

# ELECTRIC LOAD AND POWER FORECASTING USING ENSEMBLE GAUSSIAN PROCESS REGRESSION

T. Ma,<sup>1</sup> D.A. Barajas-Solano,<sup>1,\*</sup> R. Huang,<sup>1</sup> &  
A.M. Tartakovsky<sup>1,2</sup>

<sup>1</sup>Pacific Northwest National Laboratory, Richland, Washington 99354, USA

<sup>2</sup>University of Illinois Urbana-Champaign, Champaign, Illinois 61820, USA

\*Address all correspondence to: D.A. Barajas-Solano, Pacific Northwest National Laboratory, PO Box 999, MSIN: K7-90, Richland, Washington 99354, USA;  
Tel.: +1 509 375 3783, E-mail: David.Barajas-Solano@pnnl.gov

Original Manuscript Submitted: 11/13/2021; Final Draft Received: 5/10/2022

Accurate week-long forecasting of load demand and generation scheduling is critical for efficient operation of power grid systems. In this work we present an ensemble Gaussian process regression (EGPR) method for week-ahead forecasting of periodic time series data. The proposed EGPR method is based on the GPR method, and employs an ensemble constructed by periodic windowing of the time series data to compute the GPR prior mean and covariance. To improve estimates of prior statistics from a potentially small ensemble and avoid rank-deficiency issues, we propose a leave-one-out cross-validation shrinkage approach to regularizing covariance estimates. Furthermore, we evaluate existing shrinkage estimates available in the literature. A synthetic data set describing the dynamics of a power grid with 700 buses and 134 generators and a real total system load data set from Duke Energy Ohio are used to test the EGPR forecasting method. Both data sets contain load data collected every hour over a 365-day period. The synthetic data set also contains power generation profiles for each generator. We demonstrate that the proposed EGPR method is capable of accurately forecasting weekly total load demand and power generation profiles and outperforms traditional forecasting methods, including the standard data-driven GPR, autoregressive integrated moving average (ARIMA), and TBATS (exponential smoothing state space model with Box-Cox transformation, ARMA errors, trend, and seasonal components) methods.

**KEY WORDS:** *Gaussian process regression, time series forecasting, shrinkage*

## 1. INTRODUCTION

Accurate week-long forecasting of electric power generation and load demand plays an essential role in the efficient management and operation of power grids. Many decisions including system maintenance, security and reliability analysis, and generation scheduling are made based on forecasting (Amjady, 2001). Nonlinear dynamics of power grid in combination with external disturbances, which range from varying weather conditions to fluctuation of economic nature, yields data with complicated patterns. This makes forecasting of the power grid dynamics very challenging. Existing forecasting techniques are typically classified into two categories: statistical approaches and machine learning (ML) methods. Statistical techniques include multiple

linear regression (MLR) models (Charlton and Singleton, 2014; Hong et al., 2013, 2010; Wang et al., 2016; Xie et al., 2016; Xie and Hong, 2016, 2017), semiparametric additive models (Fan and Hyndman, 2011; Goude et al., 2013; Hyndman and Fan, 2009; Nedellec et al., 2014), autoregressive integrated moving average (ARIMA) models (Boroojeni et al., 2017, 2014), and exponential smoothing models (Taylor, 2008; Taylor and McSharpy, 2007). ML techniques include artificial neural network (ANN) (Bianchi et al., 2017; Din and Marnerides, 2017; Qingle and Min, 2010; Ryu et al., 2017; Shi et al., 2017; Yun et al., 2008; Zheng et al., 2017), fuzzy regression models (Hong and Wang, 2014; Song et al., 2005), support vector machines (SVMs) (Chen et al., 2004, 2017), gradient boosting machines (Lloyd, 2014; Taieb and Hyndman, 2014), and Gaussian process regression (GPR) (Lloyd, 2014; Shepero et al., 2018; Yang et al., 2018b).

In MLR models, the load is usually modeled/regressed as a function of weather and calendar variables (e.g., workdays, weekends, and national and local holidays). For example, an integrated MLR framework was proposed for short-term load forecasting (STLF) in Hong et al. (2010) that emphasizes the interactions (or cross effects) among weather and calendar variables. A parametric model for STLF in Charlton and Singleton (2014) estimated the electricity demand as a function of the temperature and calendar variables. A linear regression model with a macroeconomic indicator was developed for long-term load forecasting (LTLF) in Hong et al. (2013). This model was further extended in Wang et al. (2016) by including a large number of lagged temperature and moving average temperature variables in the MLR models. A forecasting model including a relative humidity variable and its polynomial transform was developed in Xie et al. (2016). The shifted-date temperature scenario method was used in Xie and Hong (2016) for probabilistic load forecasting. A probabilistic error measure was employed in Xie and Hong (2017) to identify relevant variables for probabilistic forecasting. In MLR models, a challenging problem is to leverage the increased computing power to build large regression models to enhance the load forecast accuracy. The semiparametric additive model falls within the regression framework, but is designed to accommodate some nonlinear relationships and serially correlated errors. In load forecasting, nonlinear and non-parametric terms are allowed within the framework of additive models for estimating the relationship between the load and explanatory variables such as temperature and calendar variables. The relationships between demand and the driver variables for STLF and LTLF were estimated in Hyndman and Fan (2009) and Fan and Hyndman (2011), respectively.

Many types of ANNs have been used for load forecasting, such as feed-forward neural networks, radial basis function (RBF) networks, and recurrent neural networks (RNN). A RBF neural network model with the adaptive neural fuzzy inference system (ANFIS) was developed in Yun et al. (2008) for STLF to account for the influence of real-time electricity prices on short-term load. A novel pooling-based deep recurrent neural network was proposed in Shi et al. (2017) for household load forecasting. A comparative analysis of RNN for STLF can be found in Bianchi et al. (2017). Although these proposed systems were tested on real data, some of them did not provide comparisons with standard benchmarks, while others did not follow standard statistical procedures in reporting the analysis of errors. In addition, ANN models may “overfit” the data, possibly due to either overtraining or overparameterization. Fuzzy regression was introduced in order to overcome some of the limitations of linear regression, such as a vague relationship between dependent and independent variables, insufficient numbers of observations, and hard-to-verify error distributions. For example, fuzzy linear regression was used in Song et al. (2005) for STLF during holidays. A fuzzy interaction regression method for STLF was proposed in Hong and Wang (2014) and shown to outperform fuzzy and linear regression models that did not account for interaction effects. SVMs are supervised learning models with

associated learning algorithms that are commonly used for pattern recognition, classification, and regression. An SVM model for mid-term load forecasting (predicting daily maximum load of the next 31 days) was proposed in Chen et al. (2004). In Chen et al. (2017) a new support vector regression (SVR) forecasting model was proposed with the ambient temperature of two hours before demand response (DR) event as input variables. The gradient boosting method was used in Taieb and Hyndman (2014) and Lloyd (2014) for the load forecasting track of GEFCom2012. This machine learning technique for regression problems provides a forecast in the form of an ensemble of weak prediction models.

Since some types of load demand are driven strongly by the weather, e.g., heating and cooling, changes in weather conditions have a significant effect on the load profiles. As such, methods that do not use weather or calendar variables as explanatory variables may be at a disadvantage. Nevertheless, in the present work we restrict our attention to developing forecasting methods that do not use explanatory variables, as they have lower data requirements.

ARIMA and the exponential smoothing method are examples of univariate models that do not rely on explanatory variables. ARIMA provides a parsimonious description of a stationary stochastic process in terms of two polynomials, autoregression and moving average. Based on maximum-likelihood estimator (MLE) and ARIMA models, Boroojeni et al. (2014) proposed a novel two-tier scheme for forecasting the power demand and generation in a general residential electrical grid which uses the distributed renewable resources as the primary energy resource. In Boroojeni et al. (2017) ARIMA is employed to model historical load data in the form of a time series with different cycles of seasonality (e.g., daily, weekly, quarterly, annually) in a given power network without requiring any additional inputs such as historical weather data (which might not be available in some cases). The exponential smoothing method assigns weights to past observations that decrease exponentially over time (Hong and Fan, 2016; Taylor, 2008; Taylor and McSharry, 2007).

GPR is another univariate model that does not rely on explanatory variables for load forecasting (Lloyd, 2014; Shepero et al., 2018; Yang et al., 2018b). In the standard data-driven GPR method, a time series is assumed to be a realization of a Gaussian process with prescribed parameterized mean and covariance functions. The parameters of these mean and covariance functions are learned from the time series measurements by maximizing the marginal likelihood function of the measurements or other pseudolikelihoods.

To improve the forecasting performance of GPR, we propose the ensemble GPR (EGPR) method. In EGPR, we compute the prior mean vectors and covariance matrices from ensembles composed of subsets of the observation time series. For example, for weekly load forecasting, we treat the  $N$  weeks prior to the beginning of the forecast week as realizations of the same Gaussian process. Then, the mean and covariance are computed as the ensemble statistics of the ensemble formed by these  $N$  weeks. The number  $N$  cannot be too large as the statistics could be affected by seasonal variations. On the other hand,  $N$  should not be too small as to result in inaccurate, noisy ensemble statistics. A synthetic data set describing a power grid with 700 buses and 134 generators and a real load data set from Duke Energy Ohio are used to validate the proposed forecasting method. Both data sets are collected at an hourly rate over a 365-day period. We demonstrate that the proposed EGPR outperforms traditional forecasting methods, including standard data-driven GPR and the ARIMA and TBATS Livera et al. (2011) methods. We also employ shrinkage estimation (Chen et al., 2009; Steland, 2018) to regularize the covariance matrix estimates with the goal of overcoming the rank-deficiency problem arising from estimating covariance matrices from small ensembles. The shrinkage coefficient is estimated via a novel leave-one-out cross validation (LOOCV)-based approach and by using the

Rao-Blackwell Ledoit-Wolf (RBLW) and oracle approximating shrinkage (OAS-EGPR) estimates presented in Chen et al. (2009). Comparisons among these methods are carried out to illustrate their performances.

The EGPR method is conceptually similar to the physics-informed GPR method (Tartakovsky and Tipireddy, 2019; Yang et al., 2019). In physics-informed GPR, ensemble realizations of the states to be estimated are generated by repeatedly sampling stochastic models of the dynamics to be forecasted. For example, in Tartakovsky and Tipireddy (2019) physics-informed GPR was used for the forecasting of the dynamics of a single-machine infinite-bus (SMIB) system powered by random mechanical wind power fluctuations. The mechanical wind power was treated as a random process and the ensemble was generated by repeatedly solving the SMIB swing equations. An open question is how to apply this ensemble-based approach to GPR for forecasting problems where either there is no stochastic model or the stochastic model is computationally costly to evaluate, but historical time series observations are available. One such problem is the forecasting of load demand and generation scheduling for large power grid systems as considered in this work. For this problem, repeatedly solving the power flow and economic dispatch problems is computationally costly, and we aim to calculate forecasts using historical data.

This paper is organized as follows. In Section 2, we introduce the EGPR method and the LOOCV-EGPR, RBLW-EGPR, and OAS-EGPR extensions. Section 3 describes the synthetic data set of load demand and generation scheduling for a power grid system with 700 buses and 134 generators. EGPR is applied to this synthetic data set for weekly forecasting of load demand and generation scheduling. Section 4 presents the application of EGPR to the Duke Energy Ohio data set for weekly forecasting of load demand. Comparisons between EGPR and other methods are carried out in both Section 3 and Section 4. Final conclusions are given in Section 5.

