IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received January 5, 2022, accepted January 14, 2022, date of publication January 18, 2022, date of current version January 26, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3144206

Regression Model-Based Short-Term Load Forecasting for University Campus Load

MITHUN MADHUKUMAR[1], ALBINO SEBASTIAN[1], XIAODONG LIANG[2], (Senior Member, IEEE), MOHSIN JAMIL[1], (Senior Member, IEEE), AND MD NASMUS SAKIB KHAN SHABBIR[2], (Student Member, IEEE)

[1]Department of Electrical and Computer Engineering, Memorial University of Newfoundland, St. John's, NL A1B 3X5, Canada
[2]Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada

Corresponding author: Xiaodong Liang (xil659@mail.usask.ca)

This work was supported in part by the Natural Science and Engineering Research Council of Canada (NSERC) Discovery Grant RGPIN-2016-04170.

ABSTRACT Load forecasting is a critical aspect for power systems planning, operation and control. In this paper, as part of research efforts of an ambitious project at Memorial University of Newfoundland in St. John's, Canada, to achieve more energy efficient and environmental friendly ''Sustainable Campus'', we present a day-ahead load forecasting approach for the energy management system of the project. The hourly load consumption dataset from January 1, 2016 to March 31, 2020 is used in the paper, which was collected from two power meters on campus. Using the load consumption dataset along with the collected meteorological dataset, a total of 19 regression model-based day-ahead load forecasting algorithms for Memorial University of Newfoundland's campus load are developed and evaluated in this paper. These 19 models belong to five families of regression models in MATLAB Regression Toolbox: Linear Regression, Regression Trees, Support Vector Machines (SVM), Gaussian Process Regression (GPR), and Ensemble of Trees. It is found that the family of GPR models shows the best load forecasting performance because they are nonparametric kernel-based probabilistic models. Two GPR models, Rational Quadratic GPR and Exponential GPR, are recommended as the best models for load forecasting through this study.

INDEX TERMS Short-term load forecasting, regression model, day-ahead load forecasting, Gaussian process regression, probabilistic models, university campus load.

I. INTRODUCTION

Load forecasting is essential to maintain the balance of power supply and demand in power grids, and serves as the foundation of power market operation. Power systems planning and operation rely on accurate load forecasting on various time horizons [1]. Load forecasting accuracy has significant impact on power grids' stability and operating cost, for example, a 1% reduction in the load forecasting error reduced 10 million pounds operating cost per year for one utility company in the United Kingdom [2].

The principle of load forecasting was introduced in 1894 by Samuel Insull, an innovator in the electric utility industry. He analyzed that different load use trends, such as domestic and commercial end-users, concluded that the maximum consumption was observed in the day time for domestic consumption while industries have maximum consumption at night [3], [4]. Load forecasting enables the electric utility to make efficient unit commitment decisions, minimize the spinning reserve, and properly schedule the maintenance plan; and optimize the transmission network's power flow to minimize underloads and overloads. Accurate load forecasting can lead to substantial savings in operational and maintenance costs, enhanced power supply efficiency, proper transmission and distribution future planning decisions.

Grid modernization, such as renewable energy sources integration, demand response, and electric vehicles, has introduced increased uncertainty due to electricity demand becoming more active and less predictable, and load forecasting is more challenging than ever before. Tomorrow's smart distribution systems feature distributed renewable energy generation and demand response control, can operate independently from the bulk power system as microgrids, where the load may contain more stochastically abrupt deviations due to much larger impact from the behavior of end users. Successfully operating such a system requires much more

The associate editor coordinating the review of this manuscript and approving it for publication was Lei Wu.

accurate and high-resolution load forecasting than today's existing techniques [5]. With the implementation of advanced data acquisition systems (such as smart meters and advanced metering infrastructure (AMI) meters), and the advancement of data analytic techniques and artificial intelligence, load forecasting has attracted significant research interests from academia and industry.

Based on the forecasting time horizon, load forecasting can be categorized into short-term, medium-term, and long-term load forecasting. The short-term load forecasting deals with the forecasting of several minutes up to one week into the future; while medium-term and long-term load forecasting are two weeks and three years, respectively [2], [6]. Among the three types of load forecasting, the short-term load forecasting is the most investigated subject in the literature due to its great importance in the operation planning, dealing with demand response, unit commitment, economic dispatch, and energy trading. Operation activities of power grids, such as regulation bids, energy arbitrage, and market-clearing mechanisms, are conducted on hourly bases, which rely on short-term forecasting techniques to provide accurate estimations of future load demands [7]. In case of a fault, fast restoration methods can be deployed to minimize the total amount of expected unsupplied demand through self-healing schemes, it requires fast and precise short-term load forecasting to realize a robust final restoration [7].

Knowledge-based expert system utilizes the knowledge of skilled human experts into a software for load forecasting. However, there is tremendous difficulty to transfer such knowledge to developing rules during load forecasting. The advancement of artificial intelligence techniques has brought the concept of soft computing-based load forecasting, which makes it an adaptable learning method for training the data set. The drawback of artificial intelligence techniques is that it sometimes cannot identify the mathematical expression between dependent and independent variables [4].

Whereas the regression analysis is an essential mathematical tool in determining the statistical relationship between various dependent and independent variables with the following advantages: 1) It gives the strength and direction of a relationship between variables; 2) Unlike simple correlations, it allows the use of more than one predictor, and allows the prediction of an outcome even when the multiple predictors are correlated with each other; 3) It can be used to correct errors based on previous assumptions; and 4) Excellent outcomes can be obtained with relatively small set of data. Therefore, regression algorithms are used in this paper for the short term load forecasting.

In the literature, load forecasting methods can be divided into three streams: 1) point (deterministic) load forecasting [1], [2], [5], 2) probabilistic load forecasting [6]–[9], and 3) hybrid methods by combining point and probabilistic load forecasting [10]. The brief summary of the three streams are shown in Fig. 1. The point load forecasting is most researched for decades by forecasting the expected value of future load using various statistic and machine learning techniques, while the probabilistic load forecasting only attracts research attention recently. In smart grid environment, the probabilistic load forecasting is essential as variability and uncertainty associated with the electricity demand increase, the demand at individual household or distribution feeder level can be quite volatile due to demand response programs and feeder reconfiguration activities. Probabilistic load forecasting provides electric load forecasting output in the form of intervals, scenarios, density functions, or probabilities; its typical applications include stochastic unit commitment, probabilistic price forecasting, and probabilistic transmission planning [8].

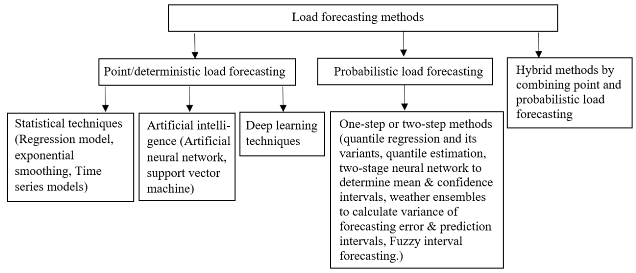


FIGURE 1. **Short-term load forecasting methods in the literature.**

Regression-based load forecasting models have been developed to provide load forecasting for the entire country, a specific region, a university campus, and small loads, such as a microgrid connected to a community etc. The models can be developed for various time horizons (short-term, medium-term, and long-term), where a wide range of influential environmental and temporal parameters have been considered. Short-term load forecasting is crucial to the energy security and stability to the load due to its immediate impact [11]. However, it has been observed that the performance of the developed short-term load forecasting models for the university campus load degrades with a smaller load size [12]. Therefore, many robust regression algorithms have been utilized to improve the effectiveness of the developed data driven short-term load forecasting models. It is found that Artificial Neural Network (ANN) is the mostly adopted regressor to develop short-term load forecasting models for the university campus load [11]–[13]. Despite being a powerful regressor, many researchers are not convinced about the direct application of ANN, as it may not be adaptive to all datasets [14]. While compared, several other regressors appear to be more effective than ANN to develop short-term load forecasting models for the university campus load. For example, the Support Vector Regressor (SVR) performs better than ANN in [15]. The Multiple Linear Regression (MLR), ANN and SVR are compared in [16], and SVR shows the most prominent performance. The Autoregressive Integrated Moving Average with exogenous variables (ARMAX), MLR, ANN, SVR are studied in [17], and SVR appeared to be the most suitable one. ANN, SVR and Ensemble (Random Forest) are studied in [18], and ensemble performs better. In [19], Gaussian Process Regression (GPR) is chosen after comparing among SVR, RF and GPR. Ref [11] concludes that simple machine learning based regressors are more effective to develop short-term load forecasting models than compound regressors. Therefore, in this paper, to find the best short-term load forecasting algorithms for the university campus load profile and predictors, the effectiveness of the Linear Regression, Support Vector Regression, Ensemble Trees, and Gaussian Process Regression (GPR) has been analyzed and compared.

In this paper, we focus on short-term load forecasting for the university campus load at Memorial University of Newfoundland, St. John's, Canada. As part of research efforts of an ambitious project at Memorial University of Newfoundland to achieve more energy efficient and environmental friendly "Sustainable Campus", we have developed an energy management system involving newly designed solar power and energy storage. To achieve effective control of the energy management system, the short-term load forecasting is critical.

To find the best short-term load forecasting algorithms for the university campus load profile and predictors, we have investigated the accuracy of day-ahead short-term load forecasting using 19 regression models in five regression families, including Regression, Linear Regression, Support Vector Regression (SVR), Ensemble Trees, and Gaussian Process Regression (GPR). The first four families belong to point forecasting category; while the GPR models are nonparametric kernel-based probabilistic models [20]. Since support vector machine (SVM) methods are commonly used for load forecasting [21], [22], SVM models serve as forecasting benchmarks for short-term load forecasting in this paper. The historical load datasets from January 2016 to March 2020 recorded at the two meters installed on the university campus are used in this study along with historical meteorological dataset. The load forecasting accuracy is evaluated using the statistic error indices, Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE), to find the best fitted regression models. After initial selection among the 19 regression models, the six top performing models are further evaluated using various response and residual plots to find the recommended final regression models.

