) $h_t(z|y_t, p_t)$ in Figure 3.8. The density estimation for the daily peak demand successfully matches the histogram of actual data across all years.

Both comparisons in Figures 3.7 and 3.8 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).

We further compare the peak load density estimation performance of the proposed approach with the alternative the trend-based approach discussed in [90]. AlRashidi *et al.* [90] 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 3.9 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.



(a) 2011

(b) 2012

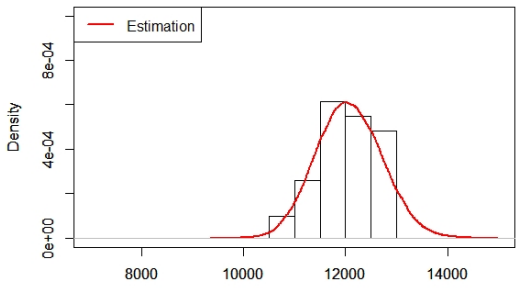
(c) 2013

(d) 2014

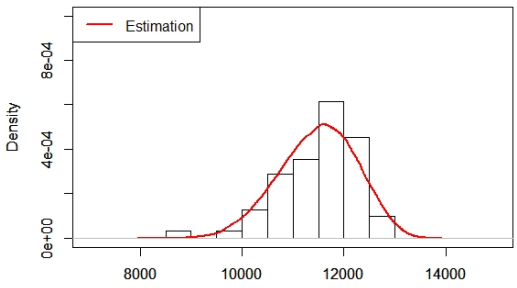
(e) 2015

(f) 2016

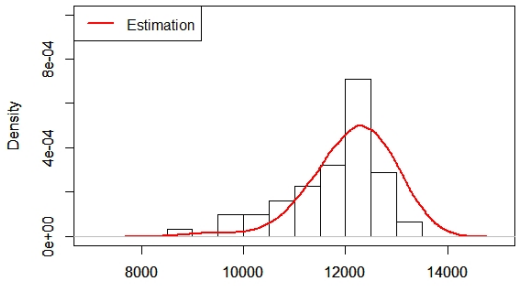
Figure 3.7: Comparison between empirical and estimated densities of daily peak temperature in the testing set



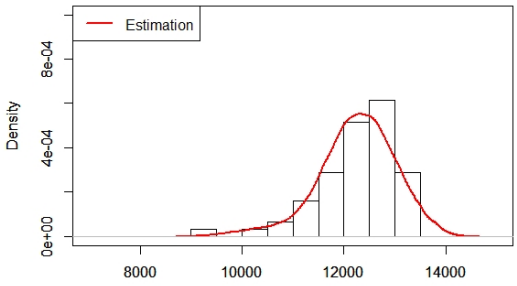
(a) 2011



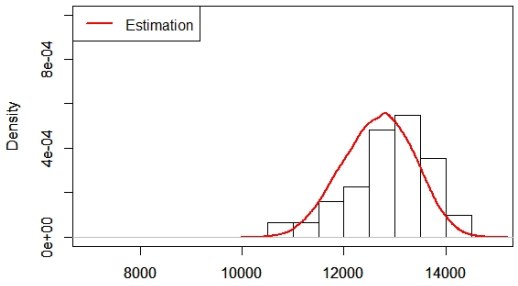
(b) 2012



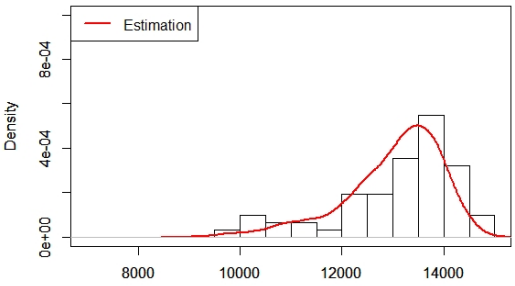
(c) 2013



(d) 2014

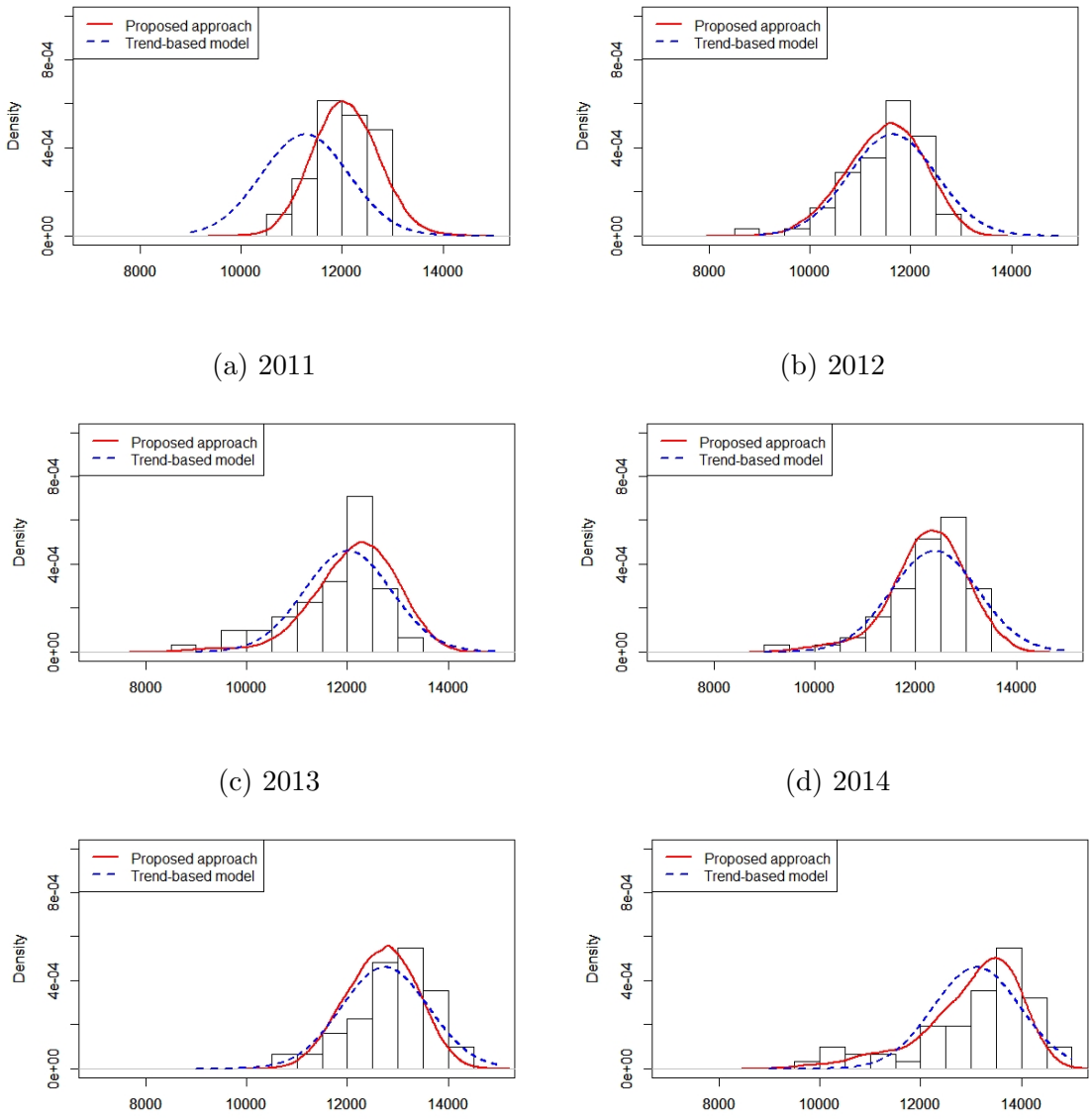


(e) 2015



(f) 2016

Figure 3.8: Comparison between empirical and estimated densities of daily peak demand in the testing set



(a) 2011

(b) 2012

(c) 2013

(d) 2014

(e) 2015

(f) 2016

Figure 3.9: Comparison of the estimated densities of daily peak load from the trend-based and proposed approach in the testing set

3.5.2 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 [116, 117]. In this chapter, 20 years was chosen to illustrate the demand change pattern during the typical planning horizon. First, Figure 3.10 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 3.11 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 3.11a. The predictive mean and median of daily peak demands in 2040 are about 22.4 GW, if there would be no DSM efforts. Figure 3.11b 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.

To account for the effect of the demand saving from DSM programs, we consider different values of d_s in (3.5). Figure 3.12 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 3.13 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 signifi-

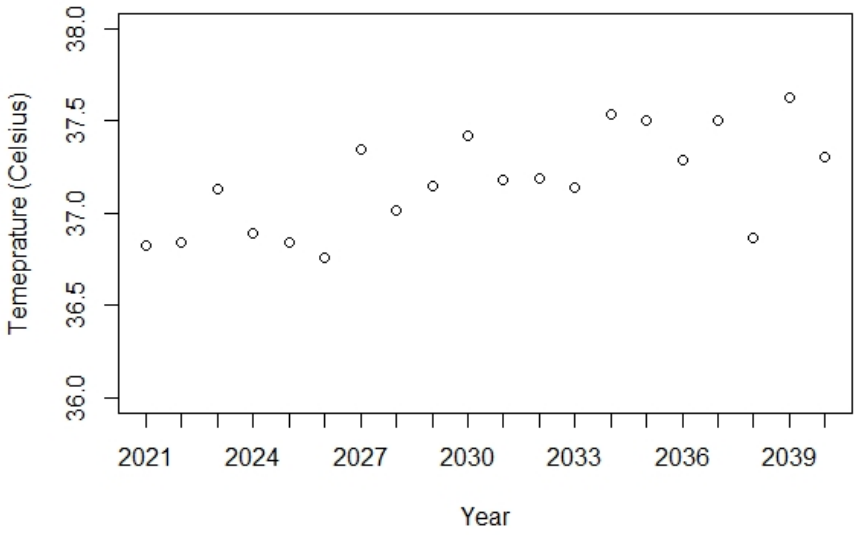
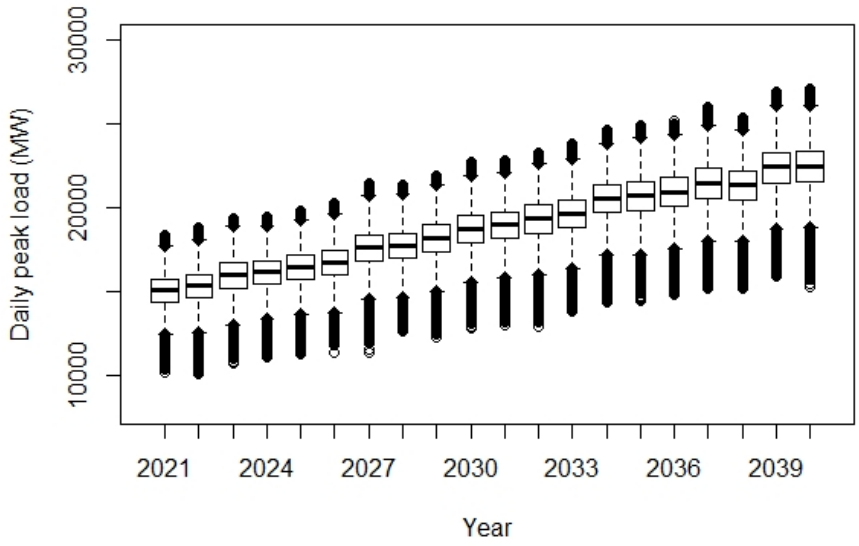


Figure 3.10: Grand ensemble average of daily peak temperature projections from GCMs

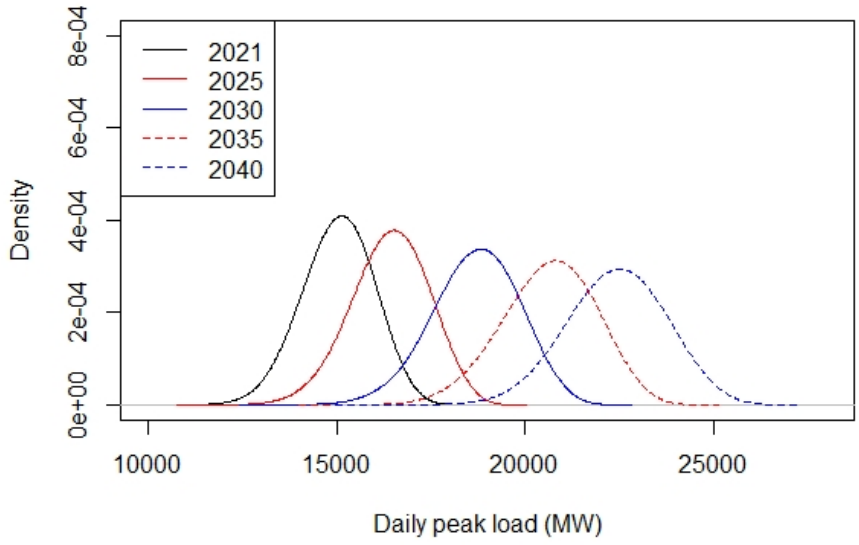
cantly different with different values of d_s, so the corresponding densities are highly overlapped. This is due to the fact that a small number of participants join the demand saving program in early years. As more people join DSM programs, the effect of the demand saving becomes more substantial so 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.

Table 3.1 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 3.14 shows the predictive mean trajectories in a range of d_s value.