## 2. EGPR FOR ELECTRIC POWER GRID FORECASTING

In this section, we formulate the EGPR method for forecasting states of the power grids using historical observations of the states. Furthermore, we present several methods for overcoming the rank deficiency in computing covariances from small ensembles, including LOOCV-EGPR, RBLW-EGPR, and OAS-EGPR.

### 2.1 EGPR Method

The proposed EGPR method, as all GPR-based methods, uses the covariances and cross covariances of and between observed and forecasted times to evaluate a probabilistic (Gaussian) estimate at the forecasted times. Consider the process of time  $x(t)$ , for which there are (noiseless) observations available at times  $t_i^o$ ,  $t = 1, \dots, N_o$ ; we aim to estimate  $x(t)$  at times  $t_i^f$ ,  $i = 1, \dots, N_f$ . In GPR, we model  $x(t)$  as a realization of the Gaussian process (GP) of time  $X(t)$ . Let  $X^\top = [(X^o)^\top, (X^f)^\top]$  be the vector of values of  $X(t)$  at times  $t_i^o$  and  $t_i^f$ , with  $X^o = [X(t_1^o), \dots, X(t_{N_o}^o)]^\top$ ,  $X^f = [X(t_1^f), \dots, X(t_{N_f}^f)]^\top$ , and distribution

$$\begin{bmatrix} X^o \\ X^f \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \bar{X}^o \\ \bar{X}^f \end{bmatrix}, \begin{bmatrix} K^{oo} & (K^{fo})^\top \\ K^{fo} & K^{ff} \end{bmatrix} \right). \quad (1)$$

$\bar{X}^o$  and  $\bar{X}^f$  are the so-called prior (or unconditional) mean vector of  $X^o$  and  $X^f$ , respectively;  $K^{oo}$ , and  $K^{ff}$  are the prior covariance matrices of  $X^o$  and  $X^f$ , respectively; and  $K^{fo}$

is the cross covariance between  $X^f$  and  $X^o$ . The covariance and cross covariance matrices are given by

$$K^{\alpha\beta} = \begin{bmatrix} \kappa(t_1^\alpha, t_1^\beta) & \cdots & \kappa(t_1^\alpha, t_{N_\beta}^\beta) \\ \vdots & \ddots & \vdots \\ \kappa(t_{N_\alpha}^\alpha, t_1^\beta) & \cdots & \kappa(t_{N_\alpha}^\alpha, t_{N_\beta}^\beta) \end{bmatrix},$$

for  $\alpha, \beta \in (o, f)$ , where  $\kappa(\cdot, \cdot)$  is the Gaussian process's covariance kernel.

Let  $x^o = [x(t_1^o), \dots, x(t_{N_o}^o)]^\top$  denote the vector of observed values of  $x(t)$ . Given  $x^o$  and  $\kappa(\cdot, \cdot)$ , the conditional (or posterior) mean and covariance of  $X^f$  are given by

$$\hat{X}^f = \bar{X}^f + K^{fo}(K^{oo})^{-1}(x^o - \bar{X}^o), \quad (2)$$

$$\hat{K}^{ff} = K^{ff} - K^{fo}(K^{oo})^{-1}(K^{fo})^\top. \quad (3)$$

The posterior means serves as the GPR forecast estimate, while the posterior covariance provides a measure of uncertainty of this estimate.

Estimation of the prior statistics  $\bar{X}^o$ ,  $\bar{X}^f$ ,  $K^{oo}$ ,  $K^{fo}$ , and  $K^{ff}$  is one of the main challenges in GPR-based methods. The standard data-driven GPR method assumes parameterized forms for these prior statistics; these so-called “hyperparameters” are estimated by minimizing the negative marginal likelihood or another pseudolikelihood of the measurements  $x^o$  (Lloyd, 2014; Shepero et al., 2018; Yang et al., 2018b).

In EGPR, we assume that the data are periodic, and that each period can be accurately modeled as realizations of a certain random process. Then, the prior statistics can be estimated via sample averaging from an ensemble constructed by periodic windowing of the historical data. Therefore, in EGPR the estimation of the prior statistics strongly depends on the structure of the data.

To illustrate this process consider a data set with uniform frequency and a weekly period consisting of  $M$  weeks of observations, and let the end of the data set correspond to the present time. We aim to forecast 1 week into the future from the last 24 hr of observations. The EGPR ensembles are constructed as follows: First, we split the data set into weeks, and select the last  $N \leq M$  weeks of data to construct the EPGR ensemble. We employ the counter  $i = 1, 2, \dots, N$  to denote the last week in the data set, the next to last week, etc., up to the  $N$ th last week. Next, we denote by  $L_i^f$  the  $i$ th last week of data, and by  $L_i^o$  the 24 hr of data prior to the  $i$ th last week. This procedure is illustrated in Fig. 1.

Once the EGPR ensemble is constructed, the unconditional statistics are computed as

$$\bar{X}^\alpha = \frac{1}{N} \sum_{i=1}^N L_i^\alpha, \quad K^{\alpha\beta} = \frac{1}{N-1} \sum_{i=1}^N (L_i^\alpha - \bar{L}^\alpha)(L_i^\beta - \bar{L}^\beta)^\top, \quad (4)$$

for  $\alpha, \beta \in (o, f)$ , where  $L_i^o$  and  $L_i^f$  are understood to be vector columns. Once the prior statistics  $\bar{X}^o$ ,  $\bar{X}^f$ ,  $K^{oo}$ ,  $K^{fo}$ , and  $K^{ff}$  have been estimated, the forecast mean and covariance can be computed using Eqs. (2) and (3).

As stated above, the number of ensemble members,  $N$ , can be chosen to be less or equal to the number of weeks in the data,  $M$ . This choice of  $N$  affects the performance of EGPR. For data with seasonal periods,  $N$  cannot be too large as to violate the assumption that all  $N$  periods of the ensemble can be modeled as realizations of the same Gaussian random process. On the other hand, small ensembles may result in rank deficiency in the estimation of the covariance and cross covariance matrices. In Sections 3 and 4, we select  $N$  by analyzing the structure of

the resulting ensemble covariances. In Section 2.2, we present several methods for covariance estimation from small ensembles.

## 2.2 Techniques to Overcome Rank Deficiency

It is known that the sample covariance estimate of Eq. (4) has rank of at most  $N - 1$ , possibly leading to singular covariance matrices when  $N$  is small. To address this challenge, we propose replacing the sample covariance estimate of (4) with the regularized “shrinkage” estimate,

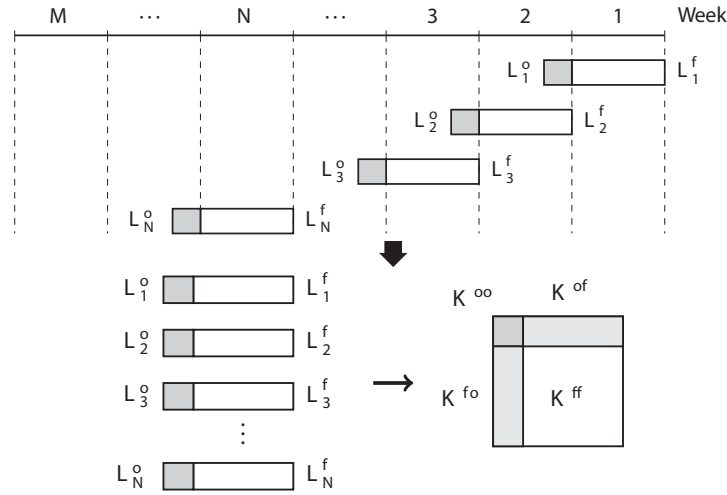
$$K_{\text{reg}} = \rho_{\text{reg}} \frac{K_r}{\text{tr}(K_r)} \text{tr}(K) + (1 - \rho_{\text{reg}}) K, \quad (5)$$

where  $K_{\text{reg}}$  denotes the regularized estimate,  $\text{tr}(\cdot)$  denotes the trace operator,  $K_r$  denotes a “correction” matrix, and  $\rho_{\text{reg}} \in [0, 1]$  is a regularization coefficient. This estimate mixes the sample estimate (4) with the correction matrix  $K_r$  to ensure that the regularized estimate is full rank.  $K_r$  is often chosen to be the identity matrix; alternatively, one can choose  $K_r$  to be a full-rank covariance matrix similar to the covariance structure of the data.

In the remainder of this section we discuss how to select  $K_r$  and  $\rho_{\text{reg}}$  in Eq. (5). First, we propose a novel leave-one-out cross validation (LOOCV)-based shrinkage estimator. Additionally, we also discuss two additional shrinkage estimators often used in practice, namely, Oracle approximating shrinkage (OAS) and Rao-Blackwell-Ledoit-Wolf (RBLW) (Chen et al., 2009).

For LOOCV, we assume that the random process we employ to model the data has a periodic and a nonperiodic component; therefore, we choose  $K_r$  as the covariance matrix corresponding to the semiperiodic kernel,

$$\kappa_r(t_1, t_2) = \theta_1^2 \exp \left( -\frac{|t_1 - t_2|^2}{2\theta_2^2} - \frac{2}{\theta_3^2} \sin^2 \left[ \frac{\pi}{24}(t_1 - t_2) \right] \right), \quad (6)$$



**FIG. 1:** Construction of the EGPR ensemble. Top: the data are split into  $M$  weeks and the most recent  $N \leq M$  weeks of data are selected. For the  $i$ th week,  $L_i^f$  denotes the  $i$ th week of data, and  $L_i^o$  denotes the 24 hr of data preceding said week. Bottom: the sample mean and covariance of the ensemble of  $N$  weeks of data are used as the prior statistics for GPR estimation.

where  $\theta_1$  denotes the kernel amplitude,  $\theta_2$  controls the correlation across 24-hr periods, and  $\theta_3$  controls the hourly (i.e., hour-to-hour) correlation. The kernel hyperparameters ( $\theta_1$ ,  $\theta_2$ , and  $\theta_3$ ) and the regularization coefficient  $\rho_{\text{LOOCV}}$  are estimated by minimizing the negative marginal likelihood of the observations, given by

$$-\log p = \frac{1}{2} \frac{(x^o - \bar{X}^o)^\top (x^o - \bar{X}^o)}{K_{\text{LOOCV}}^{oo}} + \frac{1}{2} \log |K_{\text{LOOCV}}^{oo}| + \frac{1}{2} N_o \log 2\pi, \quad (7)$$

where  $K^{\text{LOOCV}}$  is given by Eq. (5) with  $\rho_{\text{reg}} = \rho_{\text{LOOCV}}$  and  $K_r$  is given by Eq. (6), and  $K_{\text{LOOCV}}^{oo}$  is the  $oo$  portion of  $K_{\text{LOOCV}}$  as illustrated in Fig. 1.

Various alternative shrinkage approaches have been proposed in the literature. In this work we study the Oracle approximating shrinkage (OAS) and Rao-Blackwell-Ledoit-Wolf (RBLW) estimators proposed in Chen et al. (2009). Both of these estimators use  $K_r = I$  and have regularization parameters given in closed form, namely,

$$\rho_{\text{RBLW}} = \min \left\{ \frac{[(N-2)/N] \text{tr}(K^2) + \text{tr}^2(K)}{(N+2) [\text{tr}(K^2) - \text{tr}^2(K)/(N_o + N_f)]} \right\},$$

$$\rho_{\text{OAS}} = \min \left\{ \frac{(1 - 2/(N_o + N_f)) \text{tr}(K^2) + \text{tr}^2(K)}{(N+1 - 2/(N_o + N_f)) [\text{tr}(K^2) - \text{tr}^2(K)/(N_o + N_f)]} \right\},$$

where  $K$  is the sample estimate (4).