The main contributions of the paper include: 1) propose an effective load forecasting approach for the university campus load; and 2) demonstrate probabilistic forecasting models (Gaussian Process Regression models) are more accurate than other models.

The paper is organized as follows: In Section II, the background information of the university campus load is provided, the proposed load forecasting approach is introduced and explained; In Section III, a brief description of various Regression models is presented along with performance evaluation indices; Simulation results and data analysis for load forecasting are demonstrated using 19 regression models in Section IV; and Conclusions are drawn in Section V.

II. THE PROPOSED LOAD FORECASTING APPROACH FOR THE UNIVERSITY CAMPUS LOAD

In this section, we introduce the background information about the campus load of Memorial University of Newfoundland first; then the proposed load forecasting approach is presented.

A. ABOUT THE UNIVERSITY CAMPUS LOAD

The campus load at Memorial University of Newfoundland consumes a significant percentage of electricity demand in the province of Newfoundland and Labrador.

Heating is provided all year round from a high-temperature hot water plant for all structures on campus. During the typical months of a year, this plant will have two out of its four boilers up and running; while during the cold winter months, all of the four boilers will work. The campus consumes approximately 11 million liters of diesel for annual heating purposes of 60 buildings on the St. John's campus on average, which costed 75 cents per liter in 2019, or roughly \$8.5 million annually. Every year, this combustion releases more than 25,000 tons of carbon dioxide into the atmosphere, equivalent to almost 7,000 automobiles. Therefore, by converting the present system to electrical-based heating, the university can bring sustainable initiatives to the province, bring economic benefits and reduce harmful emission. Therefore, an ambitious project was conducted by our research team, aiming to make the campus more energy efficient and environmental friendly towards a future "Sustainable Campus". To achieve this goal, solar power and energy storage are designed, and an energy management system is developed. To realize this energy management system's proper operation, short-term load forecasting is required.

The load forecasting data were collected from two power meters (Meter 1 and Meter 2) on campus, which record reading every 15-minute. For our load forecasting model, hourly data were used.

In this paper, an analysis of the university campus load demand from January 1, 2016 to March 31, 2020 is conducted by focusing on monthly (Fig. 2), seasonal (Fig. 3) and annual patterns. The study results revealed a correlation between energy usage and temperature. The university campus has an average demand of 10 MW from the local utility at St John's, Newfoundland Power, and winter months usually consumes more. An unplanned outage to the campus' power system will be a severe issue as thousands of students and faculty use the campus. There are several emergency generators with a combined capacity of 3,500 kW, which provide emergency power, such as lighting on stairs and the operation of pumps for heating. Apart from this, four generators having an individual capacity of 800 kW can provide 3,200 kW emergency power to Health Science Centre.

As shown in Fig. 2, the load demand patterns of the university campus show an increase during the winter months, December, January, and February, due to space heating requirements; while show a decrease from May to October.

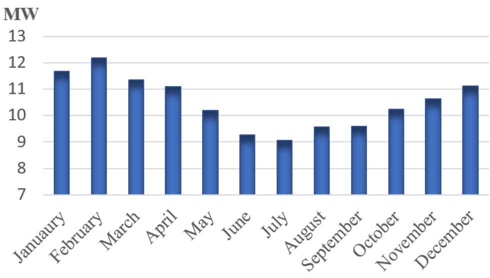


FIGURE 2. The average monthly load demand in MW of the university campus.

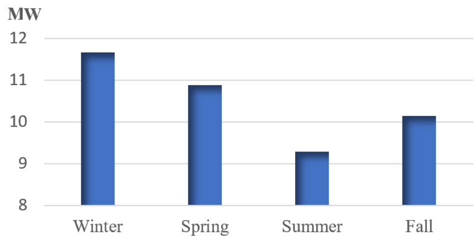


FIGURE 3. The average seasonal load demand in MW of the university campus.

The seasonal load demand of the university campus is shown in Fig. 3. The seasonal periods are defined as follows: 1) Spring (March to May); 2) Summer (June to August); 3) Fall (September to November); and 4) Winter (December-February).

B. THE PROPOSED LOAD FORECASTING APPROACH AND MODEL CREATION

In this paper, an effective load forecasting approach using regression models is proposed. The flowchart in Fig. 4 illustrates the needed steps. The procedure to implement the proposed approach is demonstrated in the following six steps:

- *Step 1: Data collection*. Two datasets need to be collected: one is historical load demand data; another is historical meteorological data.
- *Step 2: Regression models selection*. The regression models are used in the proposed method, and suitable regression models are selected for load forecasting. In this paper, 19 regression models are selected.
- *Step 3: Input parameters selection*. Important input parameters such as weather parameters are evaluated and selected.
- *Step 4: Regression models creation and load forecasting conducted using them*. The selected regression models in Step 2 will be trained and tested, and then will be used to conduct load forecasting.
- *Step 5: Comparison of the performance of regression models*. To compare the performance of the regression models, the forecasted load is compared with the actual measured load, and statistical error matrices are used to evaluate their accuracy.

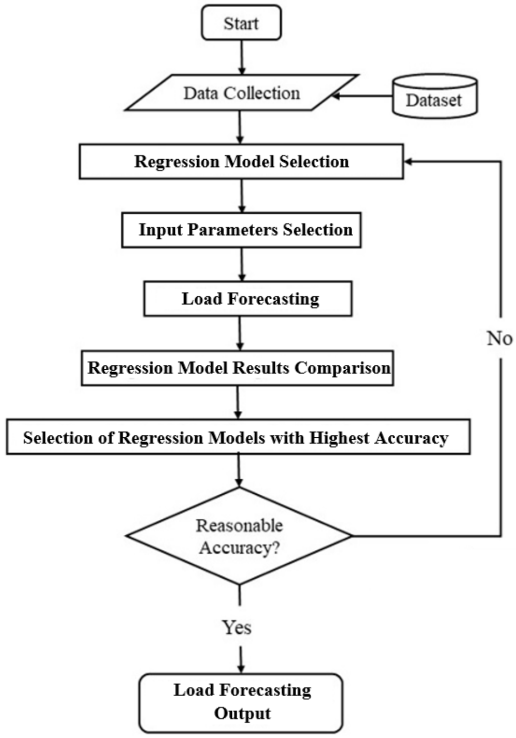


FIGURE 4. The flowchart of the proposed load forecasting method.

- Step 6: *Recommendation of regression models with the highest accuracy*. Based on previous steps, the regression models with the highest accuracy will be selected.

1) DATA COLLECTION AND PREPROCESSING

Among the six steps in the proposed approach, Step 1 on data collection and preprocessing is essential as input-output datasets are required to train the model. Data preprocessing has been a prerequisite for transforming the raw data so that the model can effectively learn the input-output relationship. In the pre-processing level, mathematical operations such as normalizing, ranking, and correlation are used [23]. The required data collection and preprocessing is shown in Fig. 5.

a: DATA COLLECTION

The initial step of datasets collection involves the weather data, time indicators and load demand data.

i) THE METEOROLOGICAL DATASET COLLECTION

A meteorological dataset is obtained from Newfoundland and Government of Canada website (https://climate.weather.gc.ca/). It is found that the meteorological data affect the load demand at the university campus, so their impact must be included in the load forecasting model [24]. The most significant independent variables for load forecasting are weather conditions. For domestic and agricultural customers, the weather effect is most common, and it can also change the user load profile. Load forecasting models use weather forecasting and other elements to predict the future load in order

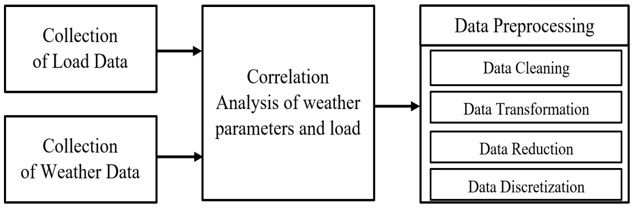


FIGURE 5. Data collection and preprocessing.

to minimize operational costs [24], [25]. In this study, the following eight weather parameters are collected for model creation:

1) Dry Bulb Temperature (DBT): the temperature of the air that is not subject to humidity or solar radiation.
2) Wet Bulb Temperature (WBT): thermometer reading where a damp cloth soaks the measuring instrument's bulb.
3) Humidity: Psychometrics from a Mollier chart is used to derive the air's relative humidity and the dew point from the DBT and WBT. The correlation to the electrical load can vary over the year because of the seasonal weather data changes.
4) Precipitation, fog, haze, or other exposure blockages, such as blowing snow or dust, may decrease visibility.
5) The wind chill: it is an index to show the average person how cold the weather conditions are. It is obtained by integrating temperature level and wind speed velocity into one number.
6) Wind direction (10's deg/tens of degrees): It reflects the average direction in two minute periods.
7) Wind speed (km/h): The airspeed in kilometers per hour (km/h) above 10 m from the ground.
8) Visibility in kilometers (km) It is the range in which a person can see and identify objects of the ideal size.

Although all weather parameters have a direct impact on the load demand, the temperature, humidity and visibility have greater impact on load variations than others.

ii) TIME INDICATOR

In short-term load forecasting, time is the critical factor to consider because its effect on the customer's load demand is highest. Some of the time indicators used in this study include: date, weekday, and time [24].

iii) LOAD PARAMETERS

The following load parameters are used for the load demand data in this study: previous day load in kW; previous day load in kVAr; previous day load in kVA; and previous day load power factor.

b: DATA CORRELATION

To study the correlation among the chosen weather parameters and the electrical load, the Pearson correlation analysis has been conducted. Pearson correlation analysis is a linear correlation analysis approach, which produces Pearson's correlation coefficients ranging from 0 to ±1, and a value close to 1 indicates a strong correlation. A positive sign indicates a proportional relationship, and a negative sign indicates an inversely proportional relationship. The outcome of the correlation analysis has been tabulated in Table 1.