(a) Box plot



(b) Density plot

Figure 3.11: Box plot (top) and density plot (bottom) of predicted daily peak demands in 2021-2040 without buildings‘ demand saving

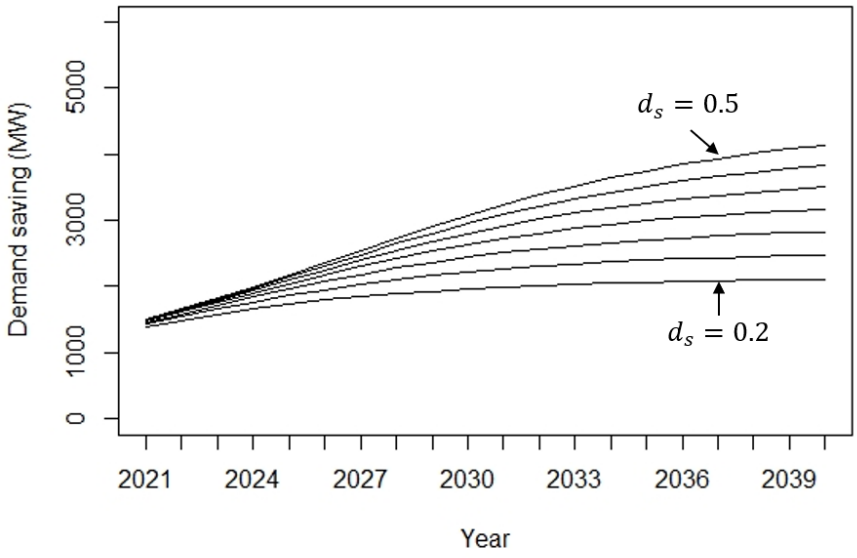
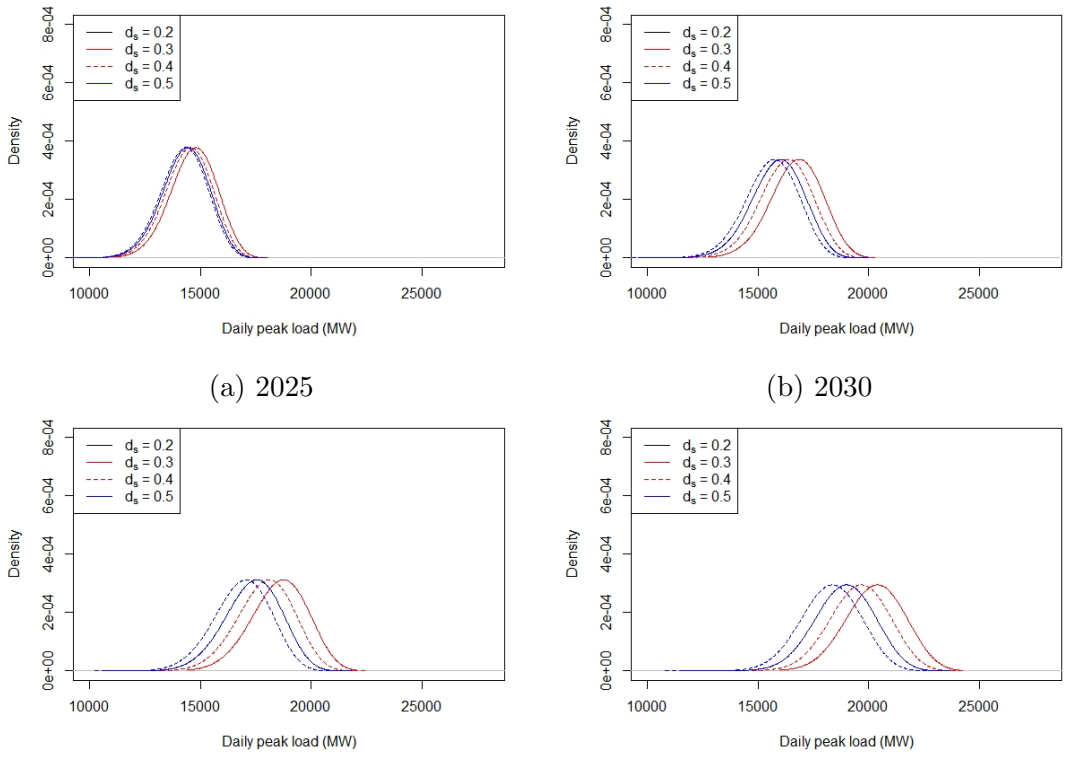


Figure 3.12: Predicted demand saving from DR programs from 2021 to 2040 in the south-central region of Texas

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.



(a) 2025

(b) 2030

(c) 2035

(d) 2040

Figure 3.13: Daily peak density prediction with and without demand saving from DR programs

Table 3.1: 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)

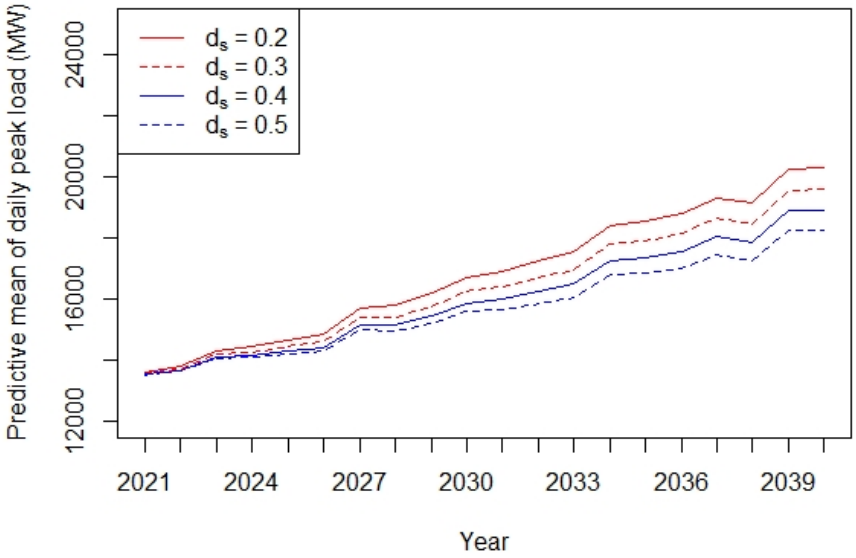


Figure 3.14: Predictive mean of the daily peak demand with and without demand savings participation

3.6 Summary

This chapter 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.

DSM activities for buildings could substantially affect daily peak loads, as we observe in Section 3.5. 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 will possibly 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. Other socio-economic variables, such as gross domestic product, industrial production, can be additionally considered to enhance the prediction capability in our future study.

CHAPTER IV

Spatio-temporal Bias Correction in Numerical Weather Prediction Model with Application to Urban Temperature Modeling During Heat Wave Events

4.1 Introduction

Ambient environmental condition affects many aspects of human life. For example, the energy consumption from residential and commercial buildings, consisting of a significant portion of the U.S. electricity demand, is largely affected by weather conditions [118]. To understand and forecast environmental conditions, physics-based models such as the WRF model are often employed [119]. The WRF is a mesoscale NWP model that also has been used to study climate change with regionally downscaled versions. However, these simulations do not completely account for local climate characteristics [71]. To better estimate localized weather conditions in urban areas, studies in [120, 121] have coupled WRF with a UCM that parameterizes urban canopy features such as building and street properties [119].

This chapter is, in particular, concerned with the spatio-temporal temperature characterization within the WRF/UCM model during extreme heat wave events. The excessive electricity consumption during extreme heat waves can result in power

blackouts and brownouts, which can put occupants of the buildings at risk [122]. Extreme heat waves can also cause adverse effects on public health, seriously affecting vulnerable populations such as seniors, infants, young children, and people with chronic health problems [122]. According to the study reported in [123], a $1°C$ increase in temperature during a heat wave can result in a 4.5% increase in mortality risk. Even worse, the so-called urban heat island effect produces additional heat stresses due to interactions between the built-environment and the atmosphere, causing higher temperatures in highly urbanized areas as compared to their surrounding areas [71].

Although UCM coupled with WRF, referred to as WRF/UCM in this chapter, can provide more accurate localized weather profiles, substantial discrepancy is possible between actual and WRF/UCM-predicted temperature. To illustrate, Figure 4.1 shows a network of ground-based weather stations around Austin, Texas, where each black circle represents the location of a weather station. Figure 4.2 compares actual temperatures (black curves) at these locations with their corresponding WRF/UCM projections (red curves) during three heat wave events. A large discrepancy is observed during the daytime on these heat wave days. Interestingly, WRF/UCM tends to underestimate the temperature from the late afternoon to the evening hours. This is mainly because heat is stored in thermal mass of buildings and asphalt paved roads, which causes the ambient temperatures to decrease more slowly. It appears that WRF/UCM cannot fully account for such local climate characteristics.

Such discrepancies likely arise due to inappropriate settings in the WRF and UCM models, e.g, inaccurate initial and/or boundary conditions and model formulation in WRF, and/or incorrect parameterization in UCM [124–126]. In particular, UCM employs several parameters such as land use, land cover information, percentage

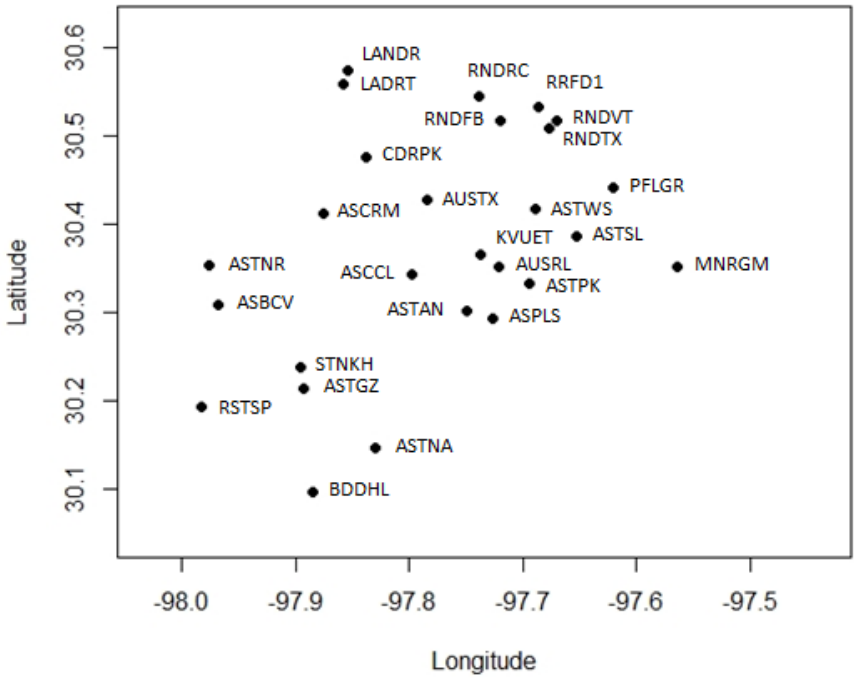
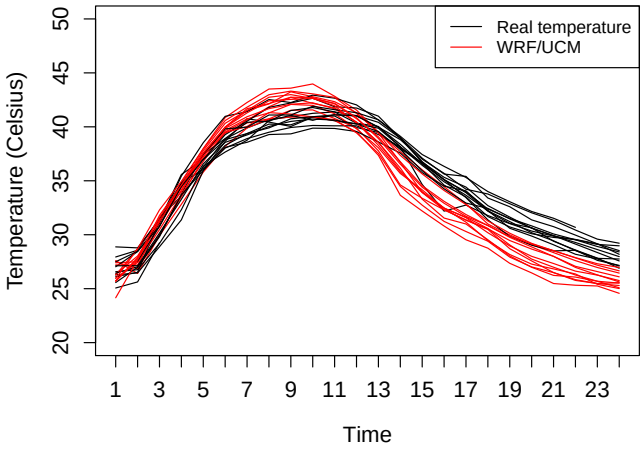


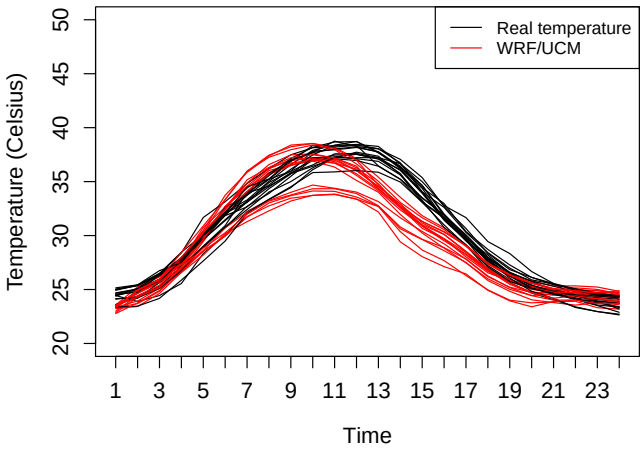
Figure 4.1: Ground-based weather stations in Central Texas

of impervious surface, building dimensions, and surface albedo. When information on these parameters is not available or is unknown, *educated guesses* from domain experts are often employed, but inappropriate assumptions can cause substantial deviations in model results from actual climate conditions [126]. One remedy is to adjust the parameter values iteratively, but doing so can be challenging and cumbersome when a large number of such parameters is involved. Furthermore, running WRF/UCM with each new parameter setting is computationally demanding.

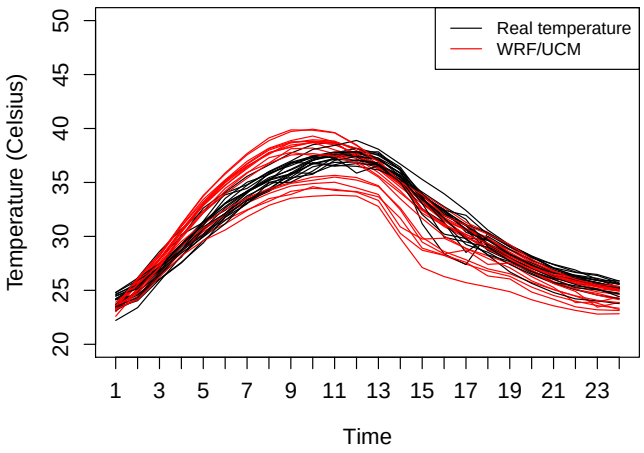
The discrepancy between measured ambient conditions and the output from numerical weather models is referred to as *bias* in the literature [127] (note that, in the present chapter, the definition of bias is different from that in the statistical literature.) To explain the bias/discrepancy, Figure 4.3 shows bias patterns on 8/8/2013 at three locations: ASTSL, ASTPK, and BDDHL identified in Figure 4.1. As ob-



(a) 8/28/2011



(b) 8/8/2013



(c) 7/24/2017

Figure 4.2: Discrepancy (or bias) between actual temperature and WRF/UCM prediction

served in Figure 4.2, WRF/UCM output suggests a clear bias pattern over the day at each location. Further, we see that bias patterns at two nearby locations, ASTSL and ASTPK, are more similar to each other but somewhat different from the pattern seen at BDDHL; this suggests possible spatial correlation in bias patterns. In summary, WRF/UCM yields unevenly distributed systematic bias patterns both spatially and temporally.

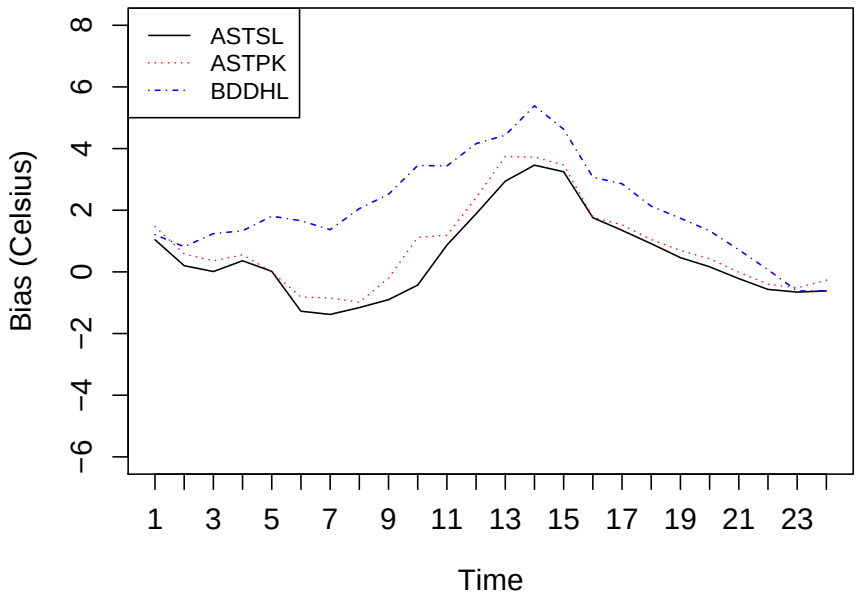


Figure 4.3: WRF/UCM bias patterns at three stations, ASTSL, ASTPK, and BDDHL on 8/8/2013

In the literature, several approaches have been proposed to reduce bias. Bias correction in the physics-based forecasting models is also called post-processing or forecast calibration [128]. The most common approach is the so-called model output statistics (MOS) approach [129]. The basic idea of MOS is to identify a relationship between actual weather observations and the corresponding NWP projections. [129] employed a linear regression model to quantify the relationship between two data sources in applications related to the probability of precipitation, surface wind,

maximum temperature, conditional frozen precipitation, and cloud cover. Similarly, [130] used linear regression to calibrate the bias in wind speed forecasts from NWP and predicted hourly and sub-hourly wind speed and wind power output.