We note that the RBLW and OAS estimators can be computed directly from the sample covariance estimate, whereas the LOOCV estimator requires solving an additional unconstrained minimization problem; nevertheless, the LOOCV estimator provides flexibility in modeling the features of the covariance. In the sequel, we will denote by “EGPR” the EGPR method without shrinkage, and by “LOOCV-EGPR,” “RBLW-EGPR,” and “OAS-EGPR” the EGPR methods that use LOOCV, RBLW, and OAS covariance estimation, respectively.

### 3. WEEKLY FORECASTING OF LOAD DEMAND AND GENERATION SCHEDULING FOR A SYNTHETIC DATA SET

In this section we apply the proposed EGPR method to a synthetic load demand and generation scheduling data set in order to produce week-ahead forecasts of both load demand and generation scheduling.

#### 3.1 Generation of the Synthetic Data Set

We consider synthetic data of a power grid system with 700 buses and 134 generators over a 365-day period, with hourly measurements of load and power generation for each node. The synthetic data set is constructed as follows: First, we generate the chronological system-level total load for 1 year with 1-hr time resolution based on the data and method provided in Grigg et al. (1999). Then, the “base case” power flow is generated as described in Young et al. (2018). The historical Duke Energy hourly load shape (Duke Energy Ohio Inc., 2021) is used to develop hourly load values with the same participation factor as the original “base case” power flow. The peak load of the system is 12,926 MW, occurring on December 17 at 18:00. Second, the total system load profile and the generator parameters are fed into a unit commitment and hourly dispatch program, which outputs the on/off status and the real power output of each generator

with hourly time resolution. Third, the chronological alternating-current power flow is computed for the power grid using the PSS/E software to generate different power flow scenarios with hourly time resolution. The chronological system-level load profile is disaggregated to produce load for each bus in the system and is used as input to PSS/E, together with the on/off status and real power of each generator from the unit commitment and hourly dispatch results. PSS/E provides a converged power flow solution for each scenario. Finally, we reinforce scenarios by adjusting voltage and line flows, multiblock switchable shunts, and lines.

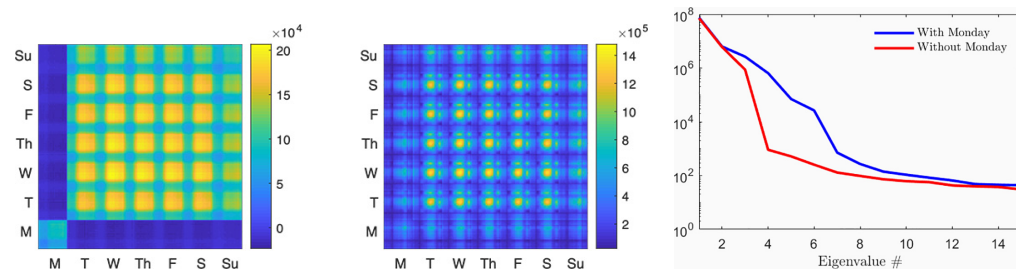
We use this data set to demonstrate and validate the proposed EGPR methods for weekly forecast of the total load and generation scheduling (specifically the real power output of each generator; the on/off status of each generator is not considered in this work). Specifically, we forecast both the total load demand and the generation scheduling for Generator 15 for the weeks of 6/16–6/22, 8/18–8/24, 10/20–10/26, and 12/22–12/28. These four weeks represent the seasons of summer, fall, and winter. For total load demand forecasting, comparisons between EGPR, standard data-driven GPR, ARIMA, and TBATS are carried out.

### 3.2 Weekly Forecasting of Total Load

#### 3.2.1 EGPR

To employ the proposed EGPR method, we split the synthetic data into weekly time series. The EGPR ensemble is constructed using second to last to the  $N + 1$ th last week of the data set, and we aim to forecast the last week of data, which we employ as reference. Figure 2 shows the ensemble load covariance for the week of 6/16–6/22, computed using  $N = 5$  and  $N = 20$ . We observe that in the synthetic data set, Monday's total load demand is weakly correlated to the total load demand of the other weekdays, while the loads on Tuesday to Friday are strongly correlated with each other. Specifically, there is practically no correlation between loads on Monday and the other weekdays (the covariance is near zero) for all considered weeks if the ensemble is not large enough (e.g.,  $N = 5$ ). Increasing  $N$  to 20 produces stronger correlation between loads on Monday and the other weekdays.

Figure 2 also shows the eigenvalues of the ensemble covariances computed using  $N = 5$  and  $N = 20$ . It can be seen that the eigenvalues of the covariance for Tuesday to Sunday decay faster than those of the covariance for Monday to Sunday. Therefore, the GP representation of Tuesday to Sunday data has a lower random dimension (the number of “significant” eigenvalues) than the GP representation of the Monday to Sunday data. Both GP representations have less than 15

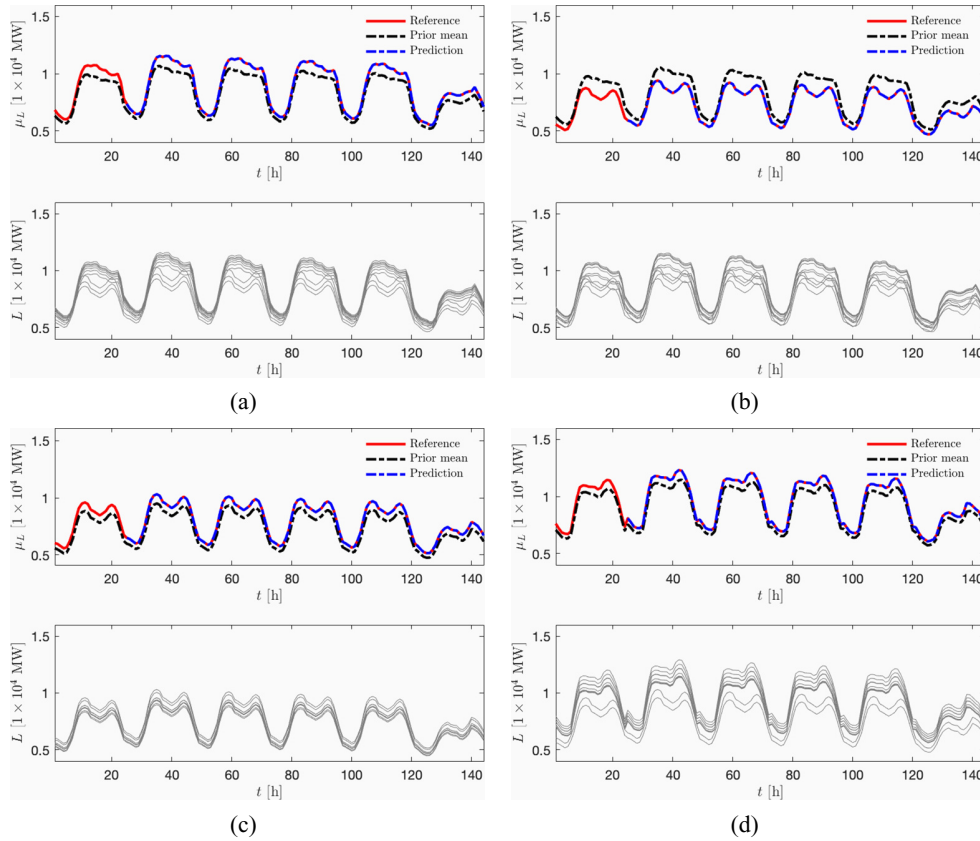


**FIG. 2:** Ensemble covariance for the week of 6/16–6/22 computed using five realizations (left) and 20 realizations (middle). Eigenspectrum of ensemble covariance for the week of 6/16–6/22 computed using 20 realizations, with and without Monday.

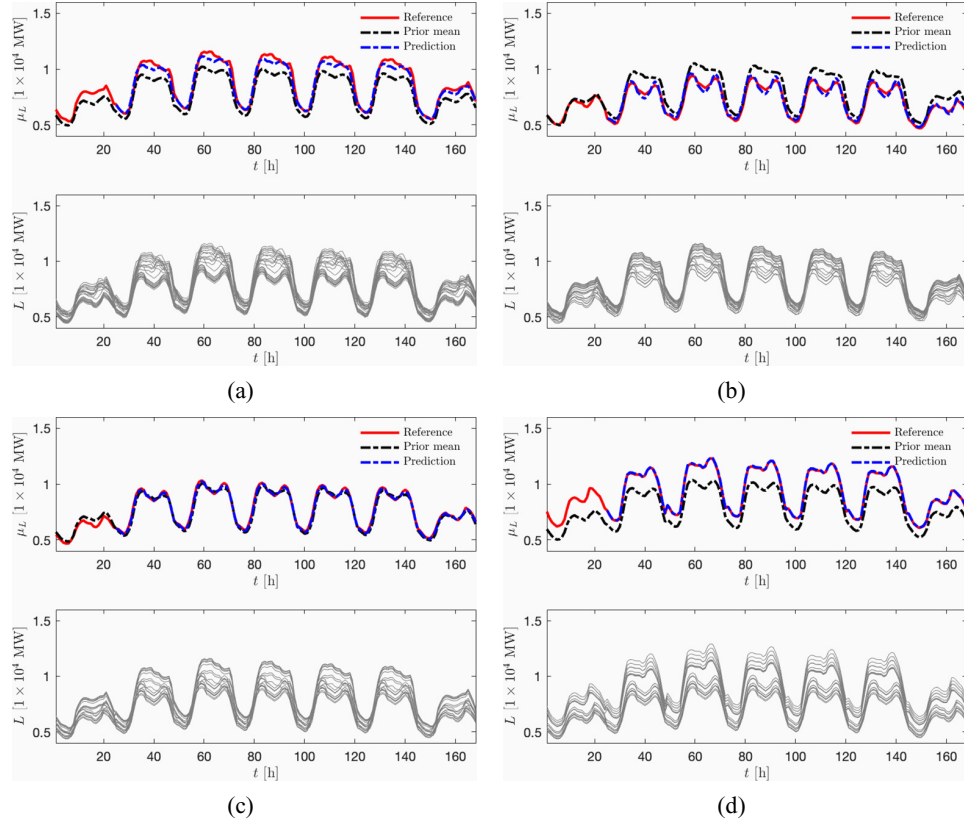
significant eigenvalues; therefore, we expect that accurate ensemble statistics can be estimated using relatively small ensembles. We note here that the real load data studied in Section 4 does not exhibit the same correlation pattern as this synthetic data set; this pattern can be in fact considered as an artifact of the synthetic data set generation process. However, we consider this synthetic data set in this section as it helps us investigate the accuracy of the proposed EGPR method.

Given the different correlation between Monday and the other weekdays, we perform two forecasts: (i) forecasting Wednesday to Sunday given 24 hr data on Tuesday with  $N = 10$ , and (ii) forecasting for Tuesday to Sunday using 24 hr data on Monday with  $N = 20$ . We perform these two load forecasts for the four different weeks previously listed and show the comparison of these forecasts against the reference data on Figs. 3 and 4, respectively. Figures 3 and 4 also show the individual ensemble members and the prior mean.