TABLE 1. Pearson's correlation coefficient.

Parameters	Pearson's Correlation Coefficient
Peak load time	0.91
Off-Peak Load time	–0.9245
Day of the week	0.75
Previous day Load	0.93
Previous day Load	0.85
Previous day Load	0.91
Previous day power factor	0.7065
Dry Bulb Temperature	–0.88
Wet Bulb Temperature	–0.8
Humidity	–0.65
Wind chill	–0.63
Visibility	–0.68

The value of the correlation coefficients indicates that parameters, such as "Day of the week", "Previous day power factor", "Humidity", "Wind chill", and "Visibility", exhibit strong correlation with the load, while the rest of the parameters exhibits a very strong correlation with the load.

c: DATA PREPROCESSING

Real-life measurements are susceptible to various degrees of discrepancies including incomplete data, noise, missing values, outliers, redundant data and inappropriate formatting, which influence the performance of the regressors. Therefore, the data must be pre-processed to ensure the data reliability [26]. In this analysis, data preprocessing includes the following three sub-phases:

- *Data Cleaning*: It includes filling missing values, noise removal, outlier detection, and resolving discrepancies within the dataset [27]. Shape-Preserving Piecewise Cubic Spline Interpolation is used to fill columns with the partial missing weather data.
- *Data Transformation*: It involves various methods including integrating multiple files into a single usable format [27], and scaling the attribute to follow specific properties. After finding the data sets with corerelation and data cleaning, the final predictors dataset was created.
- *Data Reduction*: It aims to capture most of the data properties while removing redundancies by providing a reduced representation of the data, either by reducing the number of attributes or by sampling.

2) TRAINING AND TESTING REGRESSION MODELS

The training dataset is initially used to train a model, the results are then analyzed by varying its parameters until the most efficient parameters are obtained the MATLAB tool box provided the optimized model for all the regression models.

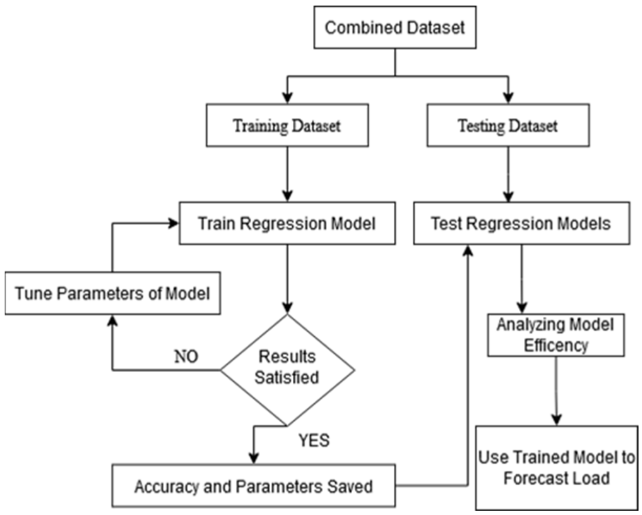


FIGURE 6. Training and testing regression models.

TABLE 2. Regression models used in this study.

Family of Regression Models	Chosen Regression Models
Linear Regression	Linear Regression model
	Interactions Regression model
	Robust Regression model
	Stepwise Linear Regression Model
Regression Trees	Medium Tree
	Coarse Tree
	Fine Tree
Support Vector Machines (SVM)	Linear SVM
	Quadratic SVM
	Cubic SVM
	Fine Gaussian SVM
	Medium Gaussian SVM
	Coarse Gaussian SVM
Gaussian Process Regression (GPR)	Rational Quadratic GPR
	Squared Exponential GPR
	Matern 5/2 GPR
	Exponential GPR
Ensemble of Trees	Boosted Trees
	Bagged Trees

After obtaining satisfied results, the trained model is used to test using testing datasets, various statistic error indices and evaluation plots are then used to evaluate the performance. The testing data are unseen by the trained model and are used to optimize the load forecasting model's control parameters, which helps to optimize and evaluate the performance of the created model.

K-Fold Cross-Validation: The initial data sample is randomly partitioned into k equivalent sized subsamples in k-fold cross-validation. A single subsample is retained from the k subsamples as recognition data for model testing, and the remaining k-1 subsamples are used for training. The cross-validation method is then repeated k times, with each of the k subsamples used as the validation data exactly once. To obtain a single estimate, the results of k can then be summed. In this paper, *five-fold cross-validation* is performed for all models to prevent the models from overfitting.

Fig. 6 shows the procedure of training and testing regression models.

III. REGRESSION MODELS

In this paper, five families of regression model algorithms provided in the MATLAB Regression Toolbox are selected to construct the short-term load forecasting model for the university campus. They are Linear Regression, Regression Trees, Support Vector Machines (SVM), Gaussian Process Regression (GPR), and Ensemble of Trees. Table 2 shows regression models used in this study.

Linear Regression: It is the simplest regression models used to forecast outcomes by modeling the relationship between the independent variable and the variable dependent. Since there are several independent and dependent variables, the model attempts to find the relationship between two or more explanatory variables and a response variable [28], [29].

Regression Trees: A single output (response) variable and multiple input (predictor) variables are used in all regression techniques. The variable output is numerical. The general regression tree structure method allows a combination of continuous and categorical variables to be input variables. A regression tree is built by a process called binary recursive partitioning. It is an iterative method that divides the data into partitions or branches, and then divides each section into smaller groups as each upper branch is approached. Initially, all training sets are grouped into the same section. In the first two partitions or branches, the algorithm starts assigning the data, using any possible binary split on every field. The algorithm chooses the division that minimizes the sum of the squared variances in the two different partitions from the mean. Each of the new branches is then added to this dividing guideline. This process continues until any node exceeds the user's minimum node size and becomes a terminal node [29].

Support Vector Machines: For nonlinear transformation, kernel functions in SVM are used. This study uses standard kernel functions, such as the linear kernel, the polynomial kernel, the Gaussian, or the radial basis function. Polynomial functions of the lower degree tend to underfit the model and do not have adequate results. With the polynomial increasing in degree, the curve is better suited [29].

Gaussian Process Regression: Gaussian process regression (GPR) models are nonparametric kernel-based probabilistic models with a finite set of random variables with a multivariate distribution. Any linear combination is distributed evenly. The Gaussian process theory is named after Carl Friedrich Gauss since it is based on the notion that an infinite-dimensional generalization of multivariate normal distributions is the Gaussian distribution. Gaussian processes are utilized in statistical modeling, regression to multiple target values, and analyzing mapping in higher dimensions [30].

Ensemble of Trees: An ensemble uses numerous algorithms to increase its accuracy in terms of efficiency and prediction. The Bagging or Bagged regression tree is a statistical classification-based regression technique designed to improve accuracy and stability of machine learning algorithms. Instead of using a single fit method, a linear combination of model fitting is created by constructing and integrating multiple predictors.

To find the best regression models for the short-term load forecasting, we performed: 1) Training a data set with cross-validation of 5 folds for all the models; 2) Plotting the behavior of regression Models with RMSE, R-Squared Value, MSE, MAE, and 3) Analyzing the results to see similarities and differences of the data [31].

A. DATASETS DESCRIPTION

In the proposed work, one-hour interval data from January 1, 2016 to March 31, 2020 are used as the simulation dataset. We eliminated columns with completely missing weather data and used Shape-Preserving Piecewise Cubic Spline Interpolation to fill columns with partial missing weather data.

1) VALIDATION DATA

First of all, out of the dataset from January 2016 to March 2020, we took out one-week data from each month from the 2016-2020 dataset in a pattern as follows: the 1st week from the 1st month, the 2nd week from the 2nd month, the 3rd week from the 3rd month, the 4th week from the 4th month, the 1st week from the 5th month, and the 2nd week from the 6th month etc. We use the above taken out data forming a new dataset, which serves as the validation dataset for the created model. The validation dataset consists of 8,976 rows and 15 columns.

2) TRAINING AND TESTING DATA

After the validation dataset is created, from the rest of the historical recorded dataset, 80% of the data is used for training, and the remaining 20% of data is used for testing. The training data for the regression models are with 5-fold cross-validation. The testing data are treated as unseen by the trained model and used to optimize the load forecasting model's control parameters, which helps to optimize and evaluate the performance of the model created.

The testing and training data can be viewed as a matrix with 28,428 rows and 15 columns. The rows represent each hour of a day from January 1, 2016 to March 31, 2020, excluding the validation dataset. The first 14 columns are the predictors or input, and the last column is the training target data, i.e., the load in kW.

3) PERFORMANCE EVALUATION DATASET

For performance evaluation after training and testing the model a random day (24 hours data) from validation dataset which falls in the month of March 2020, January 2020, October 2019, and July 2019 was selected representing each seasons of the year.

B. PERFORMANCE EVALUATION INDICES

The forecasted load is compared with the actual measured load for each regression model. By calculating three different statistical evaluations, the Root Mean Square Error (RMSE), the Mean Absolute Error (MAE), and the Mean Absolute Percentage Error (MAPE), the load forecasting capacity of each method and model accuracy can be assessed [26], [27]. If the value of RMSE > MAE, it means that there is a variation in errors.

1) MEAN ABSOLUTE ERROR (MAE)

The MAE measures the average magnitude of the errors, which can be calculated by

$$MAE = \frac{\sum_{t=1}^{n} \left|Y_t - \hat{Y}_t\right|}{n} \tag{1}$$

where $\hat{Y}_t$ is the prediction, Y_t is the true value from field recording, and n is the number of measurement points.