Note that, in the aforementioned studies, the climate model bias is assumed to be stationary, implying that the relationship between numerical model outputs and measurements is assumed not to change over time and space. In other words, they adjust bias in NWP projections by the same amount, regardless of the time of day and location. Clearly, this assumption appears to be unjustified, given what is observed in Figures 4.2 and 4.3.

Another approach, called the delta change method, computes the average bias in historical projections from NWP models and adds this to projections from NWP models [131, 132]. [131] employed the delta method to calibrate bias in temperature prediction from a regional climate model (RCM). [132] calibrated the bias in monthly mean temperature and precipitation predictions from an RCM. [133] computed the average bias at each hour to accommodate a time-varying bias pattern. Similar to the MOS approach, the delta change method does not consider temporal correlation in the bias. Further, it does not address spatially-varying bias patterns.

To address these issues, we propose a new bias correction method that leverages the spatial and temporal bias dependence for more accurate and reliable post-processing of WRF/UCM. By utilizing information collected at geographically dispersed locations in both WRF/UCM projections and actual measurements, we aim to learn heterogeneous bias pattern in a collaborative manner. The main contribution of this chapter is three-fold:

(1) Unlike the aforementioned studies that assume stationary assumptions, our approach captures nonstationary bias characteristics which are uneven across

a day-long period and over space.

(2) The proposed approach fully harnesses spatio-temporal uncertainties and generates probabilistic bias corrections.

(3) It provides an integrative framework that can produce bias-corrected estimates even in places where WRF/UCM projections are unavailable, different from existing approaches that require WRF/UCM projections for all target places.

We conduct a case study using actual and WRF/UCM predicted temperatures in the central Texas region during three extreme heat wave events. Our implementation results demonstrate that the proposed approach is capable of characterizing time- and space-varying bias patterns. The advantage of the proposed approaches is reflected by better point and probabilistic bias correction performance, compared to the WRF/UCM projections and alternative approaches.

The remainder of this chapter is organized as follows. Section 4.2 introduces the proposed bias correction approach. Section 4.3 extends the approach to provide bias-corrected temperature for sites without WRF/UCM projections. Section 4.4 presents a case study for evaluating bias correction performance of the proposed approach. Finally, Section 4.5 summarizes the chapter and outlines future research plans.

4.2 Spatio-temporal Bias Correction Method

This section presents the spatio-temporal bias correction model and model parameter estimation procedure.

4.2.1 Model Formulation

Let $y(s,t)$ and $x(s,t)$, $s = 1, \cdots, S$ and $t = 1, \cdots, T$, denote the measured and WRF/UCM-predicted temperature at site s and time t, respectively. In our analysis

we use hourly data (thus, $T = 24$), and S denotes the number of weather stations. We first consider a linear MOS model, specifically, a linear regression model [129], to post-process the WRF/UCM projection as

$$y(s,t) = \beta_0 + \beta_1 x(s,t) + \epsilon(s,t), \tag{4.1}$$

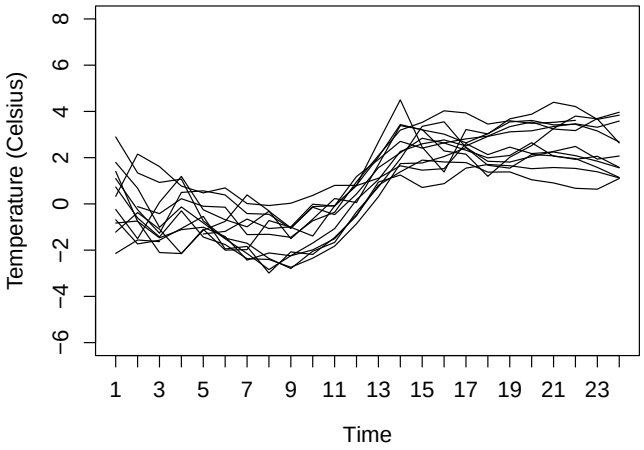
where β_0 and β_1 are the intercept and slope coefficients, respectively, and $\epsilon(s,t)$ represents a random noise.

We apply the linear MOS model to the datasets during each heat wave event shown in Figure 4.2. Figure 4.4 shows temporal patterns in the residuals, which shows underestimation in the morning and overestimation in the daytime in general. Clearly, the main issue with this linear MOS model is that it assumes homogeneous bias across time and space.

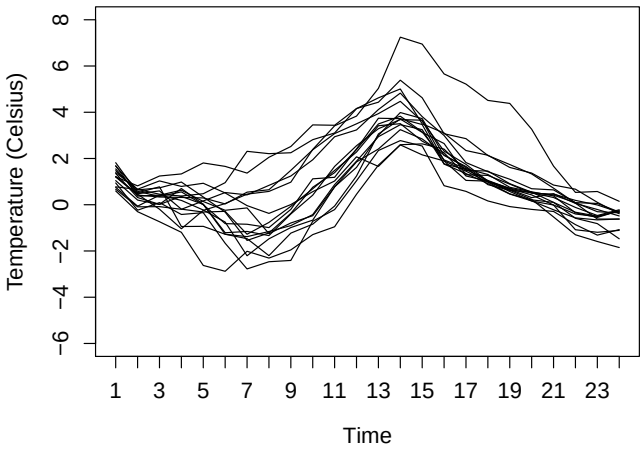
To reflect noted heterogeneous bias patterns, we allow the model coefficients to vary as a function of space and time [41]. Specifically, we decompose each model coefficient into fixed as well as spatially- and temporarily-varying components as follows:

$$\begin{aligned} y(s,t) &= \beta_0(s,t) + \beta_1(s,t)x(s,t) + \epsilon_y(s,t) \\ &= \left[\beta_0^{(F)} + \beta_0^{(R)}(s,t)\right] + \left[\beta_1^{(F)} + \beta_1^{(R)}(s,t)\right]x(s,t) + \epsilon_y(s,t) \\ &= \left[\beta_0^{(F)} + \beta_1^{(F)}x(s,t)\right] + \left[\beta_0^{(R)}(s,t) + \beta_1^{(R)}(s,t)x(s,t)\right] + \epsilon_y(s,t), \end{aligned} \tag{4.2}$$

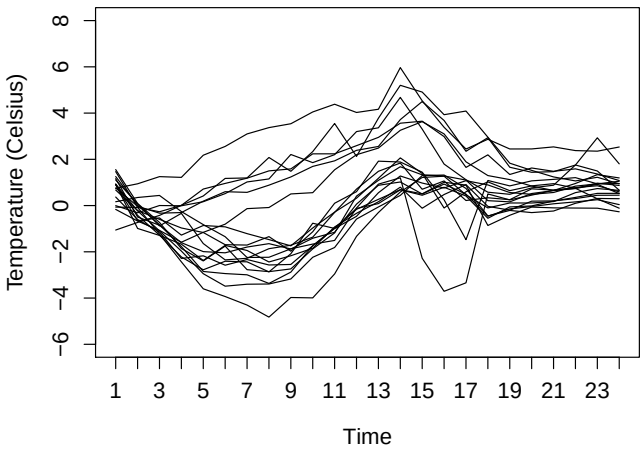
where $\beta_0^{(F)}$ and $\beta_1^{(F)}$ represent fixed effects, whereas $\beta_0^{(R)}(s,t)$ and $\beta_1^{(R)}(s,t)$ are random effects. Here, $\beta_0^{(F)} + \beta_1^{(F)}x(s,t)$ provides the global (or common) trend in temperature at all sites and hours, while $\beta_0^{(R)} + \beta_1^{(R)}x(s,t)$ captures the location- and time-specific bias. Lastly, $\epsilon_y(s,t)$ represents random noise, which is assumed to obey a Gaussian distribution with zero mean and variance σ_y^2



(a) 8/28/2011



(b) 8/8/2013



(c) 7/24/2017

Figure 4.4: Residuals in the linear regression model

We note that the bias patterns at stations located close to each other and temporally adjacent points exhibit greater similarity, compared to those for widely separated locations and/or times. This implies that the model coefficient $\beta_i^{(R)}(s,t)$ should be similar to $\beta_i^{(R)}(s',t')$ for s close to s' and t close to t', for $i=0,1$. To reflect such temporal and spatial correlation, we formulate the site- and time-specific effects $\beta_i^{(R)}(s,t)$ as latent random processes and further decompose them into temporal and spatial random effects, denoted by $\beta_i^{(T)}(t)$ and $\beta_i^{(S)}(s)$, respectively.

First, to describe the temporal correlation at consecutive time points, $\beta_i^{(T)}(t)$ should depend on $\beta_i^{(T)}(t-1)$. We employ an autoregressive (AR) process to capture temporal dependency as follows [41, 134]:

$$\beta_i^{(T)}(t) = \rho_i \beta_i^{(T)}(t-1) + \xi_i(t), \tag{4.3}$$

for $i=0,1$, where $|\rho_i| \leq 1$ represents the AR parameter, quantifying the dependence intensity between two consecutive time points, and $\xi_i(t) \sim N(0, \delta_i^2)$ is assumed to be Gaussian random noise. A higher-order AR process can be employed, however, we adopt AR(1) in this chapter due to its simplicity.

The AR(1) formulation in (4.3) implies that the collection of temporal random effects, $\mathcal{B}_i^{(T)} = [\beta_i^{(T)}(1), \beta_i^{(T)}(2), \cdots, \beta_i^{(T)}(T)]^T$ for $i=0,1$, follows a multivariate Gaussian distribution as

$$\mathcal{B}_i^{(T)} \sim N(0, \Sigma_i^{(T)}),$$

where $\Sigma_i^{(T)}$ represents a $T \times T$ temporal covariance matrix [135], whose $(t,t')^{th}$ component is defined as

$$\Sigma_i^{(T)}(t,t') = \delta_i^2 \frac{\rho_i^{|t-t'|}}{1-\rho_i^2},$$

for $t,t' = 1, \cdots, T$. This implies that the correlation intensity gets stronger, leading to a more similar bias pattern when t and t' are more adjacent.

Next, the temperature at closely spaced locations should be similar to one another, as observed in Figure 4.3. In other words, each site-specific random effect $\beta_i^{(R)}(s)$ at location s should be similar (less similar) to $\beta_i^{(R)}(s')$ for a close (remote) location s' for $i = 0, 1$ [53, 59]. To address such spatial dependency across sites, we employ a GP. Due to its nonparametric nature, a GP can flexibly represent the spatial variation. For the collection of site-specific random effects, $\mathcal{B}_i^{(S)} = [\beta_i^{(S)}(1), \beta_i^{(S)}(2), \cdots, \beta_i^{(S)}(S)]^T$ for $i = 0$ and 1, we have

$$\mathcal{B}_i^{(S)} \sim N(0, \Sigma_i^{(S)}),$$

where $\Sigma_i^{(S)}$ represents an $S \times S$ spatial covariance matrix. Among several possible covariance functions, we use the exponential covariance function in this chapter. Thus, the $(s, s')^{th}$ element of $\Sigma_i^{(S)}$ is defined as

$$\Sigma_i^{(S)}(s, s') = \tau_i^2 \exp\Big\{ - \frac{||L_s - L_{s'}||^2}{\ell_i} \Big\},$$

for $i = 0, 1$ and $s, s' = 1, \cdots, S$, where L_s is the location of site s, $|| \cdot ||$ denotes the Euclidean distance, and τ_i^2 and ℓ_i, respectively, represent the variance term and scale parameter. Note that $\beta_i^S(s)$ and $\beta_i^S(s')$ are strongly correlated when s and s' are closer.

Finally, we combine the temporal and spatial random effects to represent $\beta_i^{(R)}(s, t)$ in (4.2) for $i = 0, 1$. We consider the separable covariance structure in time and space with either additive or multiplicative structure. With the additive structure, $\beta_i^{(R)}(s, t) = \beta_i^{(S)}(s) + \beta_i^{(T)}(t)$ and its covariance function becomes

$$\begin{aligned} \Sigma_i^{(R)}((s, t), (s', t')) &= \Sigma_i^{(S)} \otimes I_{T \times T} + I_{S \times S} \otimes \Sigma_i^{(T)} \\ &= \tau_i^2 \exp\Big\{ - \frac{(L_s - L_{s'})^2}{\ell_i} \Big\} + \delta_i^2 \frac{\rho_i^{|t-t'|}}{1 - \rho_i^2}, \end{aligned}$$

for $i = 0, 1$, where $\otimes$ implies a Kronecker product. Alternatively, the multiplicative

structure $\beta_i^{(R)}(s,t) = \beta_i^{(S)}(s)\beta_i^{(T)}(t)$ can be employed and its covariance function can be derived in a similar manner.

4.2.2 Parameter Estimation and Prediction

To estimate the parameters in the proposed bias correction model with latent spatial and temporal processes, we derive the density of $y(s,t)$ in (4.2). Let us express the proposed model in a matrix form. Let $\mathbf{Y} = [y(s_1,t_1), y(s_1,t_2), \ldots, y(s_S,t_T)]^T$ denote the vector of measured temperature, $\mathbf{X}^{(F)} = [\mathbf{1}, \mathbf{x}] \in \mathbb{R}^{ST\times 2}$ denote a design matrix for the fixed effect terms with $\mathbf{1} \in \mathbb{R}^{ST\times 1}$ being a vector of ones and $\mathbf{x} = [x(s_1,t_1), x(s_1,t_2), \ldots, x(s_S,t_T)]^T$, $\mathbf{I}_{ST}$ denote a $ST \times ST$ idenity matrix, $\mathbf{D_x}$ denote the diagonal matrix whose diagonal elements are WRF/UCM projections, and $\epsilon_\mathbf{y} = [\epsilon_y(s_1,t_1), \epsilon_y(s_1,t_2), \ldots, \epsilon_y(s_S,t_T)]^T$ denote the vector of random noise. Then, the proposed model becomes

$$\mathbf{Y} = \mathbf{X}^{(F)}\mathcal{B}^{(F)} + \mathbf{I}_{ST}\mathcal{B}_0^{(R)} + \mathbf{D_x}\mathcal{B}_1^{(R)} + \epsilon_\mathbf{y}, \tag{4.4}$$

where $\mathcal{B}^{(F)} = [\beta_0^{(F)}, \beta_1^{(F)}]^T \in \mathbb{R}^{2\times 1}$ is a vector of the fixed effects, and $\mathcal{B}_0^{(R)}$ and $\mathcal{B}_1^{(R)}$ are $(ST \times 1)$-dimensional vectors of intercept random effects and slope random effects, respectively.