Figure 3 shows that the Wednesday to Sunday forecast using 24 hr of data on Tuesday closely agrees with the reference total load demand for all tested weeks. This figure also shows that there is significant variability in the load in the previous  $N$  weeks and that the prior mean provides a



**FIG. 3:** Weekly forecasts of total load for 4 weeks using 24 hourly observations on Tuesday. Top panel: prediction (blue) compared against reference (red) and the prior mean (black). Bottom panel: ensemble of ten time series used to compute the empirical covariance. (a) 6/16–6/22, (b) 8/18–8/24, (c) 10/20–10/26, and (d) 12/22–12/28.



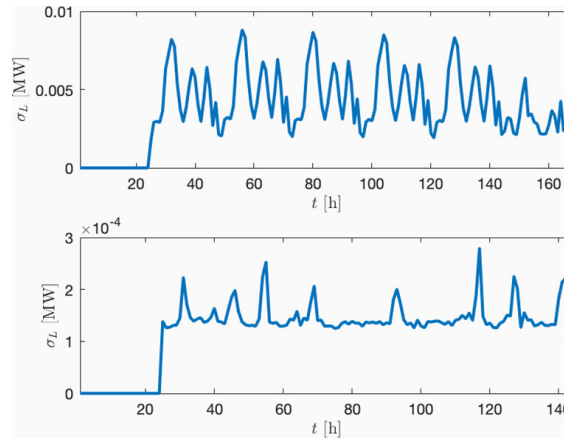
**FIG. 4:** Weekly forecasts of total load for 4 weeks using 24 hourly observations on Monday. Top panel: prediction (blue) compared against reference (red) and the prior mean (black). Bottom panel: ensemble of 20 time series used to compute the empirical covariance. (a) 6/16–6/22, (b) 8/18–8/24, (c) 10/20–10/26, and (d) 12/22–12/28.

less accurate prediction than EGPR. Similarly, Fig. 4 shows that the Tuesday to Sunday forecast using 24 hr of data on Monday data is less accurate, especially for the weeks of 6/16–6/22 and 8/18–8/24. Nevertheless, the EGPR forecast is better than the prior mean. Results in Figs. 3 and 4 demonstrate that the strong correlation between past and future values of the state to be forecasted is important for the performance of EGPR.

Figure 5 shows the posterior standard deviations for the forecast of the week of 6/16–6/22 using 24 hr data on Tuesday and Monday. The posterior standard deviation, which provides a measure of uncertainty in the forecast, is approximately two orders of magnitude smaller for the Wednesday to Sunday forecast than for the Tuesday to Sunday forecast. This shows that the forecast using Monday data is more uncertain, in addition to being less accurate, than the forecast based on the Tuesday data.

### 3.2.2 Comparison against Standard Data-Driven GPR, ARIMA, and TBATS

In this section we compare the performance of EGPR with the state-of-the-art forecasting methods, specifically against standard data-driven GPR, ARIMA, and TBATS. To evaluate the



**FIG. 5:** Posterior standard deviation of the weekly forecast for the week of 6/16–6/22. Top: using 24 hr observations on Tuesday. Bottom: using 24 hr observations on Monday.

forecasting performance of EGPR against GPR we employ the log predictive probability, which corresponds to the sum of the pointwise log probabilities of the reference values being observed given the statistical forecast (Williams and Rasmussen, 2006). For a certain estimated state  $\alpha(t)$ , it is given by

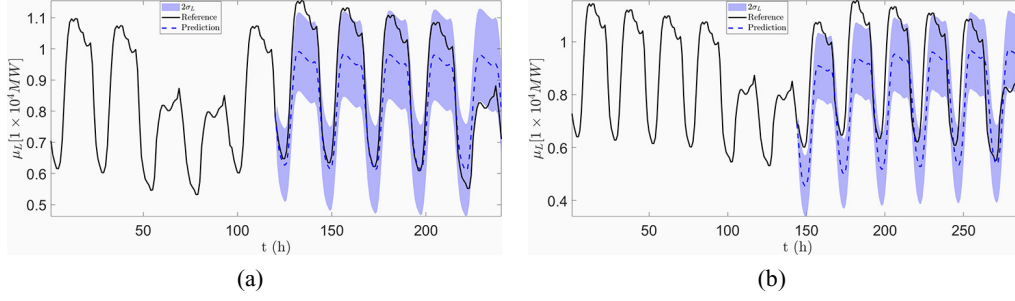
$$\log \text{ predictive probability} = - \sum_{k=1}^{N_f} \left\{ \frac{[\mu^f(t_k) - \alpha(t_k)]^2}{2 [\sigma^f(t_k)]^2} + \frac{1}{2} \log 2\pi [\sigma^f(t_k)]^2 \right\},$$

where  $N_f$  is the number of forecasted values;  $\mu^f(t_k)$  and  $\sigma^f(t_k)$  are the posterior mean and standard deviation of the forecast at time  $t_k$ , respectively; and  $\alpha(t_k)$  is the reference value at time  $t_k$ . The larger the log predictive probability, the more accurate is the model estimation or forecast.

The standard data-driven GPR forecast is computed using as parameterized covariance kernel the following combination of a squared exponential, rational quadratic, and periodic kernels:

$$\begin{aligned} \kappa(t_1, t_2) = & \theta_4^2 \exp \left( -\frac{|t_1 - t_2|^2}{2\theta_5^2} \right) + \theta_6^2 \left( 1 + \frac{|t_1 - t_2|^2}{2\theta_7\theta_8^2} \right)^{-\theta_7} \\ & + \theta_9^2 \exp \left( -\frac{2}{\theta_{10}^2} \sin^2 \left[ \frac{\pi}{24}(t_1 - t_2) \right] \right) + \theta_{11}^2 \delta(t_1, t_2), \end{aligned} \quad (8)$$

where  $\delta$  denotes the Dirac delta function, and  $\theta_i, i = 4, \dots, 11$  are the kernel hyperparameters. Such a combination of periodic and nonperiodic have been employed in other studies to model data exhibiting periodicity (see, e.g., Grunblatt et al., 2015; Klenske et al., 2016; Tolba et al., 2019; Wilson and Adams, 2013). These parameters are estimated by minimizing the negative marginal likelihood function from previous  $N_f$  observed data. Figure 6 shows the standard data-driven GPR forecast of total load for Wednesday to Sunday and Tuesday to Sunday for the week of 06/16–06/22, respectively. The log predictive probabilities of EGPR and standard data-driven GPR for weekly forecasts of load using 24 hourly observations on Tuesday and Monday are presented in Table 1. As we can see from Fig. 6 (in comparison with Figs. 3 and 4) and Table 1, the proposed EGPR method outperforms standard data-driven GPR.



**FIG. 6:** Weekly total load forecasting using data-driven GPR for the week of 06/16–06/22, using 24 hourly observations on Tuesday (a) and Monday (b)

**TABLE 1:** Log predictive probabilities for weekly forecasts of load demand using 24 hourly observations on day  $d$

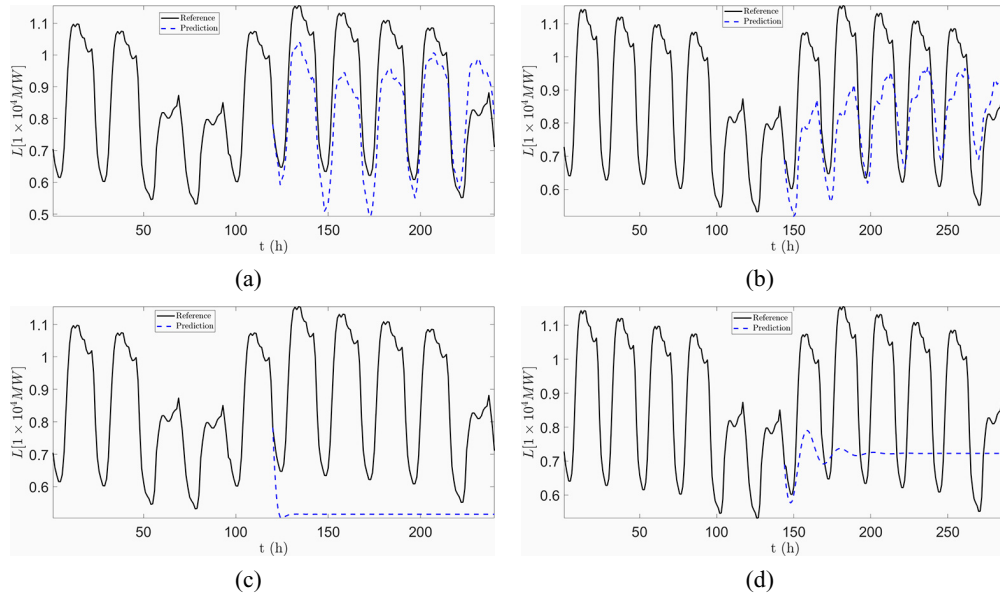
| $d$     | Method          | 06/16–06/22 | 08/18–08/24 | 10/20–10/26 | 12/22–12/28 |
|---------|-----------------|-------------|-------------|-------------|-------------|
| Tuesday | EGPR            | 642.825     | 636.445     | 706.173     | 610.729     |
|         | Data-driven GPR | 198.142     | 239.694     | 173.26      | 156.91      |
| Monday  | EGPR            | 358.162     | 385.995     | 347.498     | 378.112     |
|         | Data-driven GPR | 263.412     | 273.605     | 224.426     | 250.085     |

We now proceed to compare EGRP against ARIMA and TBATS. Specifically, we employ the nonseasonal ARIMA method (Brockwell and Davis, 2016; Hyndman and Athanasopoulos, 2018) to forecast total load demand. The ARIMA forecast is given as a linear combination of  $p$  measurements prior to the forecast time  $t_k$ , that is,

$$x(t_k) - \alpha_1 x(t_{k-1}) - \alpha_2 x(t_{k-2}) - \cdots - \alpha_p x(t_{k-p}) = e_k \\ + \beta_1 e_{k-1} + \beta_2 e_{k-2} + \cdots + \beta_q e_{k-q},$$

where  $x(t_k)$  is the forecast at time  $t_k$ ;  $p$  and  $\alpha_i$ ,  $i = 1, \dots, p$  are the order and the parameters of the autoregressive part, respectively;  $q$  and  $\beta_i$ ,  $i = 1, \dots, q$  are the order and parameters of the moving averaging part, respectively; and the  $e_i$ ,  $i = 1, \dots, q$  are zero-mean independent normally distributed error terms. Nonseasonal ARIMA models are generally denoted as ARIMA( $p, d, q$ ),  $p$  and  $q$  are given above, and  $d$  is the degree of differencing (i.e., the number of times the data have had past values subtracted). Additionally, we also employ the TBATS method to forecast total load demand. The TBATS model is a time series model for series exhibiting multiple complex seasonalities (Livera et al., 2011), which uses a combination of Fourier terms and an exponential smoothing state space model with Box-Cox transformation, ARMA errors, and trend and seasonal components, in an automated manner.

Figure 7 shows the ARIMA and TBATS forecasts of total load for Wednesday to Sunday and Tuesday to Sunday of the week of 06/16–06/22, respectively. The ARIMA models for Fig. 7 are ARIMA(40, 0, 1) and ARIMA(25, 0, 1), respectively. It can be seen that both ARIMA and TBATS perform worse than the proposed EGPR method.