2) MEAN ABSOLUTE PERCENTAGE ERROR (MAPE)

This error percentage is a measure of the prediction accuracy of a forecasting method in statistics, it produces a measure of the relative overall fit, which can be calculated by

$$MAPE = \frac{\sum_{t=1}^{n} \frac{\left|Y_t - \hat{Y}_t\right|}{Y_t}}{n} \times 100 \tag{2}$$

where $\hat{Y}_t$ is the prediction, Y_t is the true value from field recording, and n is the number of measurement points.

3) ROOT MEAN SQUARE ERROR

The RMSE is the standard deviation of the residuals (prediction errors). Residuals are a measure of how far from the regression line data points are, so RMSE is a measure of how spread out these residuals are. It can be calculated as follows:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} (Y_t - \hat{Y}_t)^2}{n}} \tag{3}$$

where $\hat{Y}_t$ is the prediction, Y_t is the true value from field recording, and n is the number of measurement points.

4) R-SQUARED

R-Squared is a statistical measure of how close the fitted regression line is to the results. R-squared lies between 0 and 1. Generally, a higher R-squared value implies that the model matches the data better.

The following criteria is used to evaluate load forecasting performance using the error indices:

- The RMSE is always positive, and a smaller RMSE value indicates a good model.
- The R-squared lies between 0 and 1. R-Squared indicates a good model near 1.
- The MSE is the square of the RMSE, and a smaller MSE value indicates a successful model

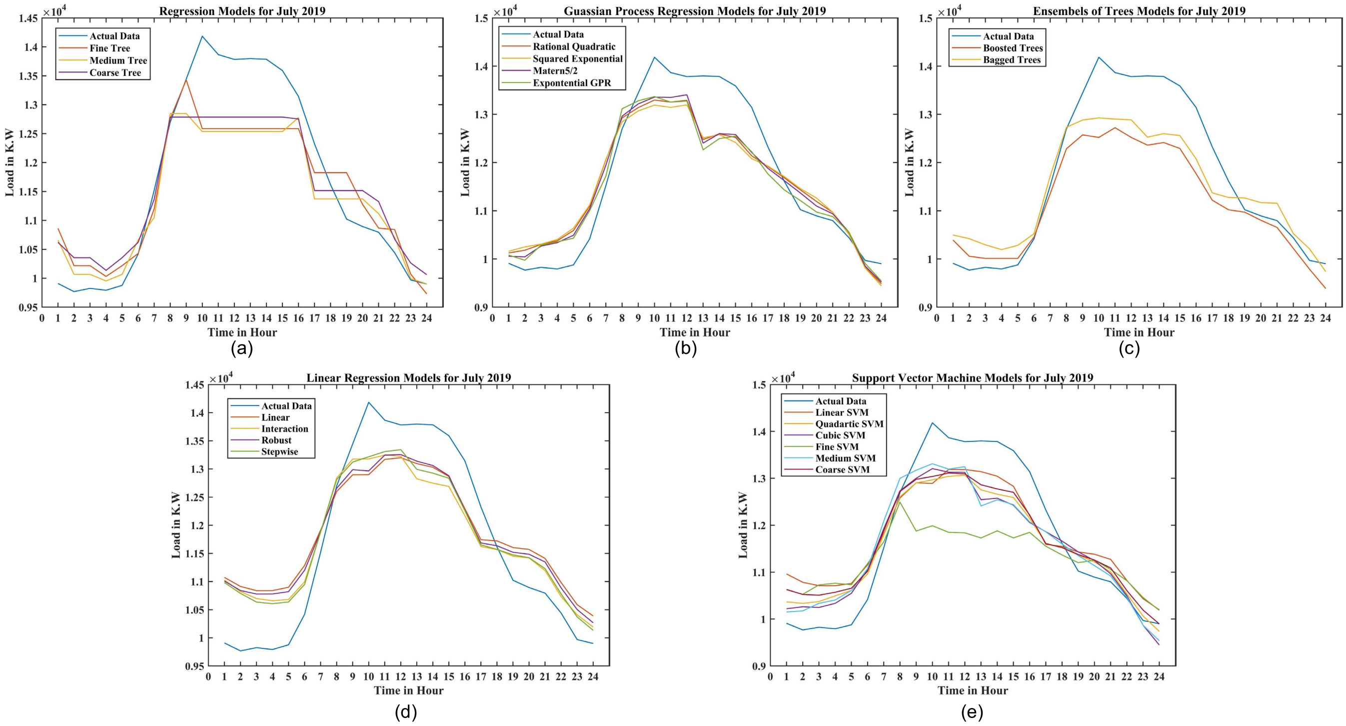


FIGURE 7. Forecasting results using five families of regression models for a day in July 2019: (a) regression tree models, (b) Gaussian Process Regression models, (c) Ensemble of Trees models, (d) Linear regression models, and (e) SVM models.

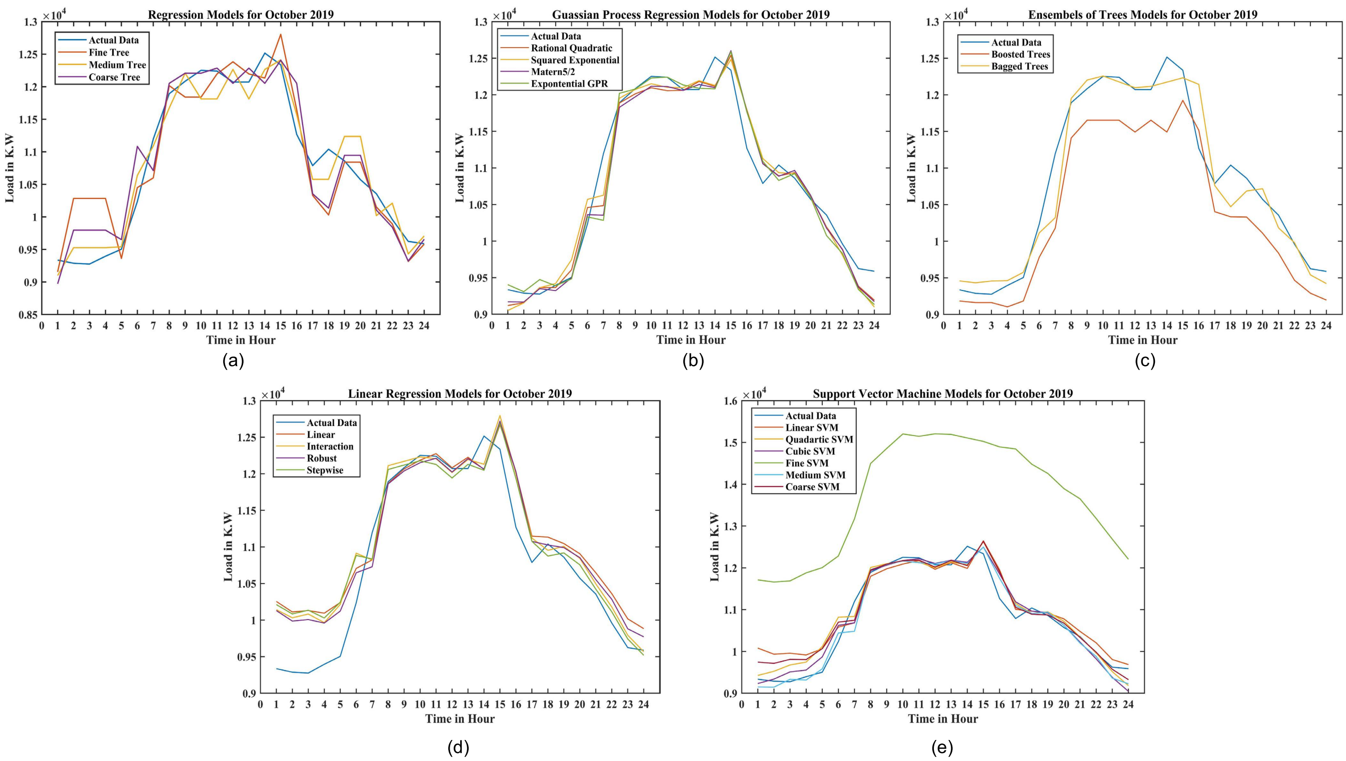


FIGURE 8. Forecasting results using five families of regression models for a day in October 2019: (a) regression tree models, (b) Gaussian Process Regression models, (c) Ensemble of Trees models, (d) Linear regression models, and (e) SVM models.

- The MAE is positive, similar to RMSE, a smaller MAE value suggests a successful model.
- An error percentage very close to zero means the predicted values are very relative to actual values.