Because the mean of the random effects is zero, the mean of $\mathbf{Y}$ becomes $X^{(F)}\mathcal{B}^{(F)}$. Recall that the covariance matrices of $\mathcal{B}_0^{(R)}$ and $\mathcal{B}_1^{(R)}$, denoted by $\Sigma_0^{(R)}$ and $\Sigma_1^{(R)}$, respectively, can be obtained from (4.4). Assuming the independence among $\mathcal{B}_0^{(R)}$, $\mathcal{B}_1^{(R)}$, and the random noise, the covariance of $\mathbf{Y}$ becomes

$$\Sigma_Y = \Sigma_0^{(R)} + \mathbf{D_x}\Sigma_1^{(R)}\mathbf{D_x} + \sigma_y^2 I_{ST}. \tag{4.5}$$

As such, $\mathbf{Y}$ follows a multivariate Gaussian distribution as

$$\mathbf{Y} \sim N(\mathbf{X}^{(F)}\mathcal{B}^{(F)}, \Sigma_Y).$$

The distribution in (4.6) holds for both additive and multiplicative covariance structure. For the additive covariance structure in (4.4), we introduce slightly modified forms to facilitate parameter estimation as follows.

$$K_0^{(R)}((s,t),(s',t')) = \exp\Big\{-\frac{(L_s - L_{s'})^2}{\ell_0}\Big\} + \frac{\delta_0^2}{\tau_0^2}\frac{\rho_0^{|t-t'|}}{1-\rho_0^2},$$

and

$$K_1^{(R)}((s,t),(s',t')) = \frac{\tau_1^2}{\tau_0^2}\exp\Big\{-\frac{(L_s - L_{s'})^2}{\ell_1}\Big\} + \frac{\delta_1^2}{\tau_0^2}\frac{\rho_1^{|t-t'|}}{1-\rho_1^2}.$$

Then, the covariance matrix of the response vector $\mathbf{Y}$ becomes

$$\begin{aligned}\Sigma_Y &= \Sigma_0^{(R)} + \mathbf{D_x}\Sigma_1^{(R)}\mathbf{D_x} + \sigma_y^2 I_{ST} \\ &= \tau_0^2(K_0^{(R)} + \mathbf{D_x}K_1^{(R)}\mathbf{D_x} + gI_{ST}) \\ &= \tau_0^2 K_Y,\end{aligned}$$

with $g = \sigma_y^2/\tau_0^2$ and $K_Y = K_0^{(R)} + \mathbf{D_x}K_1^{(R)}\mathbf{D_x} + gI_{ST}$.

Finally, we estimate the parameters and hyperparameters in the proposed model by maximizing the likelihood, which is denoted by $L_{\mathbf{Y}|\mathbf{x}}$, as follows:

$$L_{\mathbf{Y}|\mathbf{x}} = (2\pi\tau_0^2)^{-ST/2}|K_Y|^{-1/2}\exp\Big\{-\frac{1}{2\tau_0^2}(\mathbf{Y} - \mathbf{X}^{(F)}\mathcal{B}^{(F)})^T K_Y^{-1}(\mathbf{Y} - \mathbf{X}^{(F)}\mathcal{B}^{(F)})\Big\}.$$

Let $\boldsymbol{\theta}_1 = \{l_0, l_1, \rho_0, \rho_1, \delta_0, \delta_1, \tau_1^2, g\}$ denote the set of hyperparameters. We obtain the log-likelihood $\ell_{\mathbf{Y}|\mathbf{x}}(\mathcal{B}^{(F)}, \tau_0^2, \boldsymbol{\theta}_1)$ as

$$\begin{aligned}\ell_{\mathbf{Y}|\mathbf{x}}(\mathcal{B}^{(F)}, \tau_0^2, \boldsymbol{\theta}_1) &= -\frac{ST}{2}\log(2\pi) - \frac{ST}{2}\log(\tau_0^2) - \frac{1}{2}\log|K_Y| \\ &\quad -\frac{1}{2\tau_0^2}(\mathbf{Y} - \mathbf{X}^{(F)}\mathcal{B}^{(F)})^T K_Y^{-1}(\mathbf{Y} - \mathbf{X}^{(F)}\mathcal{B}^{(F)}).\end{aligned} \tag{4.6}$$

By taking partial derivatives of the log-likelihood with respective to $\mathcal{B}^{(F)}$ and τ_0^2 and setting them to 0, we obtain the maximum likelihood estimators (MLEs) for $\mathcal{B}^{(F)}$ and τ_0^2 as

$$\hat{\mathcal{B}}^{(F)} = (\mathbf{X}^{(F)T}K_Y^{-1}\mathbf{X}^{(F)})^{-1}\mathbf{X}^{(F)T}K_Y^{-1}\mathbf{Y}, \tag{4.7}$$

and

$$\hat{\tau}_0^2 = \frac{1}{ST}(\mathbf{Y} - \mathbf{X}^{(F)}\hat{\mathcal{B}}^{(F)})^T K_Y^{-1}(\mathbf{Y} - \mathbf{X}^{(F)}\hat{\mathcal{B}}^{(F)}), \tag{4.8}$$

respectively.

Plugging $\hat{\mathcal{B}}^{(F)}$ and $\hat{\tau}_0^2$ into (4.6), the log-likelihood function in (4.6) becomes a function of $\boldsymbol{\theta}_1$ as follows.

$$\begin{aligned}\ell_{\mathbf{Y}|\mathbf{x}}(\boldsymbol{\theta}_1) &= -\frac{ST}{2} - \frac{ST}{2}\log(2\pi) \\ &\quad -\frac{ST}{2}\log\Big\{\frac{1}{ST}(\mathbf{Y} - \mathbf{X}^{(F)}(\mathbf{X}^{(F)T}K_Y^{-1}\mathbf{X}^{(F)})^{-1}\mathbf{X}^{(F)T}K_Y^{-1}\mathbf{Y})^T K_Y^{-1} \\ &\quad (\mathbf{Y} - \mathbf{X}^{(F)}(\mathbf{X}^{(F)T}K_Y^{-1}\mathbf{X}^{(F)})^{-1}\mathbf{X}^{(F)T}K_Y^{-1}\mathbf{Y})\Big\} - \frac{1}{2}\log|K_Y|.\end{aligned}$$

We obtain the MLE of $\boldsymbol{\theta}_1$ using the L-BFGS-B algorithm, a numerical optimization methods. In our case study in Section 4.4, the 'optim' function in 'stats' package in the statistical software R is used for implementing the L-BFGS-B algorithm [136]. Once we obtain the MLE of $\boldsymbol{\theta}_1$, we introduce it into (4.7) and (4.8) and attain the MLEs of $\mathcal{B}^{(F)}$ and τ_0^2.

Once the parameters are estimated, we can obtain the predictive distribution of the bias-corrected temperature prediction at an unmeasured location. Suppose we have the WRF/UCM temperature projection x^* (or equivalently, $x(s^*, t^*)$) at location s^* and time point t^*, but do not have the corresponding temperature measurement. Then, given the dataset $\mathcal{D}_{ST} = [\{x(s_1, t_1), y(s_1, t_1)\}, \ldots, \{x(s_S, t_T), y(s_S, t_T)\}]^T$ used for model fitting, the predictive density of temperature y^* at x^*, follows a Gaussian distribution as

$$y^*|(x^*, \mathcal{D}_{ST}) \sim N(\mu_{y^*|x^*}, \sigma^2_{y^*|x^*}), \tag{4.9}$$

with

$$\mu_{y^*|x^*} = \mathbf{X}^*\mathcal{B}^{(F)} + \mathbf{c}_*^T\Sigma_Y^{-1}(\mathbf{Y} - \mathbf{X}^{(F)}\mathcal{B}^{(F)}), \tag{4.10}$$

and

$$\sigma^2_{y^*|x^*} = c_{**} - \mathbf{c}_*^T\Sigma_Y^{-1}\mathbf{c}_* + \sigma_y^2, \tag{4.11}$$

where $\mathbf{X}^* = [1, x^*]$ is the design matrix and $\mathbf{c}_*$ is a covariance vector whose components provide the covariance between the training and test points (i.e., between $y|x(s,t)$ and $y|x(s^*,t^*)$), defined as

$$\mathbf{c}_* = \mathbf{c}_{0,*} + \mathbf{D}_\mathbf{x}\mathbf{c}_{1,*}x^*, \tag{4.12}$$

and c_{**} is the variance at the test point as

$$c_{**} = c_{0,**} + (x^*)^2 c_{1,**}. \tag{4.13}$$

Here, $\mathbf{c}_{i,*}$ is the covariance vector constructed from the training and test points for $\beta_i^{(R)}$ for $i = 0, 1$, and $c_{i,**}$ is the variance term of test points as follows.

$$\begin{pmatrix} \mathcal{B}_i^{(R)} \\ \beta_i^{(R)}(s^*,t^*) \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_i^{(R)} & \mathbf{c}_{i,*} \\ \mathbf{c}_{i,*}^T & c_{i,**} \end{pmatrix}\right), \tag{4.14}$$

for $i = 0, 1$.

Equations (4.9)-(4.13) define the predictive distribution at a specific location and time. Predictive distributions at multiple locations and times can be easily obtained in a similar manner, but for brevity we omit those details.

4.3 Integrative Spatio-temporal Bias Correction Model

This section formulates the integrative spatio-temporal bias correction model and derive the prediction procedure at sites where WRF/UCM projections are unavailable.

4.3.1 Model Formulation

The bias correction model discussed in Section 4.2 provides the adjusted temperature at sites where the WRF/UCM projections are available. Suppose we would like to estimate the temperature at places without WRF/UCM temperature projections. In this case, one may re-run the WRF/UCM model to obtain temperature projections at those locations. However, running the WRF/UCM model is computationally expensive and time-demanding [137]. This section extends the bias correction approach presented in Section 4.2 and estimates temperature at sites where WRF/UCM projections are absent.

Our approach is to integrate the temperature field projected by WRF/UCM with the bias correction model. Specifically, we create the spatially and temporally correlated WRF/UCM temperature projection field using GP, i.e., we formulate $x(s,t)$ with a GP model. Suppose we obtain WRF/GCM projections $\mathbf{x} = [x(s_1,t_1), x(s_1,t_2), \ldots, x(s_S,t_T)]^T$. Then the marginal density of $\mathbf{x}$ follows the multivariate Gaussian denstiy as

$$\mathbf{x} \sim N(0, \Sigma_{\mathbf{x}}),$$

where $\Sigma_{\mathbf{x}}$ is the spatial and temporal covariance matrix, similar to (4.4) but with different hyperparameters.

Finally, the integrative model combines the WRF/UCM temperature field with the bias-correction model in Section 4.2 as follows.

$$\begin{aligned} \mathbf{x} &\sim N(0, \Sigma_{\mathbf{x}}) \\ y|x(s,t) &= \left[\beta_0^{(F)} + \beta_1^{(F)} x(s,t)\right] + \left[\beta_0^{(R)}(s,t) + \beta_1^{(R)}(s,t) x(s,t)\right] \\ &\quad + \epsilon_y(s,t), \end{aligned} \tag{4.15}$$

where the random effects $\beta_0^{(R)}(s,t)$ and $\beta_1^{(R)}(s,t)$ are formulated in (4.3)-(4.4).

4.3.2 Parameter Estimation and Prediction

This section discusses how to estimate the parameters in the integrative model. The likelihood of the integrative model, given $(\mathbf{x}, \mathbf{Y})$, is $L = L_{\mathbf{x}} \cdot L_{\mathbf{Y}|\mathbf{x}}$, where $L_{\mathbf{Y}|\mathbf{x}}$ is in (4.6) and $L_{\mathbf{x}}$ is

$$L_{\mathbf{x}} = (2\pi)^{-ST/2}|\Sigma_{\mathbf{x}}|^{-1/2} \exp\Big\{ -\frac{1}{2}\mathbf{x}^T\Sigma_{\mathbf{x}}^{-1}\mathbf{x}\Big\}.$$

Let $\boldsymbol{\theta}_2$ denote a set of hyperparameters in the WRF/UCM temperature field. Then the log-likelihood becomes

$$\ell_{\mathbf{x}}(\boldsymbol{\theta}_2) = -\frac{ST}{2}(2\pi) - \frac{1}{2}|\Sigma_{\mathbf{x}}| - \frac{1}{2}\mathbf{x}^T\Sigma_{\mathbf{x}}^{-1}\mathbf{x}.$$

We estimate the parameters and hyperparameters of the integrative model by maximizing the log-likelihood. From $L = L_{\mathbf{x}} \cdot L_{\mathbf{Y}|\mathbf{x}}$, the log-likelihood becomes

$$\ell(\mathcal{B}^{(F)}, \tau_0^2, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \ell_{\mathbf{Y}|\mathbf{x}}(\mathcal{B}^{(F)}, \tau_0^2, \boldsymbol{\theta}_1) + \ell_{\mathbf{x}}(\boldsymbol{\theta}_2)$$

This implies that we can estimate the parameters in the bias correction model and the WRF/UCM temperature field separately. In other words, the MLEs of $\mathcal{B}^{(F)}$, τ_0^2, and $\boldsymbol{\theta}_1$ in the integrative model are the same as those in the bias correction model obtained in Section 4.2. We get the MLE of $\boldsymbol{\theta}_2$ by maximizing $\ell\mathbf{x}(\boldsymbol{\theta}_2)$. Similar to the optimization procedure discussed in Section 4.2, we use the L-BFGS-B algorithm to attain the MLE of $\boldsymbol{\theta}_2$ using the 'optim' function in 'stats' package in the statistical software R [136].