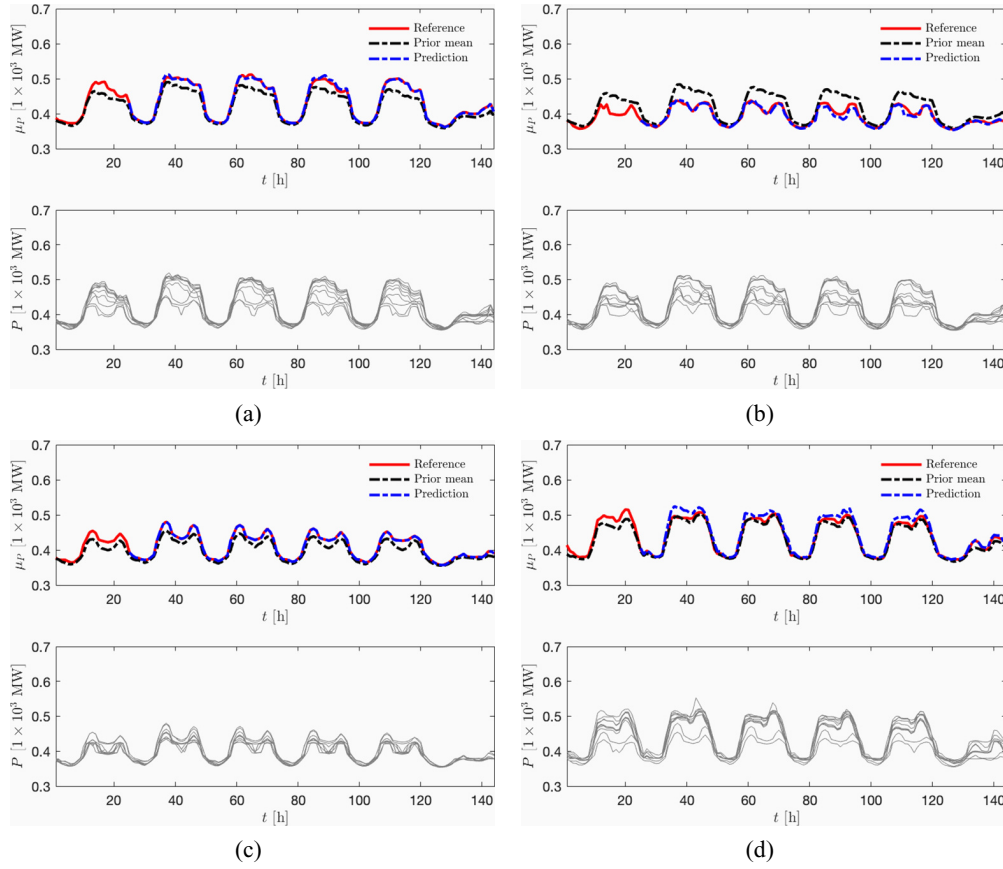
**FIG. 7:** Weekly total load forecasting using ARIMA (a), (b) and TBATS (c), (d) for the week of 06/16–06/22, using 24 hourly observations on Tuesday (a), (c) and Monday (b), (d)

### 3.3 Weekly Forecasting of Generation Scheduling

Finally, we use the EGPR method for weekly forecasting of the optimal power output of Generator 15 for the weeks of 6/16–6/22, 8/18–8/24, 10/20–10/26, and 12/22–12/28 using the power output observation data from up to 20 previous weeks. Figures 8 and 9 show the comparison of the forecast and the reference values for five (Wednesday to Sunday) and six (Tuesday to Sunday) days, respectively. As in the case with the total load forecast, we see that Monday data are weakly correlated with the rest of the data. As a result, the optimal power output forecast is more accurate for Wednesday–Sunday using the previous Tuesday data than for the Tuesday–Sunday using the previous Monday data, in accordance with Figs. 3 and 4. An accurate Wednesday–Sunday forecast can be obtained using ten previous weeks’ observations for computing the prior statistics, while 20 previous weeks observations are required for an accurate Tuesday to Sunday forecast. It can also be seen that Tuesday to Sunday forecasts for the weeks of 6/16–6/22 and 10/20–10/26 suffer from collapse to the prior mean, indicating poor forecasting performance. The poorer performance of Tuesday–Sunday forecasts compared to Wednesday–Sunday forecasts highlights the importance of analyzing the historical data in order to properly construct the ensembles for EGPR.

## 4. WEEKLY FORECASTING OF REAL LOAD DATA

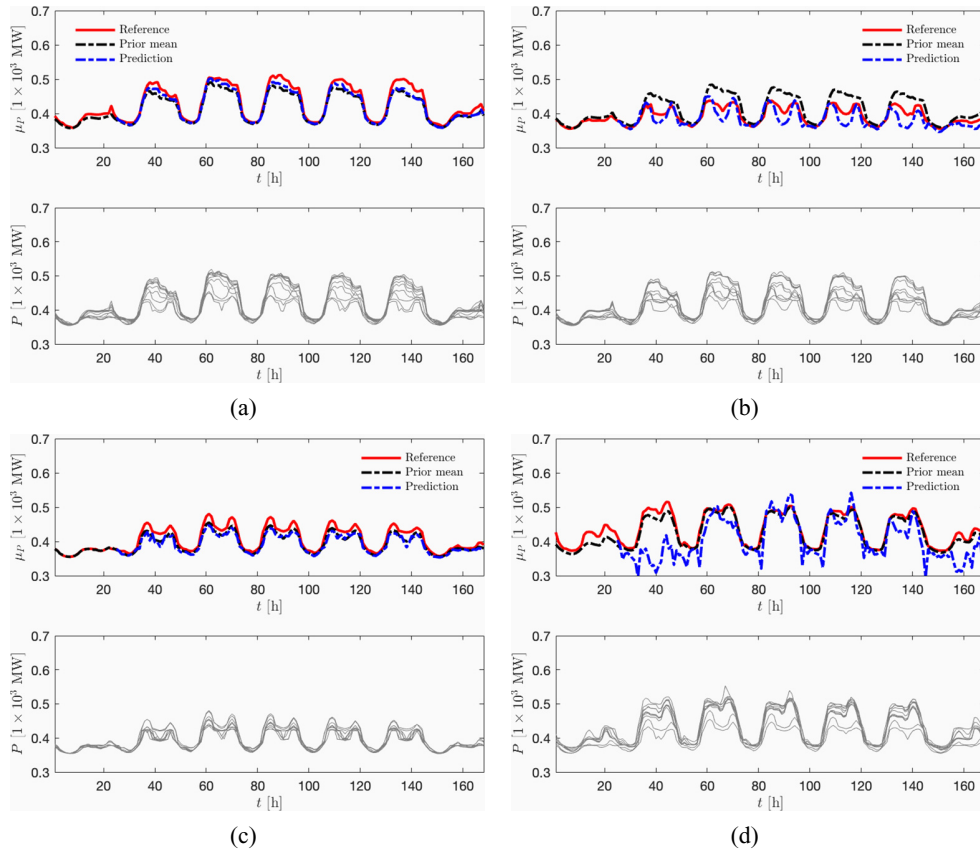
In this section we apply the proposed EGPR method to a real total system load data set from the Duke Energy system. The data consist of total system load measurements taken at an hourly rate over a 365-day period of time. Specifically, we forecast the total load demand for the weeks of 07/31–08/06, 09/18–09/24, 10/23–10/29, and 11/27–12/03 using EGPR, standard data-driven GPR, ARIMA, and TBATS, respectively.



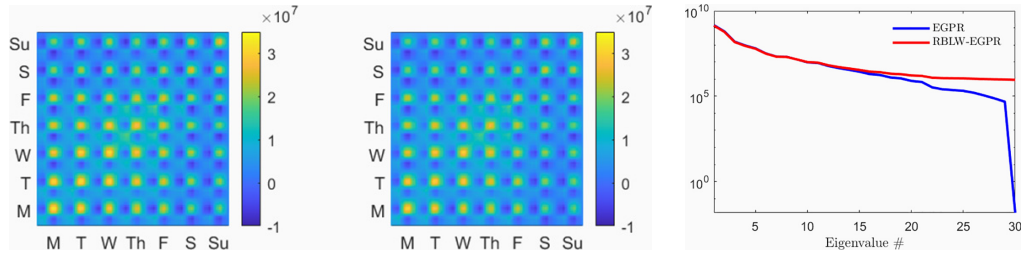
**FIG. 8:** Weekly forecasts of generation scheduling for 4 weeks using 24 hourly observations on Tuesday. Top panel: prediction (blue) compared against reference (red) and the prior mean (black). Bottom panel: ensemble of ten time series used to compute the empirical covariance. (a) 6/16–6/22, (b) 8/18–8/24, (c) 10/20–10/26, and (d) 12/22–12/28.

Figure 10 shows the ensemble load covariance for the week of 07/31–08/06 computed via EGPR and RBLW-EGPR using  $N = 30$ . It can be seen that the real load data exhibit correlations through both weekdays and weekends, a different correlation pattern than that exhibited by the synthetic data set considered in Section 3. Figure 10 also shows 30 eigenvalues of these ensemble covariances. We find that the 2-norm condition number for these two covariance matrices are  $9.6226 \times 10^{10}$  and  $1.4919 \times 10^3$ , respectively. This indicates that the covariance matrix obtained via RBLW-EGPR is better conditioned than the covariance matrix obtained via EGPR.

In the EGPR methods, the ensemble is constructed using hourly measurements of total load for second to last to the  $N + 1$ th last week of the data set. We aim to forecast the last week of data (which we use as reference) by using the previous 24 hr as observations. To study the effect of the choice of  $N$  on the predictive accuracy of the EGPR methods, we show in Fig. 11 log predictive probability as a function of realization numbers for EGPR, LOOCV-EGPR, RBLW-EGPR, and OAS-EGPR, respectively. It can be seen that for all EGPR methods and all the forecast weeks considered the log predictive probability increases with the ensemble size up to a plateau, before

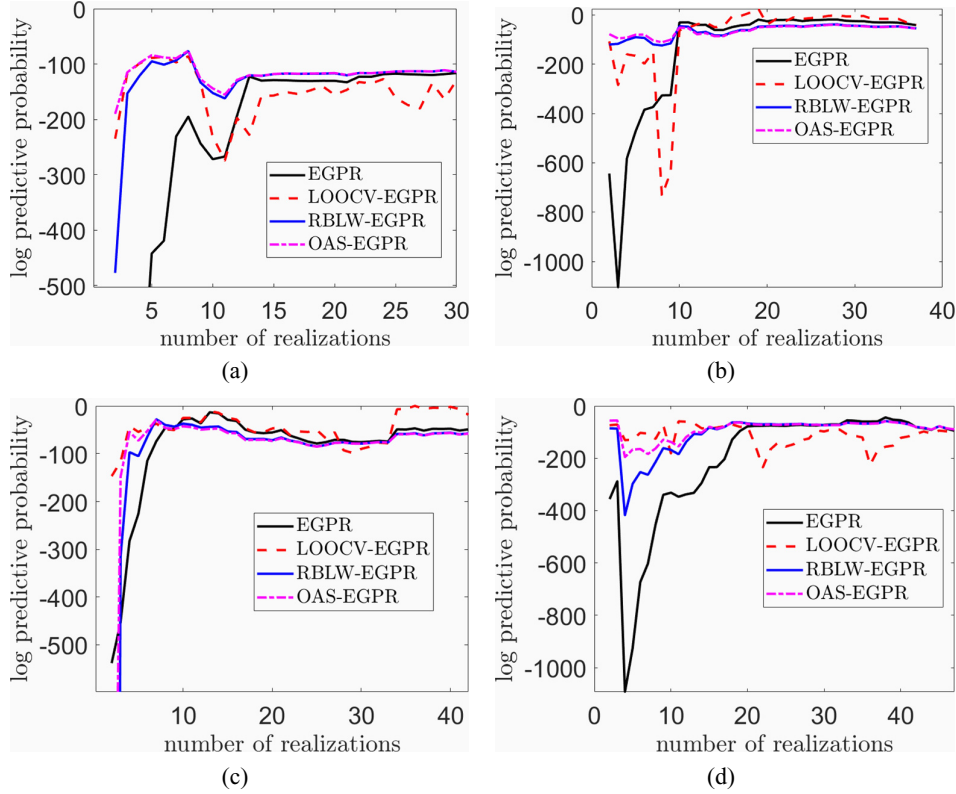


**FIG. 9:** Weekly forecasts of generation scheduling for 4 weeks using 24 hourly observations on Monday. Top panel: prediction (blue) compared against reference (red) and the prior mean (black). Bottom panel: ensemble of 20 time series used to compute the empirical covariance. (a) 6/16–6/22, (b) 8/18–8/24, (c) 10/20–10/26, and (d) 12/22–12/28.