C. PERFORMANCE EVALUATION PLOTS

1) PREDICTED VERSUS ACTUAL RESPONSE PLOTS

The plot is used to evaluate the trained model's performance, and it helps to understand how well the regression model

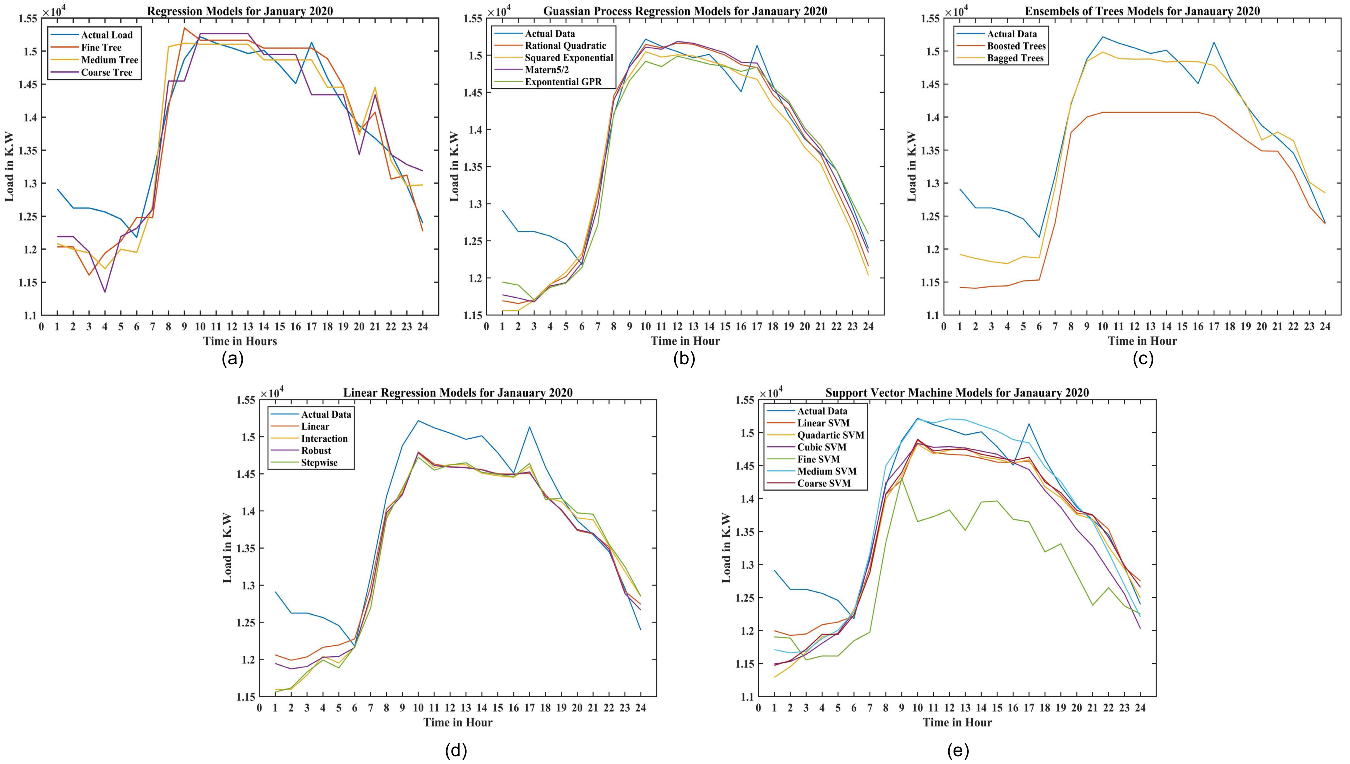


FIGURE 9. **Forecasting results using five families of regression models for a day in January 2020: (a) regression tree models, (b) Gaussian Process Regression models, (c) Ensemble of Trees models, (d) Linear regression models, and (e) SVM models.**

forecasts for different response values. The model's predicted response is plotted against the actual response. A perfect regression model has a predicted response equal to the true response, and all the points are in a diagonal line. The vertical range from the line to any point is the forecast error for that particular point. A good model has small errors, and the forecasts are scattered near the line. Typically, a great design has points spread roughly symmetrically around the diagonal line. If we can see any clear patterns in the plot, the design should be improved.

2) RESIDUAL PLOTS

For a model with good performance, it must satisfy: 1) Residuals unsymmetrically spread around zero; 2) Residuals' size alters considerably towards the right; and 3) Plots exhibit a nonlinear pattern clearly.

3) RESPONSE PLOTS

The response plot shows the predicted response versus the actual response. If the true response and predicted response of a model are identical, it indicates the model has good performance.

IV. SIMULATION RESULTS

The simulations are performed using MATLAB 2020 Regression Toolbox with the five-fold cross-validation. A day from July 2019, October 2019, January 2020, and March 2020 (4 days) from Performance Evaluation Dataset are selected to perform the regression model-based load forecasting. The outcome of the regression analysis are tabulated in Tables 3 and 4.

A. BENCHMARK MODEL

Support Vector Machine (SVM) or Support Vector Regressor (SVR) is a widely adopted regressor for developing short-term load forecasting models. Therefore, while proposing improved regressors for short-term load forecasting, SVR is mostly chosen as the benchmark model [19]. Similarly, SVM is chosen as the benchmark model in this paper.

B. SIX TOP-PERFORMANING MODELS

The comparisons of the 19 regression models-based forecasted load vs. actual load for the four chosen days (in July 2019, October 2019, January 2020 and March 2020) are shown in Figures 7 - 10. Based on the Forecasted vs. Actual Load pattern from Figure 7-10 it can be observed Rational Quadratic GPR, Exponential GPR, Matern 5/2 GPR, Squared Exponential GPR, Medium Gaussian SVM and Bagged Tree regression models where able to replicate the load more accurately.

Table 2 shows the performance evaluation through R-Squared, MSE, MAE, and RMSE values for the load forecasting using the 19 regression models. Table 3 shows the percentage errors of all 19 models.

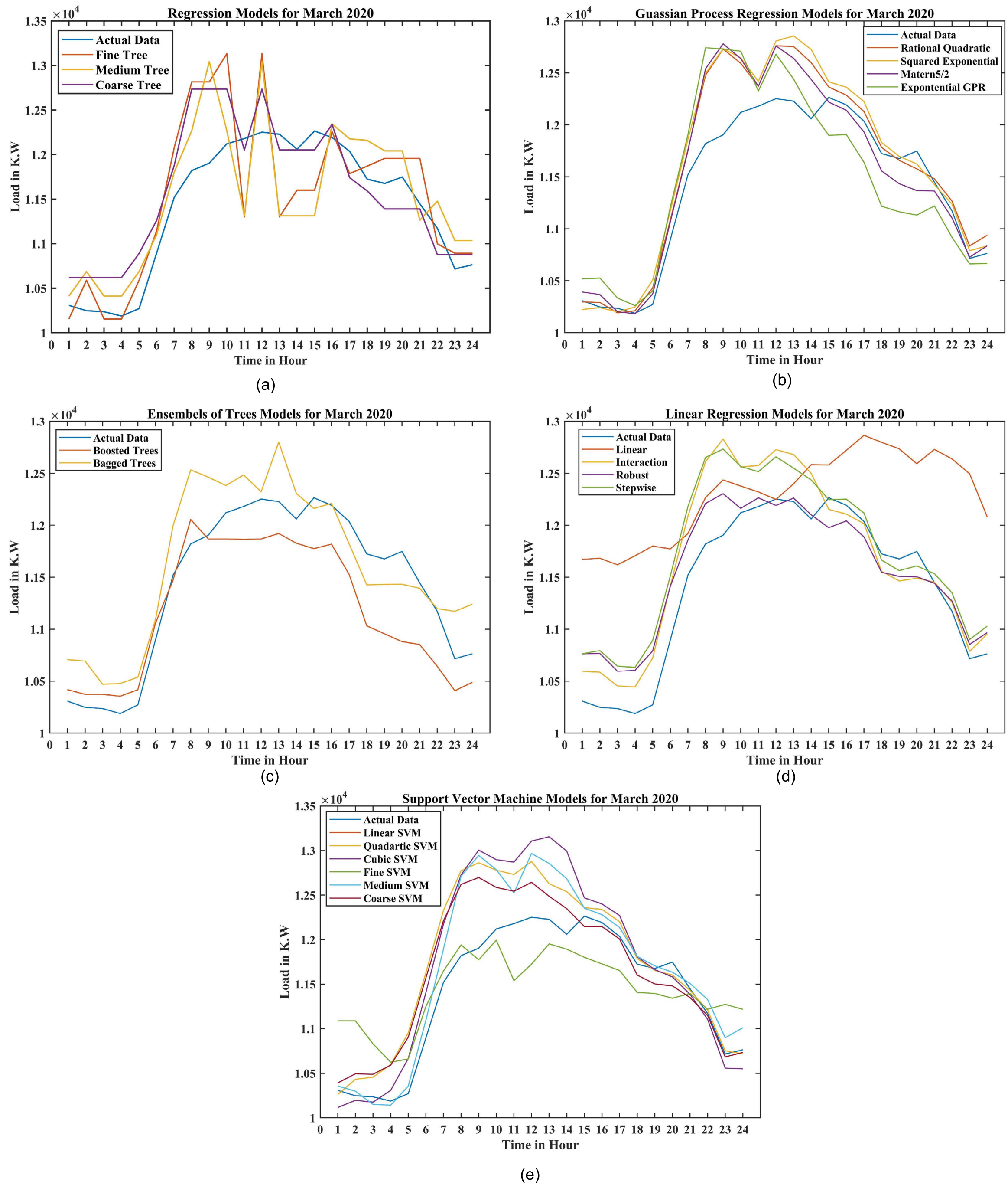


FIGURE 10. Forecasting results using five families of regression models for a day in March 2020: (a) regression tree models, (b) Gaussian Process Regression models, (c) Ensemble of Trees models, (d) Linear regression models, and (e) SVM models.

Based on the simulation results and performance evaluation indices among the 19 regression models, the six top-performing models are determined to be:

1) Rational Quadratic GPR,
2) Exponential GPR,
3) Matern 5/2 GPR,
4) Squared Exponential GPR,
5) Medium Gaussian SVM,
6) Bagged Tree.

It is found that all four GPR algorithms are chosen among the six-top performing modes. GPR models are nonparametric kernel-based probabilistic models, therefore, this type of probabilistic models work better than other models.

The selection of the six top-performing models are based on evaluation done using the following criteria:

- Evaluate RMSE, R-Squared, MSE, and MAE values of all the created models using the training and testing dataset.