We now derive the predictive distribution of bias-corrected temperature at location s^* and time t^* where the corresponding WRF/UCM projection is absent. The uncertainties in both the WRF/UCM temperature field model and the bias-corrected model complicate the derivation of the predictive density. First, following the properties in the GP, the predictive distribution of the WRF/UCM temperature projection

x^* (or $x(s^*, t^*)$) is given by

$$x^*|\mathbf{x} \sim N(\mu_{x^*}, \sigma^2_{x^*}), \tag{4.16}$$

with

$$\mu_{x^*} = \mathbf{d}_*^T \Sigma_{\mathbf{x}}^{-1} \mathbf{x}, \tag{4.17}$$

$$\sigma^2_{x^*} = d_{**} - \mathbf{d}_*^T \Sigma_{\mathbf{x}}^{-1} \mathbf{d}_* + \sigma_x^2, \tag{4.18}$$

where $\mathbf{d}_*$ is a covariance vector constructed by all pairs of the training and test points in the WRF/UCM projection and d_{**} is constructed by test points as

$$\begin{pmatrix} \mathbf{x} \\ x^* \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{\mathbf{x}} & \mathbf{d}_* \\ \mathbf{d}_*^T & d_{**} \end{pmatrix} \right). \tag{4.19}$$

Next, to obtain the predictive distribution of bias-corrected temperature y^*, we combine the predictive density of x^* in (4.16)-(4.19) with the predictive density $y^*|x^*$ in (4.9)-(4.14). Given $\mathcal{D}_{ST}$, $p(y^*|\mathcal{D}_{ST})$ becomes

$$p(y^*|\mathcal{D}_{ST}) = \int p(y^*|x^*, \mathcal{D}_{ST}) p(x^*|\mathbf{x}) dx^*.$$

Because both $x^*|\mathbf{x}$ and $y^*|(x^*, \mathcal{D}_{ST})$ follow a Gaussian distribution, $y^*|\mathcal{D}_{ST}$ also obeys a Gausssian distribution as

$$y^* = y(s^*, t^*) \sim N(\mu_{y^*}, \sigma^2_{y^*})$$

where μ_{y^*} and $\sigma^2_{y^*}$ denote the mean and variance of y^* [138].

Next, we derive μ_{y^*} and $\sigma^2_{y^*}$ using the law of total expectation and the law of total variance, respectively, as follows.

$$\begin{aligned} \mu_{y^*} &= E\Big[E[y^*|x^*]\Big] \\ &= E\Big[\beta_0^{(F)} + \beta_1^{(F)} x^* + \mathbf{c}_*^T \Sigma_Y^{-1} (\mathbf{Y} - \mathbf{X}^{(F)} \mathcal{B}^{(F)})\Big] \qquad (4.20) \\ &= \beta_0^{(F)} + \beta_1^{(F)} \mu_{x^*} + (\mathbf{c}_{0,*} + \mathbf{D}_{\mathbf{x}} \mathbf{c}_{1,*} \mu_{x^*})^T \Sigma_Y^{-1} (\mathbf{Y} - \mathbf{X}^{(F)} \mathcal{B}^{(F)}), \qquad (4.21) \end{aligned}$$

where (4.20) is obtained using (4.10), and (4.21) is due to (4.12). Further, using (4.10) and (4.11), we get

$$\begin{aligned} \sigma^2_{y^*} &= E\Big[Var[y^*|x^*]\Big] + Var\Big[E[y^*|x^*]\Big] \\ (4.22) \quad &= E\Big[c_{**} + \sigma^2_y - \mathbf{c}_*^T\Sigma_Y^{-1}\mathbf{c}_*\Big] + Var\Big[\beta_0^{(F)} + \beta_1^{(F)}x^* + \mathbf{c}_*^T\Sigma_Y^{-1}(\mathbf{Y} - \mathbf{X}^{(F)}\mathcal{B}^{(F)})\Big], \end{aligned}$$

Below we derive each term in (4.22). First, $E\Big[c_{**} + \sigma^2_y\Big]$ becomes

$$\begin{aligned} (4.23) \qquad E\Big[c_{**} + \sigma^2_y\Big] &= E\Big[c_{0,**} + (x^*)^2 c_{1,**} + \sigma^2_y\Big] \\ &= c_{0,**} + E\Big[(x^*)^2\Big]c_{1,**} + \sigma^2_y \\ &= c_{0,**} + (\mu^2_{x^*} + \sigma^2_{x^*})c_{1,**} + \sigma^2_y. \end{aligned}$$

Here, (4.23) is due to (4.13). We also attain $E\Big[\mathbf{c}_*^T\Sigma_Y^{-1}\mathbf{c}_*\Big]$ as

$$\begin{aligned} E\Big[\mathbf{c}_*^T\Sigma_Y^{-1}\mathbf{c}_*\Big] &= E\Big[(\mathbf{c}_{0,*} + \mathbf{D}_\mathbf{x}\mathbf{c}_{1,*}x^*)^T\Sigma_Y^{-1}(\mathbf{c}_{0,*} + \mathbf{D}_\mathbf{x}\mathbf{c}_{1,*}x^*)\Big] \\ &= (\mathbf{c}_{0,*} + \mathbf{D}_\mathbf{x}\mathbf{c}_{1,*}\mu_{x^*})^T\Sigma_Y^{-1}(\mathbf{c}_{0,*} + \mathbf{D}_\mathbf{x}\mathbf{c}_{1,*}\mu_{x^*}) + \sigma^2_{x^*}tr(\mathbf{c}_{1,*}^T\mathbf{D}_\mathbf{x}\Sigma_Y^{-1}\mathbf{D}_\mathbf{x}\mathbf{c}_{1,*}), \end{aligned}$$

where the last equation is obtained using the mean of a quadratic form. Finally, the last term of (4.22) becomes

$$\begin{aligned} &Var\Big[\beta_0^{(F)} + \beta_1^{(F)}x^* + \mathbf{c}_*^T\Sigma_Y^{-1}(\mathbf{Y} - \mathbf{X}^{(F)}\mathcal{B}^{(F)})\Big] \\ &\quad = Var\Big[\beta_0^{(F)} + \beta_1^{(F)}x^* + (\mathbf{c}_{0,*} + \mathbf{D}_\mathbf{x}\mathbf{c}_{1,*}x^*)^T\Sigma_Y^{-1}(\mathbf{Y} - \mathbf{X}^{(F)}\mathcal{B}^{(F)})\Big] \\ &\quad = \sigma^2_{x^*}\Big(\beta_1^{(F)} + \mathbf{c}_{1,*}^T\mathbf{D}_\mathbf{x}\Sigma_Y^{-1}(\mathbf{Y} - \mathbf{X}^{(F)}\mathcal{B}^{(F)})\Big)^2. \end{aligned}$$

In this integrative model, the uncertainty in predicting $x(s^*, t^*)$ is seemlessly incorporated into the estimation of bias-corrected temperature y^*. Compared to the predictive density in (4.9)-(4.11) for the cases where WRF/UCM projection $x(s^*, t^*)$ is available, the primary difference here is the inflated variance in (4.22) (mainly the second term of (4.22)). In Section 4.4, we study how the prediction results change, depending on whether a WRF/UCM projection is available or not.

4.4 Case Study

We use hourly temperature measurements collected at multiple locations in Austin, Texas on August 28, 2011, August 8, 2013, and July 24, 2017, when heat wave events occurred as illustrated in Figure 4.2. Due to missing data at some weather stations, we have a different number of measurements available for each heat wave event: temperature measurements from 12, 16, and 18 stations are available in 2011, 2013, and 2017, respectively. The distance between adjacent weather stations ranges from 1.16 km to 61.86 km. We also obtain WRF/UCM temperature projections at those sites. The WRF model ver. 3.9 [139] coupled with an UCM [140–142] was used to simulate a heat wave event along with urban heat effects in the Austin, Texas region. The initial and boundary conditions of the outer 36 km WRF domain were taken from Global Forecast System data provided by the NOAA National Centers for Environmental Information. For a detailed description of the WRF/UCM model considered in this chapter, we refer the readers to the study in [126].

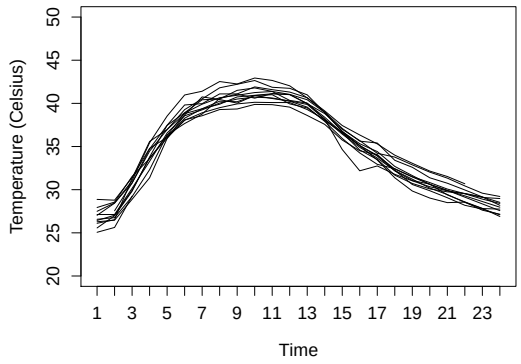
4.4.1 Implementation Results of Bias Correction Model

We implement our proposed bias-correction model developed in Section 4.2. In our implementation, we employ the additive covariance structure. As the bias patterns are different in the three heat wave events as shown in Figure 4.2, we apply the proposed procedure to each event separately. For evaluating the bias correction performance, we perform leave-one-out cross-validation. We divide the multiple sites into training and testing sites. The training set (in-sample) includes the temperature measurements at $S-1$ locations, whereas the testing set (out-of-sample) includes the measurements for 24 hours at the remaining one location. With the training set, we calibrate the bias. Then the resulting bias correction model is applied to post-correct

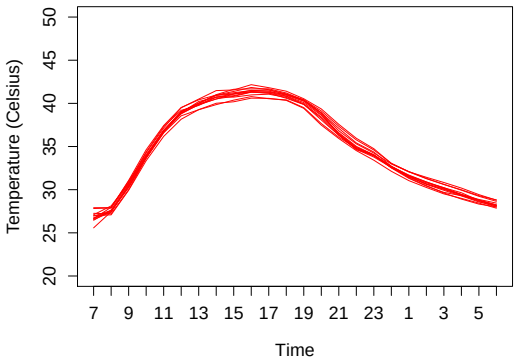
the bias of WRF/UCM projection at the testing site. This procedure is repeated S times so every site becomes the testing site. Due to the limited data (1 day of data for each heat wave event) available to us, we can only apply our approach to the spatial out-of-sample evaluation, but we intend to extend the procedure to make temporal out-of-sample predictions when more data become available to us in the future.

Figure 4.5 compares the measured temperature and the post-calibrated temperature from the proposed bias correction model when each site is selected as a testing site. Overall, the proposed approach successfully captures spatial and temporal variability in the bias, and we note that bias-corrected temperatures are close to actual observations. The proposed approach reduces the average root mean square error (RMSE) of the WRF/UCM projection by about 60% from 1.87 to 0.77. In the next section, we provide additional details on these results.

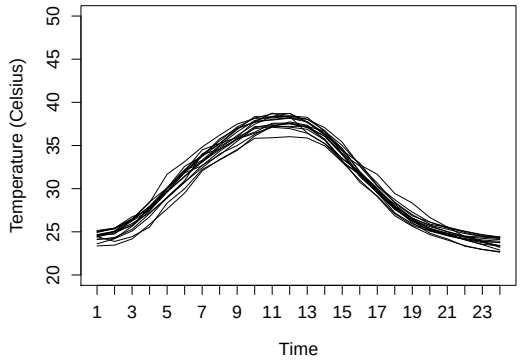
One notable advantage of the proposed bias correction model is that it can quantify different types of variability. The overall variability (i.e., the variance of the data) of actual temperature at multiple sites can be explained by temporal, spatial, and unstructured contributions. Table 4.1 shows that temporal variability explains more than 90% of the total variability. This is due to the diurnal temperature pattern over a day. While spatial variability appears to be less substantial, it still exceeds the unstructured variability from the random noise in two cases. This implies that accounting for spatial correlation can help improve estimation performance.



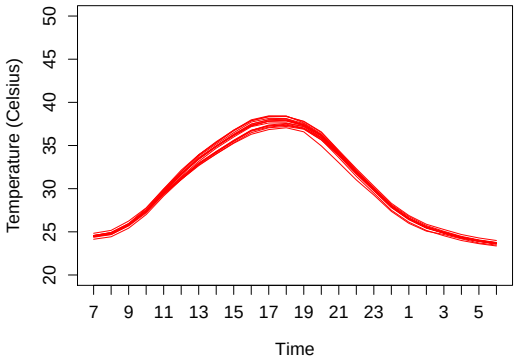
(a) Measured temperature (2011)



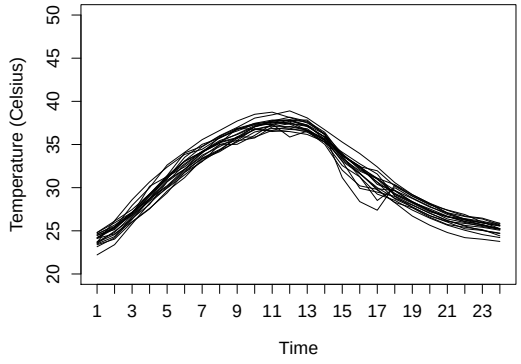
(b) Estimated temperature (2011)



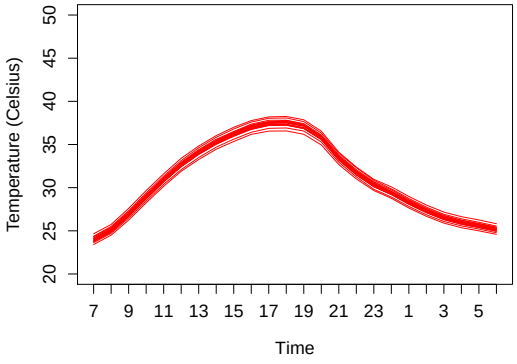
(c) Measured temperature (2013)



(d) Estimated temperature (2013)



(e) Measured temperature (2017)



(f) Estimated temperature (2017)

Figure 4.5: Comparison between measured temperature (left panel) and bias-corrected temperature estimate (right panel) by the proposed approach

Table 4.1: The average ratio (expressed in percent) of the spatial, temporal, and unstructured variabilities (the values parentheses are standard deviations at the testing sites)

Type of variability	Formula	8/28/2011	8/8/2013	7/25/2017
Spatial	$\frac{\tau_0^2+[x(s,t)]^2\tau_1^2}{Var(y(s,t))}$	4.69 (1.58)	1.87 (0.73)	3.07 (1.12)
Temporal	$\frac{\frac{\delta_0^2}{1-\rho_0^2}+[x(s,t)]^2\frac{\delta_1^2}{1-\rho_1^2}}{Var(y(s,t))}$	89.21 (3.90)	96.68 (0.92)	95.75 (1.15)
Unstructured	$\frac{\sigma^2}{Var(y(s,t))}$	6.08 (3.00)	1.45 (0.43)	1.18 (0.19)

We further study the role of spatio-temporal random effects in the intercept and slope parameters of our model. Table 4.2 shows the average ratio of variability in $\beta_0^{(R)}$ and $\beta_1^{(R)}$, and the random noise. In the heat wave event cases of 2011 and 2017, most of the variability is explained by the random effect in the intercept coefficient $\beta_0^{(R)}$, whereas the random effect in the slope coefficient $\beta_1^{(R)}$ captures a greater portion of the variability in 2013. This result demonstrates that including random effects in both the intercept and slope coefficients provides greater flexibility in the model. In all cases, the two random effects explain greater than 90% of the variability in total, greatly exceeding the unstructured variability resulting from the random noise.