**FIG. 10:** Ensemble covariance for the week of 07/31–08/06 computed using 30 realizations (left: EGPR, middle: RBLW-EGPR). Eigenspectrum of ensemble covariance for the week of 07/31–08/06 computed by EGPR and RBLW-EGPR using 30 realizations (right).

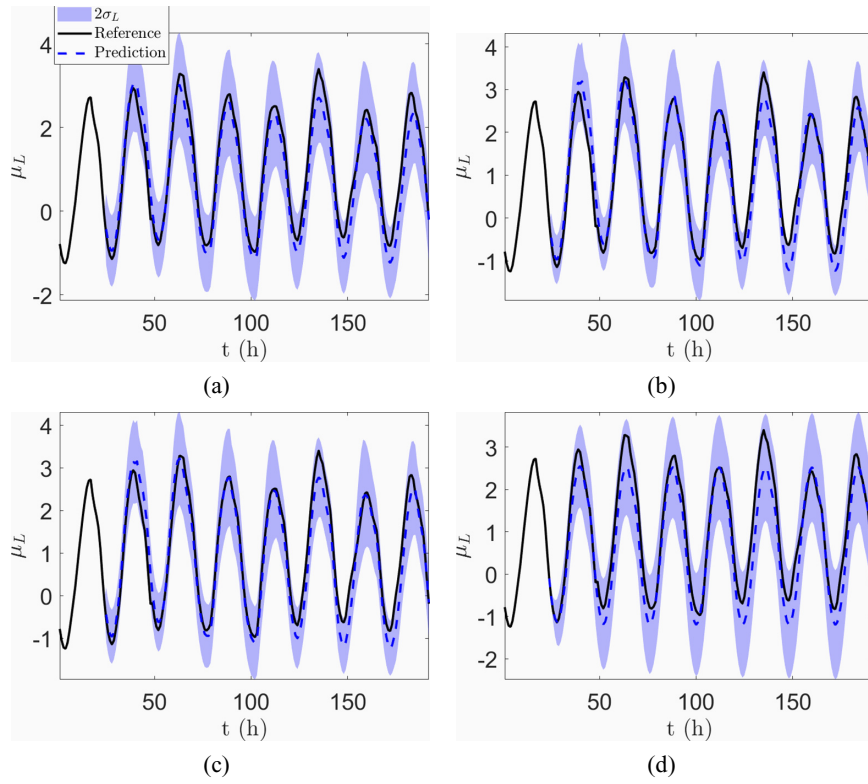
slightly decreasing. Based on this analysis, we choose  $N$  for EGPR, LOOCV-EGPR, RBLW-EGPR, and OAS-EGPR to be (30, 10, 13, 38), (8, 19, 36, 11), (8, 10, 7, 38), and (8, 10, 7, 38) for the weeks of 07/31–08/06, 09/18–09/24, 10/23–10/29, and 11/27–12/03, respectively.



**FIG. 11:** Log predictive probability as a function of realization numbers by using the EGPR methods. (a) 07/31–08/06, (b) 09/18–09/24, (c) 10/23–10/29, and (d) 11/27–12/03.

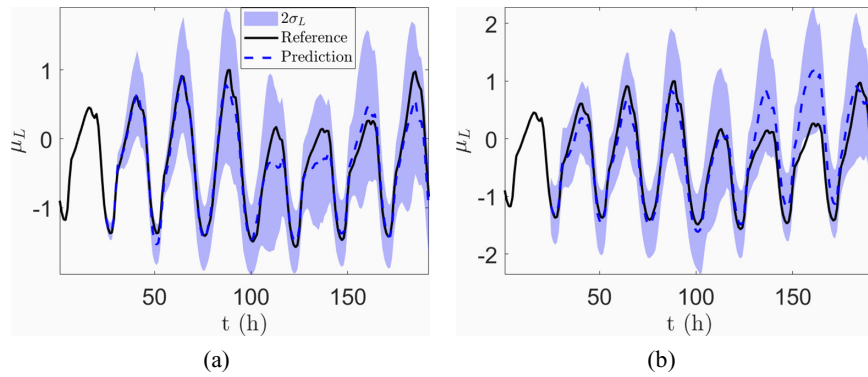
Figures 12–15 show the load forecasting for the weeks of 07/31–08/06, 09/18–09/24, 10/23–10/29, and 11/27–12/03 using LOOCV-EGPR, RBLW-EGPR, OAS-EGPR, and standard data-driven GPR, respectively. The standard data-driven GPR forecast is computed using the covariance kernel (8) and the second to last week of hourly measurements as observations. It can be seen that the regularized EGPR methods exhibit comparable performances among themselves, and in general perform better than standard data-driven GPR. We also note that standard data-driven GPR yields better performance for the weeks of 07/31–08/06 and 09/18–09/24 than for the weeks of 10/23–10/29 and 11/27–12/03; this is due to the periodic term in Eq. (8), which induces strong periodicity in the forecast.

Log predictive probabilities of the EGPR methods and standard data-driven GPR for the first day, first two days, the first five days, and the whole week of load forecasts during the weeks of 07/31–08/06, 09/18–09/24, 10/23–10/29, and 11/27–12/03 are presented in Table 2. It can be seen that RBLW-EGPR and OAS-EGPR demonstrate comparable performances. For the first day load forecasting, standard data-driven GPR has the largest log predictive probabilities for 07/31, 09/18 and 11/27, while LOOCV-EGPR has the largest log predictive probabilities for 10/23. This indicates that for very short forecast windows standard data-driven GPR performs better than the proposed EGPR method. Nevertheless, as the length of the forecast window increases, we find that the EGPR methods perform better than standard data-driven GPR.

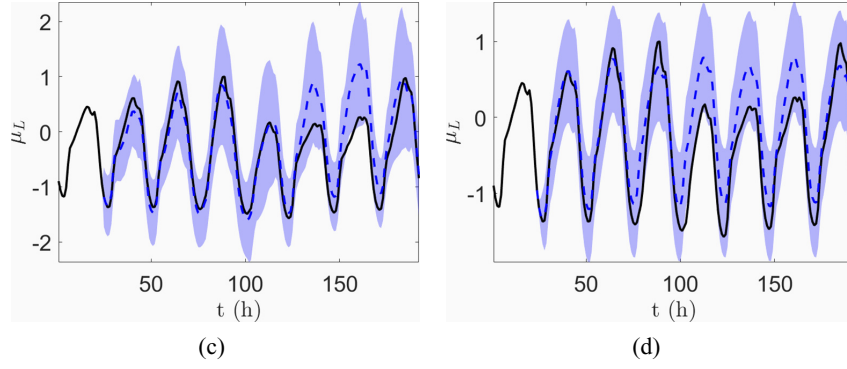


**FIG. 12:** Weekly forecasts of total load for 07/31–08/06 using 24 hr of observations with different EGPRs and data-driven GPR. Prediction (blue) compared against reference (black). (a) LOOCV-EGPR, (b) RBLW-EGPR, (c) OAS-EGPR, and (d) Data-driven GPR.

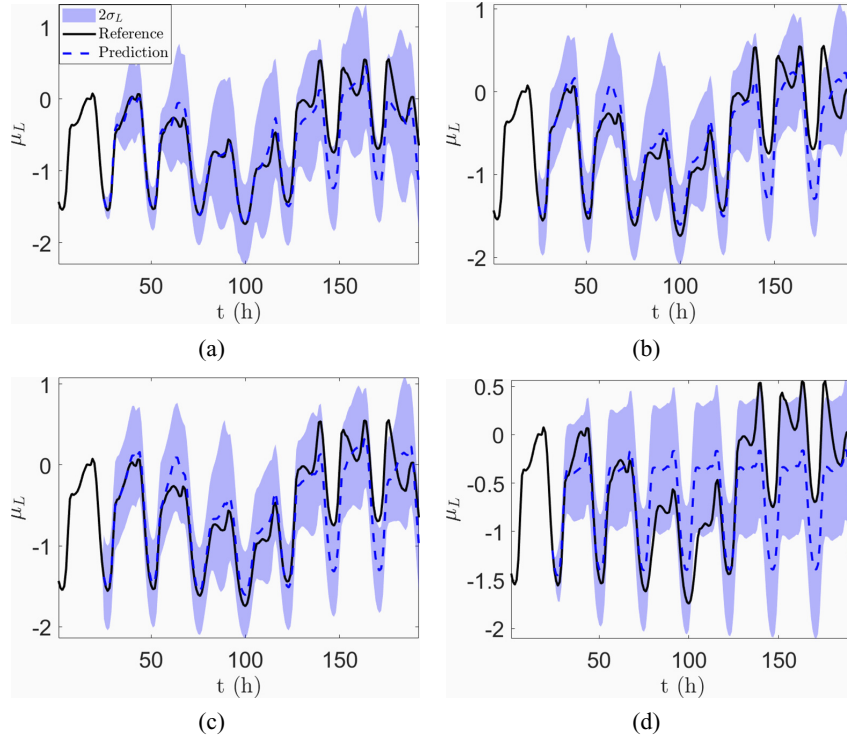
In addition to the log predictive probability, we also compute the mean absolute percentage error (MAPE) in order to compare the performance of the GPR-based forecasting methods. For