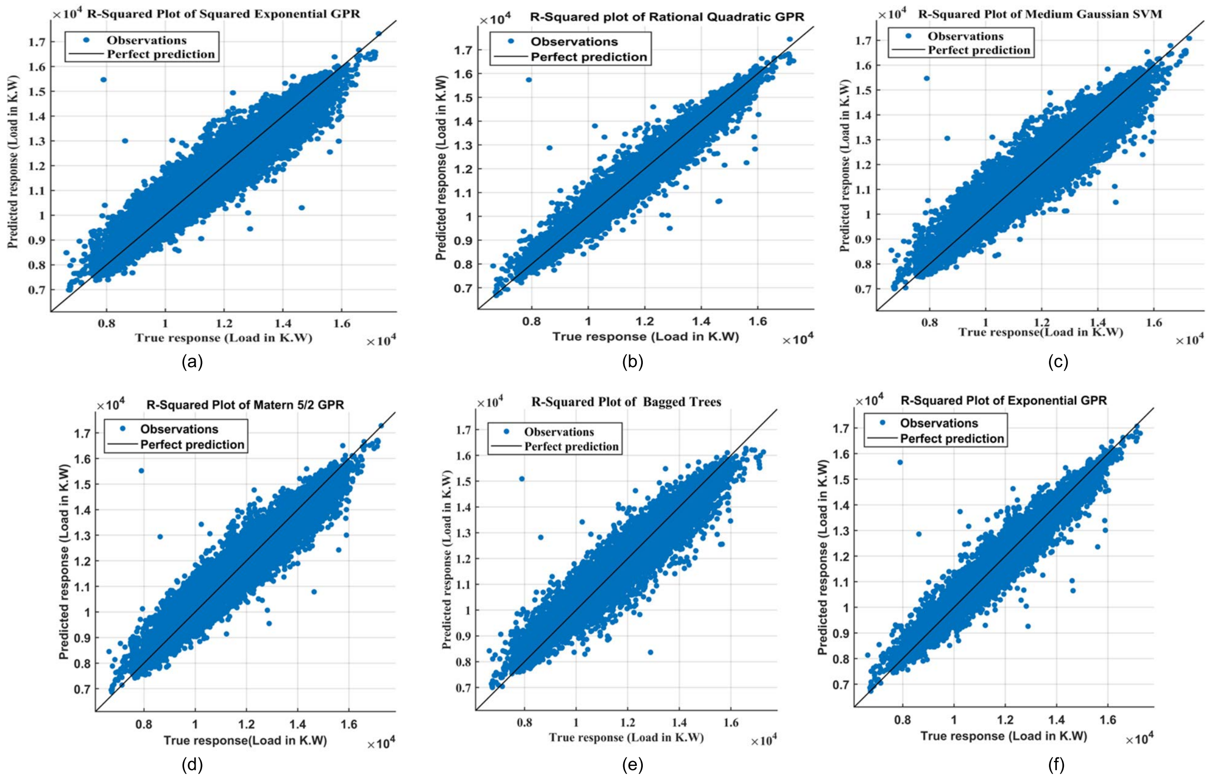


FIGURE 11. R-Squared plots for the six best performing models: (a) Squared Exponential GPR, (b) Rational Quadratic GPR, (c) Medium Gaussian SVM, (d) Matern 5/2 GPR, (e) Bagged Trees, and (f) Exponential GPR.

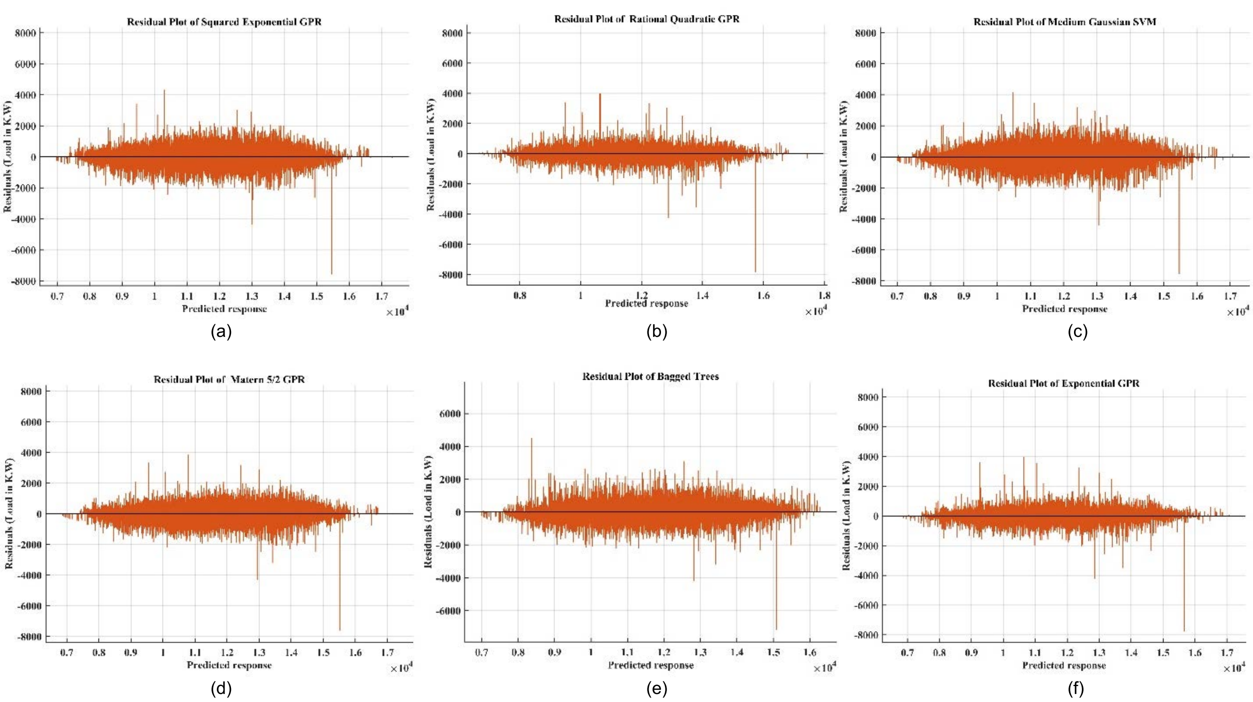


FIGURE 12. Predicted Response for the six best performing models: (a) Squared Exponential GPR, (b) Rational Quadratic GPR, (c) Medium Gaussian SVM, (d) Matern 5/2 GPR, (e) Bagged Trees, and (f) Exponential GPR.

- Analyze the MAPE of the Performance Evaluation Dataset to find the model with the least error percentage. This helps to analyze how much percentage error can occur when the load is forecasted using a particular model.
- Evaluate performance evaluation graphs of all created models using the training and testing dataset.

Performance evaluations for the top six performing models through R-squared plots, residual plots of predicted model

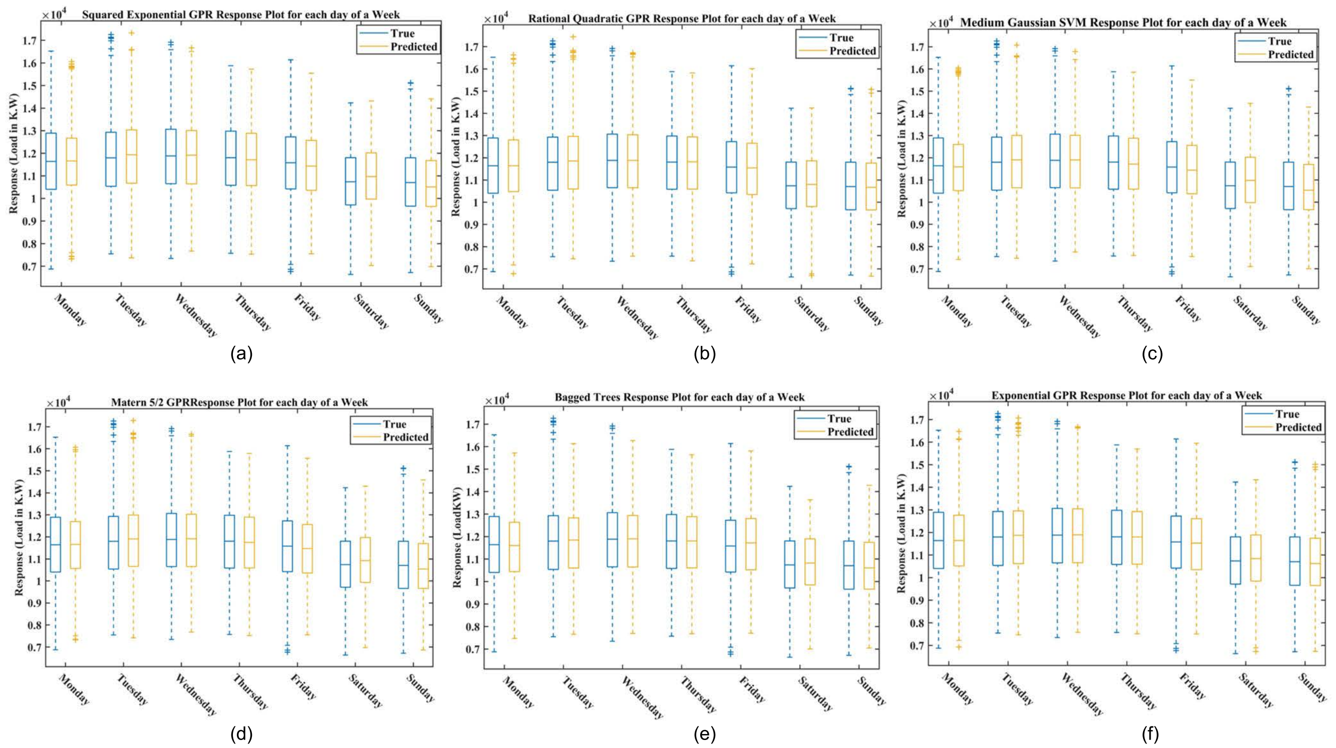


FIGURE 13. Response plots for each day in a week for the six best performing models: Squared Exponential GPR, Rational Quadratic GPR, Medium Gaussian SVM, Matern 5/2 GPR, Bagged Trees, and Exponential GPR.