Table 4.2: The average ratio (expressed in percent) of the variability in random effects $\beta_0^{(R)}$ and $\beta_1^{(R)}$ and unstructured variability (the values in parentheses are the standard deviations at the testing sites)

Variability type	Formula	8/28/2011	8/8/2013	7/25/2017
$\beta_0^{(R)}$	$\frac{\tau_0^2+\frac{\delta_0^2}{1-\rho_0^2}}{Var(y(s,t))}$	63.61 (35.56)	44.65 (20.68)	96.84 (1.63)
$\beta_1^{(R)}$	$\frac{[x(s,t)]^2\left(\tau_1^2+\frac{\delta_0^2}{1-\rho_0^2}\right)}{Var(y(s,t))}$	30.30 (36.77)	53.90 (20.56)	1.98 (1.57)
Unstructured	$\frac{\sigma^2}{Var(y(s,t))}$	6.09 (3.00)	1.45 (0.43)	1.18 (0.19)

4.4.2 Comparison with Alternative Approaches

We compare the bias correction performance of the proposed approach with WRF/UCM projection and two alternative approaches, namely, linear MOS and the delta change approach. First, the linear MOS assumes a linear relationship between the measured temperature $y(s,t)$ and its prediction $x(s,t)$ from WRF/UCM, modeled as in (4.1). As it assumes a stationary relationship, it does not account for spatially and temporally heterogeneous bias patterns.

The second alternative approach is the delta change approach [133]. It simply adds an average bias to the real observation at each time point. Specifically, let $\bar{B}(s,t)$ denote the average bias computed from all sites except site s at each time point t. Then, this delta change approach merely adds $\bar{B}(s,t)$ to the temperature prediction from WRF/UCM at site s, i.e.,

$$\hat{y}(s,t) = x(s,t) + \bar{B}(s,t), \tag{4.24}$$

where

$$\bar{B}(s,t) = \frac{1}{S-1} \sum_{s' \in S^c} [y(s',t) - x(s',t)], \tag{4.25}$$

where S^c is the set of sites not including the target site s. While the delta method appears to provide only a point estimation at first glance, we note that it is equivalent to the following linear regression model that does not include a slope parameter.

$$y(s,t) = \alpha_0(t) + x(s,t) + \epsilon(s,t),$$

for $s \in S^c$ at each time t. Then the least squares estimate of the intercept $\alpha_0(t)$ becomes equivalent to $\bar{B}(s,t)$ in (4.25), and the temperature at the target site s at each time point t can be obtained using (4.24).

Because the delta change approach does not correct the slope parameter, it is less flexible than the linear MOS in some sense. However, it adjusts the bias at each hour to account for the temporally varying bias pattern. Still, the temporal *correlation* of bias is not taken into consideration in the delta change approach, because it computes the temporal bias separately in each hour. Thus, this approach addresses the temporal variation only partially. Moreover, it does not address spatial correlation. On the contrary, our approch provides a collabrative learning framework [143] to identify WRF/UCM bias. Table 4.3 shows whether or not each method accounts for spatially and temporally varying bias patterns.

Table 4.3: Comparison of alternative methods (X: not addressed, △: partially addressed, ◯: fully addressed)

	Spatial bias pattern	Temporal bias pattern
Linear MOS	X	X
Delta change approach	X	△
Proposed approach	◯	◯

We measure the bias correction performance in terms of both point and probabilistic estimation capability using several metrics. First, for the point estimation performance, we use RMSE, defined as

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}\{y(s,t) - \hat{y}(s,t)\}^2}{T}},$$

for $s = 1, \cdots, S$, where $\hat{y}(s,t)$ denotes the estimated temperature at testing site s at hour t. For evaluating the probabilistic performance, the continuous ranked probability score (CRPS) is employed [61], defined as

$$\begin{aligned} CRPS &= \frac{1}{T}\sum_{t=1}^{T}\hat{\sigma}(s,t)\Big[\frac{y(s,t)-\hat{y}(s,t)}{\hat{\sigma}(s,t)}\Big\{2\Phi\Big(\frac{y(s,t)-\hat{y}(s,t)}{\hat{\sigma}(s,t)}\Big)+1\Big\} \\ &\quad +2\phi\Big(\frac{y(s,t)-\hat{y}(s,t)}{\hat{\sigma}(s,t)}\Big)-\frac{1}{\sqrt{\pi}}\Big], \end{aligned}$$

where ϕ and Φ, respectively, denote the standard Gaussian probability density function and cumulative density function, and $\hat{\sigma}(s,t)$ is the estimated standard deviation of the temperature at station s and time t. A smaller value of CRPS indicates better performance. In addition, other probabilistic performance metrics, including the width of the prediction interval (PI) and PI coverage rate are also examined for assessing probabilistic bias correction performance.

First, Table 4.4 summarizes the average RMSE from the spatial leave-one-out evaluation, where each site is considered as a testing site. Our proposed approach provides noteworthy improvement over the original WRF/UCM and alternatives. The RMSE values for our proposed approach are substantially lower than those of the original WRF/UCM in all three heat wave events. The proposed approach also results in the lowest RMSE value compared to other alternatives. Next, we evaluate the probabilistic performance. Table 4.5 presents average CRPS values at the testing sites. Note that the WRF/UCM provides deterministic projections, so its CRPS cannot be computed. The proposed bias correction model consistently performs better than the alternative approaches in all cases, demonstrating strong probabilistic prediction capability.

Table 4.4: Average RMSE over the testing sites from spatial out-of-sample testing (the values in parentheses are the standard deviation values of the RMSE at the testing sites, unit: celsius)

	8/28/2011	8/8/2013	7/25/2017	Average
WRF/UCM	2.12 (0.47)	1.81 (0.55)	1.69 (0.55)	1.87 (0.52)
Linear MOS	1.51 (0.21)	1.63 (0.43)	1.59 (0.51)	1.58 (0.38)
Delta change approach	0.98 (0.21)	0.90 (0.52)	1.34 (0.66)	1.09 (0.55)
Bias correction model	**0.90 (0.13)**	**0.64 (0.33)**	**0.80 (0.35)**	**0.77 (0.33)**

Table 4.5: Average CRPS over the testing sites from spatial out-of-sample testing (the values in parentheses are the standard deviation values of CRPS at the testing sites)

	8/28/2011	8/8/2013	7/25/2017	Average
WRF/UCM	-	-	-	-
Linear MOS	0.87 (0.18)	0.93 (0.25)	0.93 (0.32)	0.91 (0.25)
Delta change approach	0.59 (0.12)	0.57 (0.27)	0.81 (0.35)	0.66 (0.25)
Bias correction model	**0.52 (0.09)**	**0.39 (0.22)**	**0.49 (0.20)**	**0.46 (0.21)**

Tables 4.6 summarizes average PI widths that are indicators of prediction sharpness. The proposed model provides the narrowest PIs in all cases. In comparison with the PIs from the linear MOS and the delta change approaches, the proposed bias correction model results in 50% and 35% reductions, respectively, in PI widths on average. This result is an indication of greatly reduced uncertainty in the bias-corrected prediction.

Table 4.6: Average width of 95% PIs over the testing sites from spatial out-of-sample testing (the values in parentheses are the standard deviations at the testing sites, unit: celsius)

	8/28/2011	8/8/2013	7/25/2017	Average
WRF/UCM	-	-	-	-
Linear MOS	5.94 (0.08)	6.52 (0.02)	6.41 (0.02)	6.29 (0.04)
Delta change approach	4.35 (0.75)	4.13 (1.57)	5.79 (2.53)	4.83 (2.01)
Bias correction model	**3.49 (0.21)**	**2.73 (0.26)**	**3.29 (0.61)**	**3.15 (0.36)**

Table 4.7 further shows the coverage rate with the 95% PI. An ideal coverage rate is 95%, implying that 95% of the actual measurements are located within their 95% PIs. Our coverage rates are slightly lower than the delta change approach: however, they are close to the desired rate of 95%. It should be also noted that, while the delta change approach achieves the highest coverage rate on average, it does so with much greater PIs; its average PI width (4.83) is much wider than the proposed approach's

(3.15) (see Table 4.6).

Table 4.7: Average 95% PI coverage rate (%) over the testing sites from spatial out-of-sample testing (the values in parentheses are the standard deviations at the testing sites)

	8/28/2011	8/8/2013	7/25/2017	Average
WRF/UCM	-	-	-	-
Linear MOS	**98.21 (2.69)**	80.73 (6.78)	89.58 (17.11)	89.51 (8.86)
Delta change approach	96.15 (6.24)	**94.79 (10.15)**	**94.44 (17.68)**	**95.04 (12.64)**
Bias correction model	95.14 (5.57)	92.97 (16.30)	92.59 (13.30)	93.39 (11.72)

To elaborate further on the issue of PI width, Figure 4.6 compares actual and bias-corrected temperatures and their prediction intervals at the site ASPLS in 2013, when ASPLS is a testing site. The black circle denotes the measured temperature and the blue curve represents the WRF/UCM temperature projections. The red solid and dashed curves are the bias-corrected temperature estimates from each method and their corresponding 95% PIs. Figure 4.6a shows that the linear MOS vertically shifts the WRF/UCM output by the same amount, failing to capture large biases between 11:00 and 16:00. This is because it does not take the temporally heterogeneous bias pattern into consideration. Figure 4.6b shows that the delta change approach provides a better point correction performance, compared to the linear MOS, because it calibrates the WRF/UCM projection in each time point. However, there is still a discrepancy between the real and bias-corrected estimate between 9:00 to noon. We believe that this is because the delta method does not account for spatial variation but adjusts the WRF/UCM output by the same amount over multiple locations at each time.

In contrast, the proposed approach provides better calibration performance as is evident in Figure 4.6c. Its point estimates align well with actual measurements. We also observe that the PI from the proposed approach is narrower, exhibiting sharper

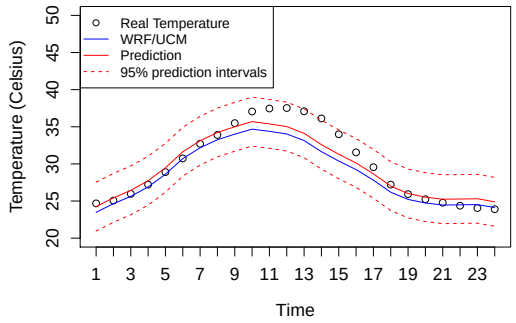
estimation, than those from the alternative approaches. Despite its narrow interval, our approach includes most measurements within the 95% PIs. We will explain Figure 4.6d in the next section.

4.4.3 Implementation Results with the Integrative Bias Correction Model

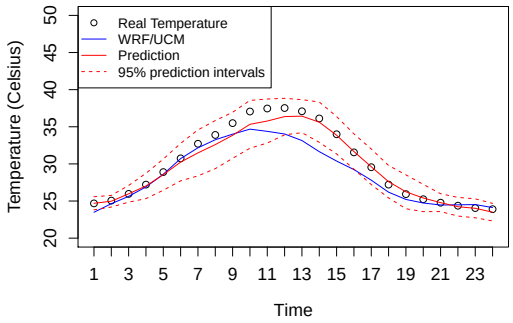
This section evaluates the estimation performance of the integrative bias correction model discussed in Section 4.3 to predict temperature when WRF/UCM projections at testing sites are absent. The alternative linear MOS and delta change approaches discussed in Section 4.4.2 cannot correct the bias for those sites, which attests to another advantage of the proposed approach. To measure the predictive performance of the integrative model, we use the bias correction model in Section 4.2 as a baseline model. Recall that the original bias correction model provides the predictive density of $y^*|x^*$, given the WRF/UCM projection $x^* = x(s^*, t^*)$, whereas the integrative model generates the predictive density of y^* without knowing x^*.

Table 4.8 compares the average RMSE, CRPS, 95% PI width, and the PI coverage rate in the leave-one-out testing. The integrative model provides slightly higher RMSE and CRPS values, indicating that both point and probabilistic prediction performance gets slightly worse. This is understandable primarily due to the additional estimation of x^* needed with the integrative model. However, the difference is not significant. Interestingly, even with the absence of the WRF/UCM projection, our integrative approach provides much better prediction results than alternative approaches with WRF/UCM projections (see Tables 4.4∼4.7).

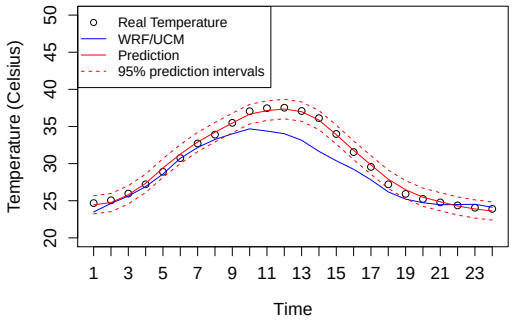
As expected, the integrative model results in a larger variance in the predictive density, leading to wider PIs than the original bias correction model, because of the inflated uncertainty indicated in (4.22). On the other hand, the wider PIs of



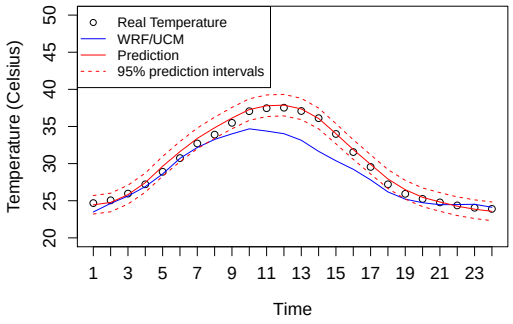
(a) Linear MOS



(b) Delta change approach



(c) Bias correction model



(d) Integrative bias correction model

Figure 4.6: Comparison of measured and bias-corrected temperatures and their prediction intervals at the testing site ASPLS in 2013 from spatial out-of-sample estimation

Table 4.8: Comparison of the proposed bias correction and integrative models: Average over the testing sites from spatial out-of-sample testing (the values in parentheses are standard deviations at the testing sites)