**FIG. 13.**



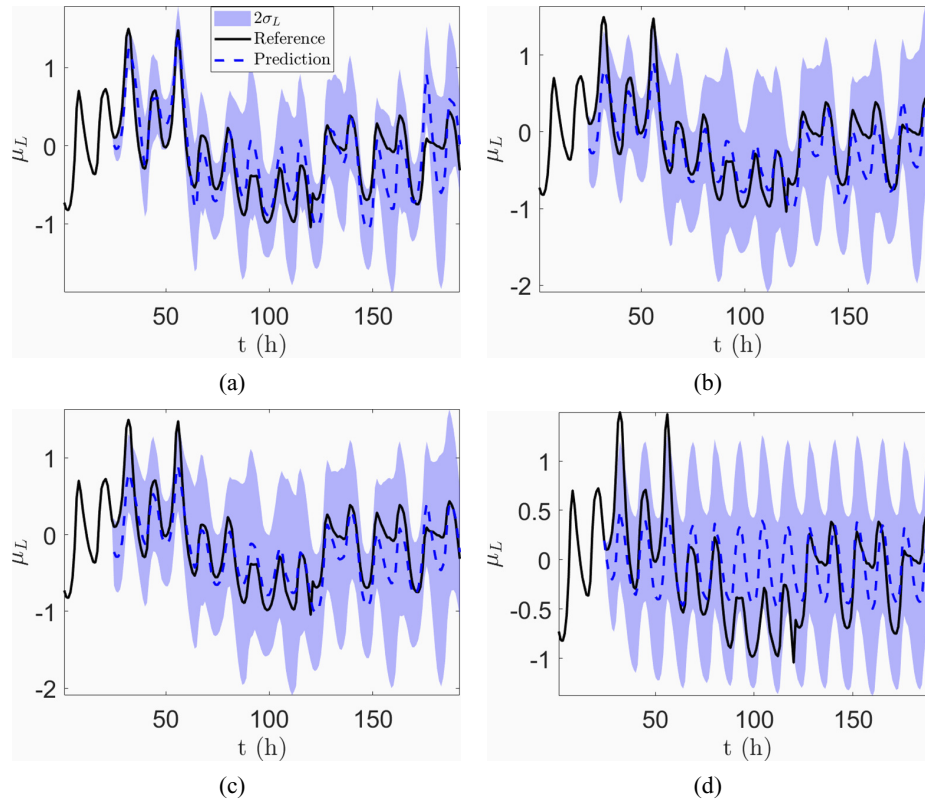
**FIG. 13:** Weekly forecasts of total load for 09/18–09/24 using 24 hr of observations with different EGPRs and data-driven GPR. Prediction (blue) compared against reference (black). (a) LOOCV-EGPR, (b) RBLW-EGPR, (c) OAS-EGPR, and (d) Data-driven GPR.



**FIG. 14:** Weekly forecasts of total load for 10/23–10/29 using 24 hr observations with different EGPRs and data-driven GPR. Prediction (blue) compared against reference (black). (a) LOOCV-EGPR, (b) RBLW-EGPR, (c) OAS-EGPR, and (d) Data-driven GPR.

a certain forecasted state  $\alpha(t)$ , MAPE is given by

$$\text{MAPE} = \frac{1}{N^f} \sum_{k=1}^{N^f} \left| \frac{\alpha(t_k) - \hat{\alpha}(t_k)}{\alpha(t_k)} \right|, \quad (9)$$



**FIG. 15:** Weekly forecasts of total load for 11/27–12/03 using 24 hr observations with different EGPRs and data-driven GPR. Prediction (blue) compared against reference (black). (a) LOOCV-EGPR, (b) RBLW-EGPR, (c) OAS-EGPR, and (d) Data-driven GPR.

where  $N_f$  denotes the number of forecast times,  $\hat{\alpha}(t_k)$  is the forecast at time  $t_k$ , and  $\alpha(t_k)$  is the reference value at time  $t_k$ . The lower the MAPE, the more accurate is the forecast. We present the MAPEs of standard data-driven GPR and the EGPR methods in Table 3. It can be seen that RBLW-EGPR has the lowest MAPEs for the weeks of 07/31–08/06, LOOCV-EGPR has the lowest MAPE for the week of 09/18–09/24 and 10/23–10/29, and EGPR has the lowest MAPEs for the weeks of 11/27–12/03. These results, together with the results of Table 2, indicate that EGPR (specially the regularized EGPR methods) outperform standard data-driven GPR for forecasting tasks. Of the regularized EGPR methods, LOOCV-EGPR is either comparable or better than RBLW-EGPR and OAS-EGPR, at the cost of having to estimate the LOOCV shrinkage coefficient via marginal likelihood maximization; on the other hand, as noted in Section 2.2, the RBLW and OAS shrinkage coefficients are given in closed form and do not require solving additional minimization problems.

Finally, we also compare the EGPR methods to demonstrate against the ARIMA and TBATS methods. The ARIMA and TBATS forecasts for the 4 weeks under consideration are shown in Fig. 16. The ARIMA models are ARIMA (2, 0, 0), ARIMA (2, 0, 3), ARIMA (4, 0, 0) and ARIMA (4, 1, 1), respectively. By comparing Fig. 16 against Figs. 12–15 it can be seen that the EGPR method outperforms both ARIMA and TBATS.

**TABLE 2:** Log predictive probabilities for total load forecasts of different lengths using different GPR methods. Largest values are indicated in italics

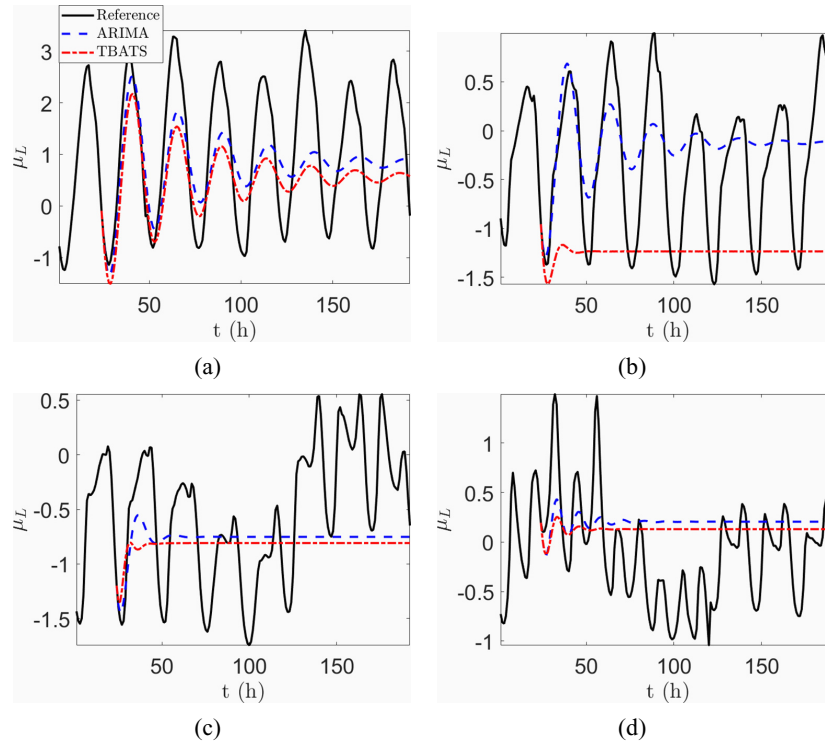
| Forecast window | Method          | 07/31         | 09/18          | 10/23          | 11/27         |
|-----------------|-----------------|---------------|----------------|----------------|---------------|
| First day       | EGPR            | -8.28989      | 4.4535         | 16.6451        | -15.8943      |
|                 | LOOCV-EGPR      | -9.9158       | 1.0484         | <i>19.5029</i> | -27.6376      |
|                 | RBLW-EGPR       | -12.0248      | 3.6806         | 10.8394        | -19.9724      |
|                 | OAS-EGPR        | -10.8668      | 2.8115         | 8.2497         | -20.1926      |
|                 | Data-driven GPR | <i>0.0272</i> | <i>11.9359</i> | 10.2287        | -13.3667      |
| First 2 days    | EGPR            | -13.8783      | 4.45345        | 24.0393        | -16.6298      |
|                 | LOOCV-EGPR      | -19.4996      | -7.0471        | <i>30.7093</i> | -39.7709      |
|                 | RBLW-EGPR       | -15.8853      | <i>5.4647</i>  | 15.328         | -25.1011      |
|                 | OAS-EGPR        | -15.8947      | 3.9194         | 11.1685        | -25.7069      |
|                 | Data-driven GPR | -20.518       | 13.6451        | 15.3259        | -29.051       |
| First 5 days    | EGPR            | -77.5001      | 0.1726         | 28.8518        | -29.1826      |
|                 | LOOCV-EGPR      | -54.5577      | <i>26.8323</i> | <i>29.0811</i> | -48.5338      |
|                 | RBLW-EGPR       | -40.5786      | -7.9555        | 24.3287        | -42.1176      |
|                 | OAS-EGPR        | -42.5929      | -12.1781       | 15.7516        | -43.1073      |
|                 | Data-driven GPR | -71.9269      | -40.4197       | -41.0583       | -74.4628      |
| Full week       | EGPR            | -116.135      | -31.0167       | -12.9992       | -44.1647      |
|                 | LOOCV-EGPR      | -85.9404      | 24.4772        | 0.363776       | -58.9145      |
|                 | RBLW-EGPR       | -76.2976      | -46.0508       | -27.4528       | -57.9922      |
|                 | OAS-EGPR        | -75.835       | -51.8968       | -28.0166       | -59.0892      |
|                 | Data-driven GPR | -58.5933      | <i>28.099</i>  | <i>42.4793</i> | <i>2.3705</i> |

**TABLE 3:** Mean absolute percentage error (MAPE) for the weekly load forecasts using different GPR methods. Smallest errors are indicated in italics

| Method          | 07/31–08/06   | 09/18–09/24   | 10/23–10/29   | 11/27–12/03   |
|-----------------|---------------|---------------|---------------|---------------|
| EGPR            | 0.0762        | 0.0545        | 0.0440        | <i>0.0409</i> |
| LOOCV-EGPR      | 0.0530        | <i>0.0279</i> | <i>0.0347</i> | 0.0568        |
| RBLW-EGPR       | <i>0.0470</i> | 0.0548        | 0.0449        | 0.0452        |
| OAS-EGPR        | 0.0483        | 0.0555        | 0.0457        | 0.0456        |
| Data-driven GPR | 0.0628        | 0.0646        | 0.0855        | 0.0718        |

## 5. DISCUSSION AND CONCLUSIONS

We have proposed the ensemble Gaussian process regression (EGPR) method for the forecasting of time series with a periodic structure. The proposed EGPR method is based on GPR, but instead of employing parameterized models for the prior mean and covariance, these prior statistics are computed via sample averaging from an ensemble constructed by periodic windowing of the time series. We have employed EGPR for the weekly forecasting of total load demand and generation scheduling (optimal generator output) for a power grid with 700 buses and 134 generators using synthetic historical data. We have also applied EGPR for the weekly forecast of total load demand using real historic data. These numerical experiments demonstrate that EGPR



**FIG. 16:** Weekly forecasts of total load for four weeks using ARIMA and TBATS. (a) 07/31–08/06, (b) 09/18–09/24, (c) 10/23–10/29, and (d) 11/27–12/03.

provides accurate forecasts and in general outperforms the standard data-driven GPR, ARIMA, and TBATS forecasting methods. Furthermore, we employ shrinkage estimation to overcome the rank-deficiency problem of estimating covariance matrices with small ensembles. To estimate the shrinkage coefficient we employ leave-one-out cross-validation (LOOCV), and both the Rao-Blackwell Ledoit-Wolf (RBLW) and Oracle approximating shrinkage (OAS) estimates.

## ACKNOWLEDGMENTS

This work was partially supported by the U.S. Department of Energy (DOE) Office of Science, Office of Advanced Scientific Computing Research (ASCR) as part of the Multifaceted Mathematics for Rare, Extreme Events in Complex Energy and Environment Systems (MACSER) project. A.M. Tartakovsky was partially supported by the NSF Grant OIA-2033607. Pacific Northwest National Laboratory is operated by Battelle for the DOE under Contract DE-AC05-76RL01830.