TABLE 3. Load forecasting model performance evaluation.

Regression models	RMSE (kW)	R-Square	MSE (kW^2)	MAE (kW)
Rational Quadratic GPR	**301.76**	**0.97**	**91058**	**203.67**
Exponential GPR	**320**	**0.96**	**102480**	**221.11**
Bagged Tree	435.59	0.93	189740	314.43
Matern 5/2 GPR	437.44	0.93	191350	325
Medium Gaussian SVM	480.58	0.92	230960	348.78
Squared Exponential GPR	483.31	0.92	233590	362.47
Fine Tree	522.43	0.9	272940	365.35
Medium Tree	525.97	0.9	276650	383.88
Fine Gaussian SVM	559.16	0.89	312660	376.31
Coarse Tree	564.51	0.89	318670	419.54
Cubic SVM	544.74	0.89	296760	400.04
Quadratic SVM	616.72	0.87	382210	453.96
Coarse Gaussian SVM	637.8	0.86	406790	470.13
Interactions Regression	657.34	0.85	432090	491.67
Stepwise Linear Regression	661.02	0.85	436950	496.34
Boosted Tree	764.86	0.79	585010	605.14
Linear Regression	768.45	0.79	590520	588.56
Robust Regression	775.66	0.79	601640	576.06
Linear SVM	782.84	0.78	612840	579.3

TABLE 4. Load forecasting model evaluation using mean absolute percentage error (%).

Regression models	July 2019	October 2019	January 2020	March 2020	Average Error
Rational Quadratic GPR	**4.54**	**1.76**	**2.06**	**1.97**	**2.58**
Exponential GPR	**4.16**	**1.72**	**1.99**	**2.92**	**2.69**
Bagged Tree	4.90	1.76	2.16	2.64	2.86
Matern 5/2 GPR	4.12	1.77	2.01	2.03	2.48
Medium Gaussian SVM	4.60	1.67	2.14	2.53	2.73
Squared Exponential GPR	4.83	1.85	2.37	2.24	2.82
Fine Tree	4.80	3.40	2.39	3.79	3.59
Medium Tree	4.95	2.42	2.53	3.76	3.415
Fine Gaussian SVM	8.25	26.55	7.05	3.25	11.275
Coarse Tree	4.88	2.97	2.83	2.94	3.40
Cubic SVM	4.67	1.88	3.14	3.49	3.295
Quadratic SVM	4.78	2.20	2.86	3.09	3.23
Coarse Gaussian SVM	4.89	2.27	2.6	2.65	3.10
Interactions Regression	5.44	3.15	3.0	2.83	3.60
Stepwise Linear Regression	5.06	3.11	3.12	3.09	3.595
Boosted Tree	5.43	4.26	5.53	2.92	4.53
Linear regression	6.10	3.49	2.42	7.68	4.92
Robust regression	5.59	3.05	2.64	2.11	3.347
Linear SVM	5.50	2.822	2.32	3.09	3.433

and response plots of the trained model are shown in Figures 11-13.

C. RECOMMENDED TWO FINAL MODELS

Fig. 11 shows the R-Squared plots for the six top performing models. A good forecasting model is scattered tightly and symmetrically near the diagonal line. As shown in Fig. 11, although all six chosen models' forecasts are symmetrically scattered, scatter points for the Rational Quadratic GPR in Fig. 11 (b) and for Exponential GPR in Fig. 11 (f) are much more tightly placed around the diagonal line than other four models, which demonstrate improved performance.

Similar characteristics can be observed in the predicted response plot given in Fig. 12. Both Rational Quadratic GPR (Fig. 12 (b)) and Exponential GPR (Fig. 12 (f)) exhibits

TABLE 5. Hyperparameter setting of the proposed models.

	Exponential GPR		Rational Quadratic GPR		
Kernel Function	Exponential		Rational Quadratic		
Kernel Parameters	SigmaL	SigmaF	SigmaL	AlphaRQ	SigmaF
	35.9	3177.32	9.54	0.024	5103.01
Optimizer	Quasi-newton		Quasi-newton		
Basis Function	Constant		Constant		
Prediction Method	Block Coordinate Descent (BCD)		Block Coordinate Descent (BCD)		
BCD Block Size	1000		1000		
Number of Greedy BCD	100		100		
Step Tolerance	0.001		0.001		
BCD Iteration Limit	1×10^6		1×10^6		
Fit Method	Standard Deviation		Standard Deviation		
Active Set Method	Random		Random		
Random Search Set Size	59		59		
Tolerance	1×10^{-6}		1×10^{-6}		
Beta	1.11×10^4		1.065×10^4		
Sigma	16.86		4.28×10^2		

asymmetric distribution around the zero line more tightly than other four models, which indicate a better performance.

In Fig. 13, true and predicted response of the chosen six models are box-plotted. To be a good model, true and predicted response should be identical. Squared Exponential GPR exhibits dissimilarities on Monday, Tuesday and Saturday (Fig. 13 (a)), Medium Gaussian SVM on Monday and Saturday (Fig. 13 (c)), Matern 5/2 GPR on Saturday (Fig. 13 (d)), and Bagged Trees on Monday and Saturday (Fig. 13 (e)). However, true and predicted response of Rational Quadratic GPR (Fig. 13 (b) and Exponential GPR (Fig. 13 (f)) remain identical every day.

Comparing the six top-performing models, the analysis confirms that Rational Quadratic GPR and Exponential GPR algorithms are the two recommended final models. They are more accurate and reliable for predicting the university campus load demand throughout every season than other models.

The Rational Quadratic model showed excellent results in RMSE, R-Squared, MSE, and MAE values when it came to the validation dataset. The error of the predicted model was analyzed for four months out of all the six top-performing models, Exponential GPR can produce more accurate results with less error percentage.

Compared to other models, Rational Quadratic GPR and Exponential GPR was able to mimic the actual load pattern more effectively.

We have chosen SVM model as a benchmark in this paper, the two GPR models (Rational Quadratic GPR and Exponential GPR) are nonparametric kernel-based probabilistic models, and they have outperform the SVM models.

V. CONCLUSION

In this paper, an effective short-term load forecasting approach for the university campus at Memorial University of Newfoundland, St. John's, Canada is proposed. In the proposed approach, 19 regression models are firstly used to create load forecasting models using measured historical load demand data, these models are then evaluated through error indices. Among the 19 models, six top-performing models are determined.

The recommended final models are GPR models, which are nonparametric kernel-based probabilistic models. Through this research, it is found that the GPR is a viable load forecasting methodology. It is nonparametric, *i.e.,* it is not limited by a functional form, so rather than calculating the probability distribution of parameters of a specific function, GPR calculates the probability distribution over all admissible functions that fit the data. GPR is able to learn useful patterns through the utilization of the available training datasets and performs data extrapolation. A GPR implements a prediction model that is computationally inexpensive, while it is able to make predictions using small and scarce training datasets. In addition, it provides a predictive distribution defined by the mean value together with the respective variance. Therefore, Rational Quadratic GPR and Exponential GPR algorithms are recommended for load forecasting in this paper.

APPENDIX

Hyperparameters are internal parameters of a regression algorithm. A regressor set the coefficient values of a data driven regression model using its hyperparameters. Hyperparameter of the proposed GPR models are tabulated in the table below.

REFERENCES

[1] K. Chen, K. Chen, Q. Wang, Z. He, J. Hu, and J. He, "Short-term load forecasting with deep residual networks," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3943–3952, Jul. 2019.

[2] M. Gilanifar, H. Wang, L. M. K. Sriram, E. E. Ozguven, and R. Arghandeh, "Multitask Bayesian spatiotemporal Gaussian processes for short-term load forecasting," *IEEE Trans. Ind. Electron.*, vol. 67, no. 6, pp. 5132–5143, Jun. 2020.

[3] S. Khatoon and A. K. Singh, "Analysis and comparison of various methods available for load forecasting: An overview," in *Proc. Innov. Appl. Comput. Intell. Power, Controls Impact Humanity (CIPECH)*, Nov. 2014, pp. 243–247.

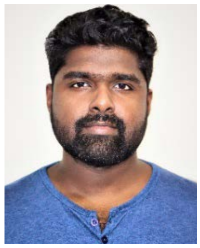
[4] A. E. Okoye and T. C. Madueme, "A theoretical framework for enhanced forecasting of electrical loads," *Int. J. Sci. Res. Publications*, vol. 6, no. 6, pp. 554–560, 2016.

[5] H. Jiang, Y. Zhang, E. Muljadi, J. J. Zhang, and D. W. Gao, "A short-term and high-resolution distribution system load forecasting approach using support vector regression with hybrid parameters optimization," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3341–3350, Jul. 2018.

[6] Z. Deng, B. Wang, Y. Xu, T. Xu, C. Liu, and Z. Zhu, "Multi-scale convolutional neural network with time-cognition for multi-step short-term load forecasting," *IEEE Access*, vol. 7, pp. 88058–88071, 2019.

[7] J. C. López, M. J. Rider, and Q. Wu, "Parsimonious short-term load forecasting for optimal operation planning of electrical distribution systems," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1427–1437, Mar. 2019.