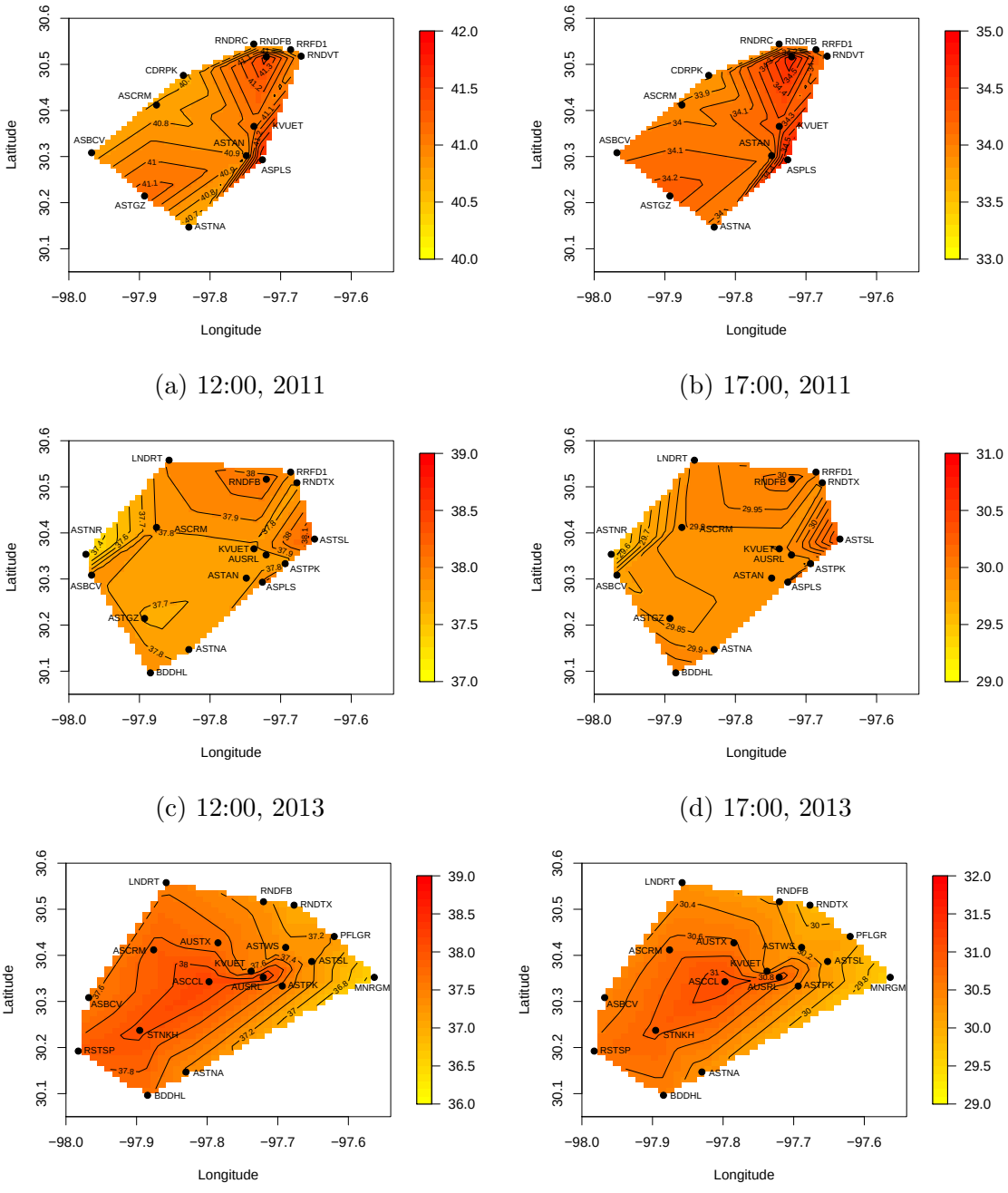
Measure	Model	8/28/2011	8/8/2013	7/25/2017	Average
RMSE	Original model	**0.90 (0.13)**	**0.64 (0.33)**	**0.80 (0.35)**	**0.77 (0.33)**
	Integrative model	0.96 (0.26)	0.70 (0.31)	**0.80 (0.33)**	0.80 (0.31)
CRPS	Original model	**0.52 (0.09)**	**0.39 (0.22)**	**0.49 (0.20)**	**0.46 (0.21)**
	Integrative model	0.56 (0.16)	0.42 (0.20)	**0.49 (0.21)**	0.48 (0.20)
95% PI width	Original model	**3.49 (0.21)**	**2.73 (0.26)**	**3.29 (0.61)**	**3.15 (0.36)**
	Integrative model	4.04 (0.51)	2.90 (0.23)	**3.29 (0.60)**	3.35 (0.45)
95% PI coverage rate	Original model	95.14 (5.57)	92.97 (16.30)	**92.59 (13.30)**	93.39 (11.72)
	Integrative model	**95.14 (6.36)**	**93.49 (14.43)**	92.36 (13.95)	**93.48 (11.58)**

the integrative model slightly improve the coverage rates, although the difference is minimal. Figure 4.6d shows the bias-corrected temperature and its 95% PI when the site ASPLS is used as a testing site. Results are similar to that from the baseline bias correction model in Figure 4.6c, consistent with our summary in Table 4.8. Overall, while the added uncertainty increases prediction error with the integrative model, its performance is comparable to the original bias correction model and is substantially better than alternative approaches.

Finally, Figure 4.7 depicts heat maps showing bias-corrected temperatures for the three heat wave events in the Austin, Texas region. We can clearly see spatially heterogeneous temperature patterns in each case. Ambient temperatures in the region are seen to vary by $1°C \sim 2°C$, which can be explained by the urban heat island effect. Interestingly, the spatial variation pattern for each heat wave event is different. We believe that this is because the regional characteristics (e.g., population, building density) have rapidly changed over the years in this study region. The population in the city of Austin grew by 30% over the 2010-2019 period with rapid urbanization [144], intensifying the urban heat island effects when the heat wave event occurred. This explains the higher temperatures in the central part of Figures 4.7e and 4.7f.

4.5 Summary

Although NWP models can provide useful climate information, they inevitably exhibit systematic bias patterns because of incomplete characterization of local or regional variations. This chapter presents a new probabilistic bias correction approach that takes into consideration the spatially and temporally heterogeneous bias. We formulate the bias correction model with random effect coefficients that vary over



(a) 12:00, 2011

(b) 17:00, 2011

(c) 12:00, 2013

(d) 17:00, 2013

(e) 12:00, 2017

(f) 17:00, 2017

Figure 4.7: Heatmaps for the bias-corrected temperatures as predicted by the integrative bias correction model at 12:00 and 17:00 for three heat wave events in the Austin, Texas region (the darker color is, the higher temperature is; each curve shows the contour plot of the same temperature.)

space and time. A case study using actual temperature data and associated temperature predictions from WRF/UCM demonstrates that the proposed framework can successfully yield reliable solutions for limiting the systematic bias presented in WRF/UCM, while substantially reducing the average RMSE by about 60% from $1.87°C$ to $0.77°C$. Our approach also provides accurate and reliable probabilistic bias correction performance, outperforming the alternative linear MOS and delta change approaches in most performance criteria evaluated in this chapter, including consideration of CRPS, PI width, and coverage rates.

In the future, we plan to evaluate temporal prediction performance (e.g., day-ahead bias correction performance) with larger datasets. In this chapter, due to the limited WRF/UCM data available to us, we were not able to validate the bias correction performance on multiple consecutive days. We intend to extend our approach for providing day(s)-ahead bias-corrected prediction by adaptively learning daily bias patterns. We will also incorporate other factors, such as seasonal effects and socio-economic factors, in the bias correction model in our future work. Lastly, the extension of this work for electricity usage prediction and decision-making in power system operations remains a subject for future study.

Acknowledgement

We thank Earth Networks for providing data collected at ground-based weather stations in central Texas.

CHAPTER V

Conclusion and Future Research Directions

This dissertation study develops statistical methods dealing with multiple data sources in energy and environmental systems, when field environmental data and physics-based model projections are available and/or when data is collected at multiple dispersed locations. We summarize the major accomplishments of this dissertation as follows.

- In Chapter II, we propose a probabilistic approach that can quantify the daily and spatially varying diurnal patterns of the wind speed collected from multiple weather stations. The varying coefficient model enables the proposed approach to provide the capability to evaluate the wind variability at non-observational locations, which provides useful insights for selecting wind farm locations.

- In Chapter III, the integrated model for the long-term density prediction of the daily peak demand is proposed by incorporating the uncertainties associated with the climate, population growth, and socio-economic changes. The long-term temperature prediction is obtained by GCM temperature projections, one of the physics-based models. The proposed integrated model deals with the bias in GCM projections and provides the predictive density of the daily peak demand for the long-term future.

- Chapter IV addresses the systematic bias from the physics-based model output due to incomplete characterization of local or regional variations. We present a new probabilistic bias correction approach that takes account of the spatially and temporally heterogeneous bias from WRF/UCM in an integrative framework. The model can be extended to the case when the WRF/UCM projections at testing locations are absent, which can reduce additional effort and time to re-run the WRF/UCM to obtain the WRF/UCM projections at those locations.

The approaches developed in this dissertation could be further extended in multiple directions, as outlined below.

- We plan to incorporate different levels of prediction time horizons to further improve prediction accuracy. For example, in the wind power system, the physics-based NWP model generates multiple-periods ahead wind forecasting, e.g., 1 hour ahead to 12 hour ahead. In such cases, we can calibrate the bias from the physics-based model sequentially whenever new projections become available. We will investigate online, adaptive learning approaches [145] to further reduce the NWP model bias.

- When data size becomes large, one can face the curse of dimensionality. Recent developments in Bayesian analysis, such as sparse GP approximations [146, 147], will help alleviate the computational burden. We will also explore stochastic optimization approaches that use a subset of data. Variance reduction techniques can be useful when identifying effective samples in stochastic optimization [148–150].

- In this study, stationary spatial and temporal correlation structures are as-

sumed. That is, the correlation intensity is decided based on temporal difference and distance among sites. This might not hold for some applications. For example, in wind power systems, some sites with similar terrains may exhibit analogous power generation patterns, even though they are not located nearby. In the future, we will explore nonstationary correlation patterns to provide more flexibility and improve the applicability of the proposed modeling framework.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] 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.

[2] U.S. Energy Information Administration. Consumption and efficiency. https://www.eia.gov/consumption/, accessed in April, 2021.

[3] U.S. Energy Information Administration. U.S. energy facts explained. https://www.eia.gov/energyexplained/us-energy-facts/, accessed in April, 2021.

[4] M. Maureen Hand, S. Baldwin, E. DeMeo, J.M. Reilly, T. Mai, D. Arent, G. Porro, M. Meshek, and D. Sandor. Renewable electricity futures study. volume 1. exploration of high-penetration renewable electricity futures. Technical report, National Renewable Energy Lab. (NREL), Golden, CO (United States), 2012.

[5] U.S. Energy Information Administration. https://www.eia.gov, accessed in April, 2021.

[6] 20% wind energy by 2030 report. Technical report, U.S. Department of Energy, Washington, D.C. (United States), 2008.

[7] Monica Neukomm, Valerie Nubbe, and Robert Fares. Grid-interactive efficient buildings. Technical report, U.S. Department of Energy, Washington, D.C. (United States), 2019.

[8] A.K. 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)*, pages 130–138. IEEE, 2016.

[9] Fatima Amara, Kodjo Agbossou, Yves Dubé, Sousso 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.

[10] 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.

[11] 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.

[12] 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)*, pages 1–4, 2018.

[13] 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.

[14] Elias Kyriakides and Marios Polycarpou. *Short term electric load forecasting: A tutorial*. Springer, 2007.

[15] 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.

[16] Eugene A. Feinberg and Dora Genethliou. *Load forecasting*. Springer, 2005.

[17] Martin Leutbecher and Tim N. Palmer. Ensemble forecasting. *Journal of computational physics*, 227(7):3515–3539, 2008.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] Abdullah AlShelahi, Jingxing Wang, Mingdi You, Eunshin Byon, and Romesh Saigal. Data-driven prediction for volatile processes based on real option theories. *International Journal of Production Economics*, 226:107605, 2020.

[26] Abdullah Alshelahi, Jingxing Wang, Mingdi You, Eunshin Byon, and Romesh Saigal. Integrative density forecast and uncertainty quantification of wind power generation. *IEEE Transactions on Sustainable Energy*, 2021.

[27] Amanda Lenzi, Pierre Pinson, Line H Clemmensen, and Gilles Guillot. Spatial models for probabilistic prediction of wind power with application to annual-average and high temporal resolution data. *Stochastic Environmental Research and Risk Assessment*, 31(7):1615–1631, 2017.

[28] Swasti R. Khuntia, Jose L. Rueda, and Mart AMM van der Meijden. A multivariate framework to study spatio-temporal dependency of electricity load and wind power. *Wind Energy*, 22(12):1825–1847, 2019.

[29] Ricardo J. Bessa, Artur Trindade, and Vladimiro Miranda. Spatial-temporal solar power forecasting for smart grids. *IEEE Transactions on Industrial Informatics*, 11(1):232–241, 2014.

[30] Hussein Sharadga, Shima Hajimirza, and Robert S. Balog. Time series forecasting of solar power generation for large-scale photovoltaic plants. *Renewable Energy*, 150:797–807, 2020.

[31] M. C. Alexiadis, P. S. Dokopoulos, and H. S. Sahsamanoglou. Wind speed and power forecasting based on spatial correlation models. *IEEE Transactions on Energy Conversion*, 14(3):836–842, 1999.

[32] Tilmann Gneiting, Kristin Larson, Kenneth Westrick, Marc G. Genton, and Eric Aldrich. Calibrated probabilistic forecasting at the stateline wind energy center: The regime-switching space–time method. *Journal of the American Statistical Association*, 101(475):968–979, 2006.

[33] Amanda S. Hering and Marc G. Genton. Powering up with space-time wind forecasting. *Journal of the American Statistical Association*, 105(489):92–104, 2010.

[34] Caroline Persson, Peder Bacher, Takahiro Shiga, and Henrik Madsen. Multi-site solar power forecasting using gradient boosted regression trees. *Solar Energy*, 150:423–436, 2017.

[35] Julija Tastu, Pierre Pinson, Pierre-Julien Trombe, and Henrik Madsen. Probabilistic forecasts of wind power generation accounting for geographically dispersed information. *IEEE Transactions on Smart Grid*, 5(1):480–489, 2013.

[36] Miao He, Lei Yang, Junshan Zhang, and Vijay Vittal. A spatio-temporal analysis approach for short-term forecast of wind farm generation. *IEEE Transactions on Power Systems*, 29(4):1611–1622, 2014.

[37] West Texas Mesonet. `http://www.mesonet.ttu.edu/`, accessed in May, 2008.

[38] Trevor Hastie and Robert Tibshirani. Varying-coefficient models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 55(4):757–779, 1993.

[39] Jianqing Fan and Wenyang Zhang. Statistical methods with varying coefficient models. *Statistics and its Interface*, 1(1):179, 2008.

[40] Byeong U. Park, Enno Mammen, Young K. Lee, and Eun Ryung Lee. Varying coefficient regression models: a review and new developments. *International Statistical Review*, 83(1):36–64, 2015.

[41] Youngchan Jang and Eunshin Byon. Probabilistic characterization of wind diurnal variability for wind resource assessment. *IEEE Transactions on Sustainable Energy*, 11(4):2535–2544, 2020.

[42] Youngchan Jang, Eunshin Byon, Elham Jahani, and Kristen Cetin. On the long-term density prediction of peak electricity load with demand side management in buildings. *Energy and Buildings*, 228:110450, 2020.

[43] Francois Bouffard and Francisco D. Galiana. Stochastic security for operations planning with significant wind power generation. In *2008 IEEE Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century*, pages 1–11. IEEE, 2008.

[44] Jie Zhang, Caroline Draxl, Thomas Hopson, Luca Delle Monache, Emilie Vanvyve, and Bri-Mathias Hodge. Comparison of numerical weather prediction based deterministic and probabilistic wind resource assessment methods. *Applied Energy*, 156:528–541, 2015.