## REFERENCES

- Amjady, N., Short-Term Hourly Load Forecasting Using Time-Series Modeling with Peak Load Estimation Capability, *IEEE Transact. Power Sys.*, vol. **16**, no. 3, pp. 498–505, 2001.
- Bianchi, F.M., Maiorino, E., Kampffmeyer, M.C., Rizzi, A., and Jenssen, R., *Recurrent Neural Networks*

- for *Short-Term Load Forecasting: An Overview and Comparative Analysis*, Berlin, Germany: Springer, 2017.
- Borojjeni, K.G., Amini, M.H., Bahrami, S., Iyengar, S., Sarwat, A.I., and Karabasoglu, O., A Novel Multi-Time-Scale Modeling for Electric Power Demand Forecasting: From Short-Term to Medium-Term Horizon, *Electric Power Sys. Res.*, vol. **142**, pp. 58–73, 2017.
- Borojjeni, K.G., Mokhtari, S., Amini, M.H., and Iyengar, S., Optimal Two-Tier Forecasting Power Generation Model in Smart Grids, *Int. J. Inform. Proc.*, vol. **8**, no. 4, pp. 79–88, 2014.
- Brockwell, P.J. and Davis, R.A., *Introduction to Time Series and Forecasting*, Berlin, Germany: Springer, 2016.
- Charlton, N. and Singleton, C., A Refined Parametric Model for Short Term Load Forecasting, *Int. J. Forecast.*, vol. **30**, no. 2, pp. 364–368, 2014.
- Chen, B.-J., Chang, M.-W., and Lin, C.-J., Load Forecasting Using Support Vector Machines: A Study on Eunit Competition 2001, *IEEE Transact. Power Systems*, vol. **19**, no. 4, pp. 1821–1830, 2004.
- Chen, Y., Wiesel, A., and Hero, A.O., Shrinkage Estimation of High Dimensional Covariance Matrices, *2009 IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Taipei, Taiwan, pp. 2937–2940, 2009.
- Chen, Y., Xu, P., Chu, Y., Li, W., Wu, Y., Ni, L., Bao, Y., and Wang, K., Short-Term Electrical Load Forecasting Using the Support Vector Regression (SVR) Model to Calculate the Demand Response Baseline for Office Buildings, *Appl. Energy*, vol. **195**, pp. 659–670, 2017.
- Din, G.M.U. and Marnerides, A.K., Short Term Power Load Forecasting Using Deep Neural Networks, *2017 Int. Conf. on Computing, Networking and Communications (ICNC)*, pp. 594–598, Silicon Valley, NC, January 26–29, 2017.
- Duke Energy Ohio Inc., Hourly Load and Historical Customer Switching, accessed November 11, 2021, from <https://www.duke-energyohiocbp.com/Documents/LoadandOtherData.aspx>, 2021.
- Fan, S. and Hyndman, R.J., Short-Term Load Forecasting Based on a Semi-Parametric Additive Model, *IEEE Transact. Power Sys.*, vol. **27**, no. 1, pp. 134–141, 2011.
- Goude, Y., Nedellec, R., and Kong, N., Local Short and Middle Term Electricity Load Forecasting with Semi-Parametric Additive Models, *IEEE Transact. Smart Grid*, vol. **5**, no. 1, pp. 440–446, 2013.
- Grigg, C., Wong, P., Albrecht, P., Allan, R., Bhavaraju, M., Billinton, R., Chen, Q., Fong, C., Haddad, S., and Kuruganty, S., The IEEE Reliability Test System-1996. A Report Prepared by the Reliability Test System Task Force of the Application of Probability Methods Subcommittee, *IEEE Transact. Power Sys.*, vol. **14**, no. 3, pp. 1010–1020, 1999.
- Grunblatt, S.K., Howard, A.W., and Haywood, R.D., Determining the Mass of Kepler-78b with Nonparametric Gaussian Process Estimation, *Astrophys. J.*, vol. **808**, no. 2, p. 127, 2015. DOI: 10.1088/0004-637x/808/2/127
- Hong, T. and Fan, S., Probabilistic Electric Load Forecasting: A Tutorial Review, *Int. J. Forecast.*, vol. **32**, no. 3, pp. 914–938, 2016.
- Hong, T. and Wang, P., Fuzzy Interaction Regression for Short Term Load Forecasting, *Fuzzy Optim. Decis. Making*, vol. **13**, no. 1, pp. 91–103, 2014.
- Hong, T., Wilson, J., and Xie, J., Long Term Probabilistic Load Forecasting and Normalization with Hourly Information, *IEEE Transact. Smart Grid*, vol. **5**, no. 1, pp. 456–462, 2013.
- Hong, T., Short Term Electric Load Forecasting, PhD, North Carolina State University, Raleigh, NC, USA, 2010.
- Hyndman, R.J. and Athanasopoulos, G., *Forecasting: Principles and Practice*, OTexts, 2018.
- Hyndman, R.J. and Fan, S., Density Forecasting for Long-Term Peak Electricity Demand, *IEEE Transact. Power Sys.*, vol. **25**, no. 2, pp. 1142–1153, 2009.
- Klenske, E.D., Zeilinger, M.N., Schölkopf, B., and Hennig, P., Gaussian Process-Based Predictive Control

- for Periodic Error Correction, *IEEE Transact. Control Sys. Technol.*, vol. **24**, no. 1, pp. 110–121, 2016.
- Livera, A.M.D., Hyndman, R.J., and Snyder, R.D., Forecasting Time Series with Complex Seasonal Patterns Using Exponential Smoothing, *J. Am. Stat. Assoc.*, vol. **106**, no. 496, pp. 1513–1527, 2011. DOI: 10.1198/jasa.2011.tm09771
- Lloyd, J.R., Gefcom2012 Hierarchical Load Forecasting: Gradient Boosting Machines and Gaussian Processes, *Int. J. Forecast.*, vol. **30**, no. 2, pp. 369–374, 2014.
- Nedellec, R., Cugliari, J., and Goude, Y., Gefcom2012: Electric Load Forecasting and Backcasting with Semi-Parametric Models, *Int. J. Forecast.*, vol. **30**, no. 2, pp. 375–381, 2014.
- Qingle, P. and Min, Z., Very Short-Term Load Forecasting Based on Neural Network and Rough Set, *2010 Int. Conf. on Intelligent Computation Technology and Automation*, Vol. 3, Changsha, China, pp. 1132–1135, 2010.
- Ryu, S., Noh, J., and Kim, H., Deep Neural Network Based Demand Side Short Term Load Forecasting, *Energy*, vol. **10**, no. 1, p. 3, 2017.
- Shepero, M., van der Meer, D., Munkhammar, J., and Widén, J., Residential Probabilistic Load Forecasting: A Method Using Gaussian Process Designed for Electric Load Data, *Appl. Energy*, vol. **218**, pp. 159–172, 2018.
- Shi, H., Xu, M., and Li, R., Deep Learning for Household Load Forecasting—A Novel Pooling Deep RNN, *IEEE Transact. Smart Grid*, vol. **9**, no. 5, pp. 5271–5280, 2017.
- Song, K.B., Baek, Y.S., Hong, D.H., and Jang, G., Short-Term Load Forecasting for the Holidays Using Fuzzy Linear Regression Method, *IEEE Transact. Power Sys.*, vol. **20**, no. 1, pp. 96–101, 2005.
- Steland, A., Shrinkage for Covariance Estimation: Asymptotics, Confidence Intervals, Bounds and Applications in Sensor Monitoring and Finance, *Stat. Papers*, vol. **59**, no. 4, pp. 1441–1462, 2018.
- Taieb, S.B. and Hyndman, R.J., A Gradient Boosting Approach to the Kaggle Load Forecasting Competition, *Int. J. Forecast.*, vol. **30**, no. 2, pp. 382–394, 2014.
- Tartakovsky, A. and Tipireddy, R., Physics-Informed Machine Learning Method for Forecasting and Uncertainty Quantification of Partially Observed and Unobserved States in Power Grids, *Proc. of the 52nd Hawaii Int. Conf. on System Sciences*, Grand Wailea, Maui, HI, USA, 2019.
- Taylor, J.W., An Evaluation of Methods for Very Short-Term Load Forecasting Using Minute-by-Minute British Data, *Int. J. Forecast.*, vol. **24**, no. 4, pp. 645–658, 2008.
- Taylor, J.W. and McSharry, P.E., Short-Term Load Forecasting Methods: An Evaluation Based on European Data, *IEEE Transact. Power Sys.*, vol. **22**, no. 4, pp. 2213–2219, 2007.
- Tolba, H., Dkhili, N., Nou, J., Eynard, J., Thil, S., and Grieu, S., Ghi Forecasting Using Gaussian Process Regression: Kernel Study, *iFAC Workshop on Control of Smart Grid and Renewable Energy Systems*, Vol. 52, pp. 455–460, Jeju, South Korea, June 10–12, 2019.
- Wang, P., Liu, B., and Hong, T., Electric Load Forecasting with Recency Effect: A Big Data Approach, *Int. J. Forecast.*, vol. **32**, no. 3, pp. 585–597, 2016.
- Williams, C.K. and Rasmussen, C.E., *Gaussian Processes for Machine Learning*, Vol. 2, Cambridge, MA: MIT Press, 2006.
- Wilson, A. and Adams, R., Gaussian Process Kernels for Pattern Discovery and Extrapolation, *Proc. of the 30th Int. Conf. on Machine Learning*, S. Dasgupta and D. McAllester, Eds., Vol. 28, *Proc. of Machine Learning Research*, PMLR, Atlanta, Georgia, USA, pp. 1067–1075, 2013.
- Xie, J., Chen, Y., Hong, T., and Laing, T.D., Relative Humidity for Load Forecasting Models, *IEEE Transact. Smart Grid*, vol. **9**, no. 1, pp. 191–198, 2016.
- Xie, J. and Hong, T., Temperature Scenario Generation for Probabilistic Load Forecasting, *IEEE Transact. Smart Grid*, vol. **9**, no. 3, pp. 1680–1687, 2016.
- Xie, J. and Hong, T., Variable Selection Methods for Probabilistic Load Forecasting: Empirical Evidence

- from Seven States of the United States, *IEEE Transact. Smart Grid*, vol. **9**, no. 6, pp. 6039–6046, 2017.
- Yang, X., Barajas-Solano, D., Tartakovsky, G., and Tartakovsky, A.M., Physics-Informed Cokriging: A Gaussian-Process-Regression-Based Multifidelity Method for Data-Model Convergence, *J. Comput. Phys.*, vol. **395**, pp. 410–431, 2019.
- Yang, Y., Li, S., Li, W., and Qu, M., Power Load Probability Density Forecasting Using Gaussian Process Quantile Regression, *Appl. Energy*, vol. **213**, pp. 499–509, 2018b.
- Young, S.J., Makarov, Y., Diao, R., Fan, R., Huang, R., O'Brien, J., Halappanavar, M., Vallem, M., and Huang, Z.H., Synthetic Power Grids from Real World Models, *2018 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5, August 5–10, 2018.
- Yun, Z., Quan, Z., Caixin, S., Shaolan, L., Yuming, L., and Yang, S., RBF Neural Network and Anfis-Based Short-Term Load Forecasting Approach in Real-Time Price Environment, *IEEE Transact. Power Sys.*, vol. **23**, no. 3, pp. 853–858, 2008.
- Zheng, H., Yuan, J., and Chen, L., Short-Term Load Forecasting Using EMD-LSTM Neural Networks with a Xgboost Algorithm for Feature Importance Evaluation, *Energ.*, vol. **10**, no. 8, p. 1168, 2017.