[8] B. Liu, J. Nowotarski, T. Hong, and R. Weron, "Probabilistic load forecasting via quantile regression averaging on sister forecasts," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 730–737, Mar. 2017.
[9] C. Feng, M. Sun, and J. Zhang, "Reinforced deterministic and probabilistic load forecasting via *Q*-learning dynamic model selection," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1377–1386, Mar. 2020.
[10] Z. Cao, C. Wan, Z. Zhang, F. Li, and Y. Song, "Hybrid ensemble deep learning for deterministic and probabilistic low-voltage load forecasting," *IEEE Trans. Power Syst.*, vol. 35, no. 3, pp. 1881–1897, May 2020.
[11] B. Yildiz, J. I. Bilbao, and A. B. Sproul, "A review and analysis of regression and machine learning models on commercial building electricity load forecasting," *Renew. Sustain. Energy Rev.*, vol. 73, pp. 1104–1122, Jun. 2017, doi: 10.1016/j.rser.2017.02.023.
[12] S.-J. Baek and S.-G. Yoon, "Short-term load forecasting for campus building with small-scale loads by types using artificial neural network," in *Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. (ISGT)*, Feb. 2019, pp. 1–5, doi: 10.1109/ISGT.2019.8791674.
[13] D. Palchak, S. Suryanarayanan, and D. Zimmerle, "An artificial neural network in short-term electrical load forecasting of a university campus: A case study," *J. Energy Resour. Technol.*, vol. 135, no. 3, May 2013, Art. no. 032001, doi: 10.1115/1.4023741.
[14] Y. Kim, H.-G. Son, and S. Kim, "Short term electricity load forecasting for institutional buildings," *Energy Rep.*, vol. 5, pp. 1270–1280, Nov. 2019.
[15] S. Rai and M. De, "Analysis of classical and machine learning based short-term and mid-term load forecasting for smart grid," *Int. J. Sustain. Energy*, vol. 40, no. 9, pp. 821–839, Oct. 2021, doi: 10.1080/14786451.2021.1873339.
[16] J. Massana, C. Pous, L. Burgas, J. Melendez, and J. Colomer, "Short-term load forecasting in a non-residential building contrasting models and attributes," *Energy Buildings*, vol. 92, pp. 322–330, Apr. 2015, doi: 10.1016/j.enbuild.2015.02.007.
[17] J. Massana, C. Pous, L. Burgas, J. Melendez, and J. Colomer, "Short-term load forecasting for non-residential buildings contrasting artificial occupancy attributes," *Energy Buildings*, vol. 130, pp. 519–531, Oct. 2016, doi: 10.1016/j.enbuild.2016.08.081.
[18] M. Son, J. Moon, S. Jung, and E. Hwang, "A short-term load forecasting scheme based on auto-encoder and random forest," in *Applied Physics, System Science and Computers III*, Cham, Switzerland: Springer, 2019, pp. 138–144, doi: 10.1007/978-3-030-21507-1_21.
[19] A. Prakash, S. Xu, R. Rajagopal, and H. Noh, "Robust building energy load forecasting using physically-based kernel models," *Energies*, vol. 11, no. 4, p. 862, Apr. 2018, doi: 10.3390/EN11040862.
[20] *Gaussian Process Regression Models-MATLAB & Simulink*. Accessed: Jan. 4, 2022. [Online]. Available: https://www.mathworks.com/help/stats/gaussian-process-regression-models.html
[21] B.-J. Chen, M.-W. Chang, and C.-J. Lin, "Load forecasting using support vector machines: A study on EUNITE competition 2001," *IEEE Trans. Power Syst.*, vol. 19, no. 4, pp. 1821–1830, Nov. 2004.
[22] E. Ceperic, V. Ceperic, and A. Baric, "A strategy for short-term load forecasting by support vector regression machines," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4356–4364, Nov. 2013.
[23] M. D. Olagoke, A. A. Ayeni, and M. A. Hambali, "Short term electric load forecasting using neural network and genetic algorithm," *Int. J. Appl. Inf. Syst.*, vol. 10, no. 4, pp. 22–28, Jan. 2016.
[24] S. K. Sheikh and M. G. Unde, "Short term load forecasting using ANN technique," *Int. J. Eng. Sci. Emerg. Technol.*, vol. 1, no. 2, pp. 97–107, 2012.
[25] A. Y. Saber and T. Khandelwal, "IoT based online load forecasting," in *Proc. 9th Annu. IEEE Green Technol. Conf. (GreenTech)*, Mar. 2017, pp. 189–194.
[26] S. García, J. Luengo, and F. Herrera, *Data Preprocessing in Data Mining*, vol. 72. Cham, Switzerland: Springer, 2015. [Online]. Available: https://link.springer.com/book/10.1007/978-3-319-10247-4#about
[27] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*. Amsterdam, The Netherlands: Elsevier, 2011.
[28] M. Jawad, M. S. A. Nadeem, S.-O. Shim, I. R. Khan, A. Shaheen, N. Habib, L. Hussain, and W. Aziz, "Machine learning based cost effective electricity load forecasting model using correlated meteorological parameters," *IEEE Access*, vol. 8, pp. 146847–146864, 2020.
[29] M. S. Acharya, A. Armaan, and A. S. Antony, "A comparison of regression models for prediction of graduate admissions," in *Proc. Int. Conf. Comput. Intell. Data Sci. (ICCIDS)*, Feb. 2019, pp. 1–5.
[30] Y. K. Semero, J. Zhang, and D. Zheng, "PV power forecasting using an integrated GA-PSO-ANFIS approach and Gaussian process regression based feature selection strategy," *CSEE J. Power Energy Syst.*, vol. 4, no. 2, pp. 210–218, 2018.
[31] A. M. Pirbazari, E. Sharma, A. Chakravorty, W. Elmenreich, and C. Rong, "An ensemble approach for multi-step ahead energy forecasting of household communities," *IEEE Access*, vol. 9, pp. 36218–36240, 2021.



MITHUN MADHUKUMAR received the Bachelor of Technology degree in electrical and electronics engineering from Mahatma Gandhi University, Kerala, India, in 2017, and the Master of Applied Science degree in energy systems engineering from the Memorial University of Newfoundland, Canada, in 2021. From 2018 to 2019, he was with the Kerala State Electricity Board, a public energy utility agency. He has industrial experience in power systems, embedded systems, and automation through constructive professional collaboration with various public and private sectors.

ALBINO SEBASTIAN was born in 1995. He received the Bachelor of Technology degree in electrical and electronics engineering from Mahatma Gandhi University, Kerala, India, in 2018, where he was instrumental in developing an autonomous solar panel cleaning robot, and the Master of Applied Science degree in energy system engineering from the Memorial University of Newfoundland, in 2021. His interest in renewable energy motivated him to pursue the Master of Applied Science degree. He joined as an Intern with the Kerala State Electricity Board, in 2018, where he worked on HVDC, high-voltage transmission, switchgear, and instrumentation. His focus is on developing intelligent energy management system for industrial and commercial applications.

XIAODONG LIANG (Senior Member, IEEE) was born in Lingyuan, Liaoning, China. She received the B.Eng. and M.Eng. degrees from Shenyang Polytechnic University, Shenyang, China, in 1992 and 1995, respectively, the M.Sc. degree from the University of Saskatchewan, Saskatoon, Canada, in 2004, and the Ph.D. degree from the University of Alberta, Edmonton, Canada, in 2013, all in electrical engineering.

From 1995 to 1999, she served as a Lecturer with Northeastern University, Shenyang. In October 2001, she joined Schlumberger, Edmonton, and was promoted to a Principal Power Systems Engineer with this world's leading oil field service company, in 2009. After serving Schlumberger for almost 12 years, from 2013 to 2019, she was with Washington State University, Vancouver, WA, USA, and the Memorial University of Newfoundland, St. John's, NL, Canada, as an Assistant Professor and later an Associate Professor. She was an Adjunct Professor with Memorial University of Newfoundland from 2019 to 2022. In July 2019, she joined the University of Saskatchewan, where she is currently an Associate Professor and the Canada Research Chair in technology solutions for energy security in remote, northern, and indigenous communities. Her research interests include power systems, renewable energy, and electric machines.

Dr. Liang is a registered Professional Engineer in the province of Saskatchewan, Canada.

MOHSIN JAMIL (Senior Member, IEEE) received the Bachelor of Engineering degree in industrial electronics engineering form the NED University of Engineering and Technology, Karachi, Pakistan, in 2004, the M.Sc. degree in electrical engineering from Dalarna University, Sweden, in 2006, the M.Sc. degree in electrical engineering from the National University of Singapore, in 2008, with major in automation and control engineering, and the Ph.D. degree in electrical engineering from the University of Southampton, U.K., in 2012.

From 2012 to 2016, he was working with the Robotics Department, National University of Sciences and Technology (NUST), Islamabad, Pakistan. From 2016 to 2019, he was working with the Department of Electrical Engineering, Islamic University of Madinah, Saudi Arabia. He has been an Assistant Professor with the Department of Electrical and Computer Engineering, Memorial University of Newfoundland, Canada, since August 2019. He is the author or coauthor of several IEEE publications in different journals and peer-reviewed conferences.

Dr. Jamil was a recipient of different awards and funding grants. He is an Associate Editor of IEEE Access.

MD NASMUS SAKIB KHAN SHABBIR (Student Member, IEEE) was born in Tangail, Bangladesh. He received the B.Sc. degree in electrical and electronic engineering from the Rajshahi University of Engineering and Technology, Bangladesh, in 2012, and the Ph.D. degree in electrical engineering from the Memorial University of Newfoundland, Canada, in 2021.

From 2013 to 2016, he served as a Lecturer with the Department of Electrical and Electronic Engineering, Bangladesh University of Business and Technology, Bangladesh, and was promoted to an Assistant Professor, in 2016, and served until October 2017. He was a Visiting Research Student with the Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, Canada, from September 2019 to August 2020. He is currently working as a Postdoctoral Fellow with the University of Saskatchewan. His research interests include design and control of high-performance power electronic converters and their applications in renewable energy integration to power grids.