[45] Giwhyun Lee, Eunshin Byon, Lewis Ntaimo, and Yu Ding. Bayesian spline method for assessing extreme loads on wind turbines. *The Annals of Applied Statistics*, 7(4):2034–2061, 2013.

[46] Soon-Duck Kwon. Uncertainty analysis of wind energy potential assessment. *Applied Energy*, 87(3):856–865, 2010.

[47] Sungmoon Jung, O. Arda Vanli, and Soon-Duck Kwon. Wind energy potential assessment considering the uncertainties due to limited data. *Applied Energy*, 102:1492–1503, 2013.

[48] Eduardo Alejandro Martinez-Cesena and Joseph Mutale. Wind power projects planning considering real options for the wind resource assessment. *IEEE Transactions on Sustainable Energy*, 3(1):158–166, 2011.

[49] Barbara Jimenez, Francesco Durante, Bernhard Lange, Torsten Kreutzer, and Jens Tambke. Offshore wind resource assessment with WAsP and MM5: comparative study for the German Bight. *Wind Energy*, 10(2):121–134, 2007.

[50] David Carvalho, Alfred Rocha, M. Gómez-Gesteira, and C. Silva Santos. WRF wind simulation and wind energy production estimates forced by different reanalyses: Comparison with observed data for portugal. *Applied Energy*, 117:116–126, 2014.

[51] Eunshin Byon, Eduardo Pérez, Yu Ding, and Lewis Ntaimo. Simulation of wind farm operations and maintenance using discrete event system specification. *Simulation*, 87(12):1093–1117, 2011.

[52] Johan Lindström, Adam A. Szpiro, Paul D. Sampson, Assaf P. Oron, Mark Richards, Tim V. Larson, and Lianne Sheppard. A flexible spatio-temporal model for air pollution with spatial and spatio-temporal covariates. *Environmental and ecological statistics*, 21(3):411–433, 2014.

[53] Mingdi You, Eunshin Byon, Jionghua (Judy) Jin, and Giwhyun Lee. When wind travels through turbines: A new statistical approach for characterizing heterogeneous wake effects in multi-turbine wind farms. *IISE Transactions*, 49(1):84–95, 2017.

[54] Carl Edward Rasmussen. Gaussian processes in machine learning. In *Advanced lectures on machine learning*, pages 63–71. Springer, 2004.

[55] ML Stein. Interpolation of spatial data: some theory for Kriging, 1999.

[56] Finn Lindgren and Håvard Rue. Bayesian spatial modelling with R-INLA. *Journal of Statistical Software*, 63(19), 2015.

[57] Håvard Rue, Sara Martino, and Nicolas Chopin. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):319–392, 2009.

[58] George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. *Time Series Analysis: Forecasting and Control*. Wiley, 2015.

[59] Mingdi You, Bingjie Liu, Eunshin Byon, Shuai Huang, and Jionghua Jin. Direction-dependent power curve modeling for multiple interacting wind turbines. *IEEE Transactions on Power Systems*, 33(2):1725–1733, 2018.

[60] David J. Spiegelhalter, Nicola G. Best, Bradley P. Carlin, and Angelika Van Der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639, 2002.

[61] Tilmann Gneiting, Adrian E. Raftery, Anton H. Westveld III, and Tom Goldman. Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Monthly Weather Review*, 133(5):1098–1118, 2005.

[62] Eric P. Grimit, Tilmann Gneiting, V. J. Berrocal, and Nicholas A. Johnson. The continuous ranked probability score for circular variables and its application to mesoscale forecast ensemble verification. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 132(621C):2925–2942, 2006.

[63] 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.

[64] Spencer Abraham. National transmission grid study. Technical report, U.S. DOE Office of the Secretary of Energy, Washington, D.C. (United States), 2003.

[65] Eric Hirst. US transmission capacity: present status and future prospects. *Edison Electric Institute*, 2004.

[66] 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.

[67] Da Li, Carol C. Menassa, Vineet R. Kamat, and Eunshin Byon. Heat - human embodied autonomous thermostat. *Building and Environment*, 178:106879, 2020.

[68] Iain Staffell and Stefan Pfenninger. The increasing impact of weather on electricity supply and demand. *Energy*, 145:65–78, 2018.

[69] 2015 Residential Energy Consumption Survey Data. `https://www.eia.gov/consumption/residential/data/2015/`, accessed in July, 2019.

[70] 2018 Commercial Buildings Energy Consumption Survey Data. `https://www.eia.gov/consumption/commercial/data/2018/index.php?view=methodology`, accessed in July, 2019.

[71] 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.

[72] 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.

[73] 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.

[74] 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.

[75] IPCC Sixth Assessment Report. `https://www.ipcc.ch/assessment-report/ar6/`, accessed in July, 2019.

[76] 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.

[77] 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.

[78] 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.

[79] U.S. Energy Information Administration. `https://www.eia.gov/electricity/data/eia861/`, accessed in April, 2020.

[80] Smart Electric Power Alliance. 2018 Utility Demand Response Market Snapshot. `https://sepapower.org/resource/2018-demand-response-market-snapshot/`, accessed in April, 2020.

[81] Frank M. Bass. A new product growth for model consumer durables. *Management Science*, 15(5):215–227, 1969.

[82] 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.

[83] 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.

[84] 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.

[85] 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.

[86] 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.

[87] Matthew Collins, Reto Knutti, Julie Arblaster, Jean-Louis Dufresne, Thierry Fichefet, 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. 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*, pages 1029–1136. Cambridge University Press, 2013.

[88] 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.

[89] Toly Chen. A collaborative fuzzy-neural approach for long-term load forecasting in Taiwan. *Computers & Industrial Engineering*, 63(3):663–670, 2012.

[90] 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.

[91] 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.

[92] 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.

[93] Rob J. Hyndman and Shu Fan. Density forecasting for long-term peak electricity demand. *IEEE Transactions on Power Systems*, 25(2):1142–1153, 2010.

[94] 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.

[95] 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.

[96] Daniel Cooley. Extreme value analysis and the study of climate change. *Climatic change*, 97(1-2):77, 2009.

[97] Richard W. Katz. *Statistical methods for nonstationary extremes*. Springer, 2013.

[98] Linyin Cheng, Amir AghaKouchak, Eric Gilleland, and Richard W. Katz. Nonstationary extreme value analysis in a changing climate. *Climatic change*, 127(2):353–369, 2014.

[99] Iowa Environmental Mesonet, Iowa State University. `https://mesonet.agron.iastate.edu/request/download.phtml?network=TX_ASOS`, accessed in July, 2019.

[100] Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections Archive. `https://gdo-dcp.ucllnl.org/downscaled/cmip/projections`, accessed in July, 2019.

[101] The Electric Reliability Council of Texas (ERCOT). `http://www.ercot.com/gridinfo/load`, accessed in July, 2019.

[102] William Naggaga Lubega and Ashlynn S. Stillwell. Maintaining electric grid reliability under hydrologic drought and heat wave conditions. *Applied Energy*, 210:538–549, 2018.

[103] United States Census Bureau. `https://www.census.gov`, accessed in July, 2019.

[104] Richard W. Katz. Statistics of extremes in climate change. *Climatic change*, 100(1):71–76, 2010.

[105] Stuart Coles, Joanna Bawa, Lesley Trenner, and Pat Dorazio. *An introduction to statistical modeling of extreme values*, volume 208. Springer, 2001.

[106] 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.

[107] 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.

[108] United States Census Bureau. International programs: international data base. `https://www.census.gov/programs-surveys/international-programs/about/idb.html`, accessed in July, 2019.

[109] 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.

[110] 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*, pages 1–2, 2010.

[111] Jorge J. Moré. The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical analysis*, pages 105–116. Springer, 1978.

[112] Gustav Feichtinger. Optimal pricing in a diffusion model with concave price-dependent market potential. *Operations Research Letters*, 1(6):236–240, 1982.

[113] United States Census Bureau. New Residential Construction. `https://www.census.gov/construction/nrc/index.html`, accessed in April, 2020.

[114] Christopher Z. Mooney. *Monte Carlo simulation*, volume 116. Sage publications, 1997.

[115] Frank Heinz. Hottest summer smackdown: 2011 vs. 1980. `https://www.nbcdfw.com/news/local/heat-streak-scorches-on-126138273/2121542/`, 2011.

[116] 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.

[117] 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.

[118] Shuluo Ning, Eunshin Byon, Teresa Wu, and Jing Li. A sparse partitioned-regression model for nonlinear system–environment interactions. *IISE Transactions*, 49(8):814–826, 2017.

[119] William C. Skamarock, Joseph B. Klemp, Jimy Dudhia, David O. Gill, Dale M. Barker, Wei Wang, and Jordan G. Powers. A description of the advanced research WRF version 2. Technical report, National Center For Atmospheric Research, 2005.

[120] Bruno Bueno, Matthias Roth, Leslie Norford, and Reuben Li. Computationally efficient prediction of canopy level urban air temperature at the neighbourhood scale. *Urban Climate*, 9:35–53, 2014.

[121] Feng Chen, Xuchao Yang, and Weiping Zhu. WRF simulations of urban heat island under hot-weather synoptic conditions: The case study of Hangzhou city, China. *Atmospheric Research*, 138:364–377, 2014.

[122] Mat Santamouris. Innovating to zero the building sector in Europe: Minimising the energy consumption, eradication of the energy poverty and mitigating the local climate change. *Solar Energy*, 128:61–94, 2016.

[123] G. Brooke Anderson and Michelle L. Bell. Heat waves in the United States: mortality risk during heat waves and effect modification by heat wave characteristics in 43 US communities. *Environmental health perspectives*, 119(2):210–218, 2011.

[124] T. N. Krishnamurti, J. Sanjay, A. K. Mitra, and T. S. V. Vijaya Kumar. Determination of forecast errors arising from different components of model physics and dynamics. *Monthly Weather Review*, 132(11):2570–2594, 2004.

[125] Kenneth A. Hart, W. James Steenburgh, Daryl J. Onton, and Andrew J. Siffert. An evaluation of mesoscale-model-based model output statistics (MOS) during the 2002 Olympic and Paralympic Winter Games. *Weather and Forecasting*, 19(2):200–218, 2004.

[126] Elham Jahani, Soham Vanage, David Jahn, W. Gallus, and K. S. Cetin. City-scale energy modeling to assess impacts of extreme heat on electricity consumption and production using WRF-UCM modeling with bias correction. In *Canadian Society of Civil Engineers Annual Conference*, 2019.

[127] Estatio Gutiérrez, Jorge E. González, Robert Bornstein, Mark Arend, and Alberto Martilli. A new modeling approach to forecast building energy demands during extreme heat events in complex cities. *Journal of Solar Energy Engineering*, 135(4), 2013.

[128] Thomas Nipen and Roland Stull. Calibrating probabilistic forecasts from an NWP ensemble. *Tellus A: Dynamic Meteorology and Oceanography*, 63(5):858–875, 2011.

[129] Harry R. Glahn and Dale A. Lowry. The use of model output statistics (MOS) in objective weather forecasting. *Journal of Applied Meteorology*, 11(8):1203–1211, 1972.

[130] Aurore Dupré, Philippe Drobinski, Bastien Alonzo, Jordi Badosa, Christian Briard, and Riwal Plougonven. Sub-hourly forecasting of wind speed and wind energy. *Renewable Energy*, 145:2373–2379, 2020.

[131] Peter Berg, H. Feldmann, and H-J Panitz. Bias correction of high resolution regional climate model data. *Journal of Hydrology*, 448:80–92, 2012.

[132] Claudia Teutschbein and Jan Seibert. Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology*, 456:12–29, 2012.

[133] Elham Jahani, Soham Vanage, and Kristen Cetin. City-scale high-resolution WRF-UCM urban weather predictions compared to a dense network of ground-based weather station data for assessment of urban building energy consumption. *ASHRAE Transactions*, 125:248–256, 2019.

[134] Mike West and Jeff Harrison. *Bayesian forecasting and dynamic models*. Springer Science & Business Media, 2006.

[135] George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[136] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014.

[137] Carl Malings, Matteo Pozzi, Kelly Klima, Mario Bergés, Elie Bou-Zeid, and Prathap Ramamurthy. Surface heat assessment for developed environments: Probabilistic urban temperature modeling. *Computers, Environment and Urban Systems*, 66:53–64, 2017.

[138] Christopher M. Bishop. *Pattern recognition and machine learning*. Springer, 2006.

[139] William C. Skamarock, Joseph B. Klemp, Jimy Dudhia, David O. Gill, Dale M. Barker, Wei Wang, and Jordan G. Powers. A description of the advanced research WRF version 3. NCAR technical note-475+ STR. 2008.

[140] NCAR Fei Chen. Current status of urban modeling in the community Weather Research and Forecast (WRF) model. In *Sixth Symposium on the Urban Environment*, 2006.

[141] Hiroyuki Kusaka and Fujio Kimura. Coupling a single-layer urban canopy model with a simple atmospheric model: Impact on urban heat island simulation for an idealized case. *Journal of the Meteorological Society of Japan. Ser. II*, 82(1):67–80, 2004.

[142] Hiroyuki Kusaka, Hiroaki Kondo, Yokihiro Kikegawa, and Fujio Kimura. A simple single-layer urban canopy model for atmospheric models: Comparison with multi-layer and slab models. *Boundary-layer meteorology*, 101(3):329–358, 2001.

[143] Ying Lin, Kaibo Liu, Eunshin Byon, Xiaoning Qian, Shan Liu, and Shuai Huang. A collaborative learning framework for estimating many individualized regression models in a heterogeneous population. *IEEE Transactions on Reliability*, 67(1):328–341, 2018.

[144] United States Census Bureau. Population and housing unit estimates datasets. `https://www-census-gov.proxy.lib.umich.edu/programs-surveys/popest/data/data-sets.html`, 2021. Accessed: 2021-01-30.

[145] Eunshin Byon, Youngjun Choe, and Nattavut Yampikulsakul. Adaptive learning in time-variant processes with application to wind power systems. *IEEE Transactions on Automation Science and Engineering*, 13(2):997–1007, 2016.

[146] James Hensman, Magnus Rattray, and Neil D Lawrence. Fast nonparametric clustering of structured time-series. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(2):383–393, 2014.

[147] Matthias Seeger, Christopher Williams, and Neil Lawrence. Fast forward selection to speed up sparse gaussian process regression. Technical report, 2003.

[148] Youngjun Choe, Eunshin Byon, and Nan Chen. Importance sampling for reliability evaluation with stochastic simulation models. *Technometrics*, 57(3):351–361, 2015.

[149] Qiyun Pan, Eunshin Byon, Young Myoung Ko, and Henry Lam. Adaptive importance sampling for extreme quantile estimation with stochastic black box computer models. *Naval Research Logistics*, 67:524–547, 2020.

[150] Qiyun Pan, Young Myoung Ko, and Eunshin Byon. Uncertainty quantification for extreme quantile estimation with stochastic computer models. *IEEE Transactions on Reliability*, 70:134–145, 2021.